Faculty of Mechanical Engineering and Naval Architecture

Politecnico di Milano

Fanika Lukačević

A MODEL OF COGNITIVE LOAD FOR COMPUTER-AIDED DESIGN PERFORMANCE

INTERNATIONAL DUAL DOCTORAL THESIS

Zagreb, 2026

Faculty of Mechanical Engineering and Naval Architecture

Politecnico di Milano

Fanika Lukačević

A MODEL OF COGNITIVE LOAD FOR COMPUTER-AIDED DESIGN PERFORMANCE

INTERNATIONAL DUAL DOCTORAL THESIS

Mentors:
Assoc. Prof. Stanko Škec, PhD
Assoc. Prof. Niccolò Becattini, PhD

Zagreb, 2026

Fakultet strojarstva i brodogradnje

Politecnico di Milano

Fanika Lukačević

MODEL KOGNITIVNOGA OPTEREĆENJA ZA RAČUNALNO POTPOMOGNUTO KONSTRUIRANJE

MEĐUNARODNI DVOJNI DOKTORSKI RAD

Mentori:
Izv.prof.dr.sc. Stanko Škec
Izv.prof.dr.sc. Niccolò Becattini

Zagreb, 2026.

BIBLIOGRAPHY DATA

...

|

ACKNOWLEDGMENTS

I would like to express my gratitude to my supervisors, Assoc. Prof. Stanko Škec and Assoc. Prof. Niccolò Becattini, for their support and guidance throughout my PhD studies. I extend my sincere thanks to the members of the examination committee, Prof. Mario Štorga, Prof. Gaetano Cascini, Assoc. Prof. Kosa Goucher-Lambert, Assoc. Prof. Peter Törlind, and Assist. Prof. Tomislav Martinec, for reviewing this thesis and providing valuable and insightful comments. I would also like to thank Dr. Marija Majda Škec for her help with the statistical analyses.

I am sincerely grateful to all colleagues who generously gave their time and participated in the experimental studies.

I deeply appreciate the many people, collaborations, conversations, and everyday moments that shaped this period. To my dear CADLab and B22 third-floor colleagues and friends: thank you for sharing the joys, challenges, doubts, and small victories of PhD life. Special thanks go to Jelena for her understanding, for cheering me on, and for the laughter through every circle.

To my friends in and from NG, Zagreb, and Milan: thank you for always caring, listening, showing up, keeping me grounded, and making this journey lighter.

My deepest gratitude goes to my family for their continuous support, love, and encouragement, wherever I may be in life or in the world. Thank you for always believing in me. I would not be where I am today without you.

This page intentionally left blank

CONTENTS

BIBLIOGRAPHY DATA I
ACKNOWLEDGMENTS II
CONTENTS IV
ABSTRACT X
SOMMARIO XII
PROŠIRENI SAŽETAK XVI
LIST OF FIGURES XXII
LIST OF TABLES XXVIII
LIST OF ABBREVIATIONS AND SYMBOLS XXXIII
1 INTRODUCTION 1
1.1 Research motivation, aims, and hypothesis 1
1.2 Research methodology 5
1.2.1 Research clarification 6
1.2.2 Descriptive study I 7
1.2.3 Prescriptive study 7
1.2.4 Descriptive study II 8
1.2.5 Validation procedure 8
1.3 Model and method purpose and objectives 10
1.4 Scientific contribution 11
1.5 Thesis structure 12
2 RESEARCH BACKGROUND AND RELATED WORK 15
2.1 Literature review process 15
2.2 CAD performance 16
2.2.1 Studies on CAD performance 17
2.2.2 Describing, measuring, and visualising CAD performance 21

2.2.3 Conclusions of the literature review on CAD performance 27
2.3 Human performance modelling in complex HCI tasks 29
2.3.1 Cognitive architectures for modelling human performance in HCI tasks 31
2.3.2 Task analysis frameworks.. 37
2.3.3 Existing (studies on) cognitive models of CAD performance 43
2.3.4 Conclusions of the literature review on human performance modelling 48
2.4 Cognitive load (CL)... 50
2.4.1 Theoretical models of CL.. 51
2.4.2 Studies on CL in (engineering) design tasks 63
2.4.3 Measuring cognitive load 69
2.4.4 Conclusions of the literature review on CL... 76
3 THE FIRST EMPIRICAL STUDY 79
3.1 Study motivation, objectives, and research questions... 79
3.2 Methodology 82
3.2.1 Participants 83
3.2.2 Experimental procedure 84
3.2.3 CAD modelling tasks 85
3.2.4 Experimental setup 86
3.2.5 Collected experimental data 87
3.3 Differences in CAD outcomes and processes when modelling from isometric and orthographic projections.. 88
3.3.1 Data analysis.. 88
3.3.2 Results 90
3.3.3 Discussion 93
3.4 CAD modelling process segmentation and representation...................................... 95
3.4.1 Coding the scripts.. 97
3.4.2 Data analysis... 106

3.4.3 Results ... 107

3.4.4 Discussion ... 116

3.5 Differences in engineering designers' brain activity when CAD modelling from isometric and orthographic projections ... 126

3.5.1 EEG data pre-processing ... 127

3.5.2 Data analysis ... 129

3.5.3 Results ... 130

3.5.4 Discussion ... 137

3.6 Experimental study limitations ... 141

3.7 Conclusions ... 143

4 THEORETICAL MODELS ... 145

4.1 Key elements of CAD performance and CL in CAD activities ... 145

4.2 Theoretical model of engineering designers' performance in CAD ... 148

4.2.1 Perceptual, cognitive, motor, and CAD system operators ... 150

4.2.2 Hierarchical structure of CAD tasks ... 152

4.2.3 Unit task phases: acquisition, execution, inspection ... 158

4.2.4 Mapping P, C, and M operators to CAD actions, CAD operations, and unit task parts ... 162

4.3 Theoretical model of CL in CAD ... 164

5 A METHOD FOR MEASURING AND ANALYSING COGNITIVE LOAD IN CAD ACTIVITIES ... 171

5.1 Method assumptions, input requirements, and scope ... 172

5.2 Steps to model and analyse engineering designer's CAD performance ... 176

5.2.1 Analyse HCI to create a HCI script ... 177

5.2.2 Analyse the structure of a CAD task to create a timestamped hierarchical decomposition into subtask clusters and unit tasks ... 178

5.2.3 Map the motor, perceptual, and cognitive operators to the CAD events ... 180

5.2.4 Visualise the model of CAD performance ... 183

5.3 Steps to model and analyse dynamic changes in CL ... 183
5.3.1 Calculate utilisations of the operators ... 184
5.3.2 Identify EEG-based CL indicators applicable for the CAD context ... 185
5.3.3 Calculate information-processing weights ... 191
5.3.4 Calculate CL in task segments ... 194
5.3.5 Visualise the model of CL ... 195
6 THE SECOND EMPIRICAL STUDY ... 199
6.1 Study motivation, objectives, and research questions ... 200
6.2 Methodology ... 203
6.2.1 Participants ... 203
6.2.2 Experimental procedure ... 203
6.2.3 Experimental setup ... 204
6.2.4 Collected experimental data ... 205
6.2.5 CAD modelling tasks ... 205
6.2.6 NASA TLX questionnaire ... 206
6.2.7 Experimental task validation ... 207
6.2.8 EEG data pre-processing ... 209
6.3 Identifying EEG-based CL indicators for the CAD context ... 211
6.3.1 Data analysis ... 211
6.3.2 Results ... 213
6.3.3 Discussion ... 218
6.4 Assigning information processing weights to intermediate-level subtasks and unit task phases ... 222
6.4.1 Data analysis ... 222
6.4.2 Results ... 225
6.4.3 Discussion ... 227
6.5 Relating utilisation-based and subjective CL assessments ... 229

6.5.1 Data analysis ... 230
6.5.2 Results ... 232
6.5.3 Discussion ... 234
6.6 Comparing CL between unit task phases/subtasks and their contribution to the overall CL ... 236
6.6.1 Data analysis ... 237
6.6.2 Results ... 239
6.6.3 Discussion ... 246
6.7 Modelling the dynamic changes in CL throughout CAD modelling ... 248
6.7.1 Data analysis ... 249
6.7.2 Results ... 253
6.7.3 Discussion ... 254
6.8 Relating CAD performance assessments and CL ... 256
6.8.1 Data analysis ... 256
6.8.2 Results ... 256
6.8.3 Discussion ... 258
6.9 Experimental study limitations ... 260
6.10 Conclusions ... 261
7 VALIDATION AND DISCUSSION ... 267
7.1 Validating the theoretical models ... 267
7.1.1 Validating the theoretical model of CAD performance ... 267
7.1.2 Validating the theoretical model of CL in CAD activities ... 271
7.2 Validating the method for measuring and analysing CL in CAD activities ... 275
7.2.1 Theoretical structural validity ... 275
7.2.2 Empirical structural validity ... 279
7.2.3 Empirical performance validity ... 281
7.2.4 Theoretical performance validity ... 287

7.2.5 Consolidated method, advantages, and trade-offs 290
7.3 Confirming the research hypothesis 292
7.4 Limitations 294
8 CONCLUSIONS 297
REFERENCES 302
APPENDIX 318
Appendix A 318
Appendix B 336
Appendix C 369
BIOGRAPHY 370
ŽIVOTOPIS 371
BIBLIOGRAPHY 372

ABSTRACT

Computer-Aided Design (CAD) modelling is a common activity in contemporary engineering design, yet CAD performance is not formalised in a way that makes its key elements, influencing factors, and underlying mechanisms explicit. CAD modelling is also a cognitively complex human–computer interaction (HCI) activity that can impose high demands on limited information-processing resources, making cognitive load (CL) relevant both as an individual construct and in relation to CAD performance. However, CAD-specific CL indicators are not well established and reliable approaches for continuous CL measurement and interpretation in realistic CAD tasks remain limited. This thesis addresses these gaps by observing CAD modelling from the lens of information processing as the use of perceptual, cognitive, and motor resources to transform design problem information (inputs) into design representations (outputs), through a CAD modelling process that unfolds over time. The research aimed to develop and validate 1) a theoretical model of engineering designers' performance in CAD activities, 2) a theoretical model of CL in CAD activities, and 3) a method for measuring and analysing CL in the CAD context, while it hypothesised that psychophysiological monitoring enables assessment of CAD activity performance.

The findings from the literature review were corroborated with the first empirical study, where 20 engineering designers generated 3D CAD models from technical drawings with isometric or orthographic projections, while CAD performance was assessed through outcome- and process-based measures and EEG was recorded during modelling. The study was positioned as an initial empirical step to verify prerequisites for later model and method prescription; that a practically relevant influencing factor (representation format, operationalised through projection type) produces measurable differences in CAD performance outcomes and CAD modelling processes, that EEG recording during realistic CAD modelling is feasible and sufficiently sensitive to detect condition-related differences in brain activity, and that human performance (cognitive) modelling is feasible and informative for representing CAD modelling as a structured process through mapping observable CAD behaviour to perceptual, cognitive, and motor (P, C, M) operators. The findings confirmed representation format as an influencing factor reflected in both modelling outcomes and processes across multiple hierarchical goal/task levels and showed that brain activity is sensitive to projection type, implying that representation format can modulate CL. The study also established a hierarchical task decomposition supported by timeline visualisations, thereby providing a foundation for process-based analysis and subsequent theoretical modelling.

Building on these theoretical and empirical insights, the thesis developed complementary theoretical models of CAD performance and CL and prescribed a multimodal method integrating subjective CL assessment (NASA TLX), psychophysiological measurement (EEG), and utilisation-based human performance modelling, with defined data inputs and information flow for task- and segment-level analysis. The CAD performance model links observable interaction with a CAD system to P, C, and M operators and supports systematic comparison across engineering designers and CAD tasks. The CL model is anchored in the model of CAD performance. It quantifies CL through P and C operator utilisations and explicit equations linking these utilisations to WL and CL, and it introduces segment-specific weighting to represent differences in information-processing demands across unit task phases and intermediate subtasks. The method operationalised CAD modelling by decomposing the process into hierarchical task segments, mapping observable CAD behaviour to P, C, and M operators, computing indicators based on the utilisation of these operators and dynamic CL estimates, and relating CL indicators to process- and outcome-based CAD performance metrics. A second experimental study with 24 engineering designers then operationalised and tested the integrated method under varying internal task complexity. The study identified parietal and occipital alpha task-related power (TRP) as a prominent EEG-based CL indicator in the CAD context. Utilisation-based indicators provided process-level discrimination across unit task phases and intermediate subtasks, while dynamic alignment between EEG-derived and utilisation-derived CL timeseries indicated that these metrics capture overlapping but distinct aspects of CL. Regression analyses linking utilisation-based CL estimates to overall CAD modelling performance provided evidence supporting the research hypothesis; higher estimated CL was consistently associated with lower performance, and weighted utilisations showed the clearest and most robust relationship, indicating added explanatory value when operator utilisations are combined with EEG-derived information-processing weights. Based on the results of the second empirical study, the thesis also proposed a simplified utilisation-based method variant that preserves the ability to track CL changes across the modelling sequence with reduced experimental data and analysis requirements. Therefore, the thesis contributes two validated theoretical models (engineering designer's performance and CL in CAD activities) and an operational multimodal method for measuring and analysing CL in CAD modelling.

Keywords:

Computer-Aided Design; Cognitive load; Information processing; Human performance modelling; Electroencephalography

SOMMARIO

La modellazione Computer-Aided Design (CAD) è un'attività comune nella progettazione ingegneristica contemporanea; tuttavia, la stima della prestazione relativa alla modellazione CAD non è formalizzata in modo tale da rendere espliciti i suoi elementi chiave, i fattori che la influenzano e i meccanismi sottostanti. La modellazione CAD è inoltre un'attività di interazione uomo-computer (Human–Computer Interaction, HCI) cognitivamente complessa, che può imporre elevate richieste a risorse limitate di elaborazione delle informazioni, rendendo il carico cognitivo (Cognitive Load, CL) rilevante sia come costrutto individuale sia in relazione alla prestazione della modellazione CAD. Tuttavia, gli indicatori di CL specifici per la modellazione CAD non sono ancora ben consolidati e gli approcci affidabili per la misurazione e l'interpretazione continua del CL in compiti CAD realistici rimangono limitati. Questa tesi affronta queste lacune osservando la modellazione CAD dalla prospettiva dell'elaborazione delle informazioni, intesa come utilizzo di risorse percettive, cognitive e motorie per trasformare le informazioni del problema progettuale, ossia gli input, in rappresentazioni di sistemi tecnici, ossia gli output, attraverso un processo di modellazione CAD che si sviluppa nel tempo. La ricerca ha avuto l'obiettivo di sviluppare e validare: 1) un modello teorico della prestazione dei progettisti ingegneristici nelle attività CAD, 2) un modello teorico del CL nelle attività CAD e 3) un metodo per misurare e analizzare il CL nel contesto CAD, ipotizzando che il monitoraggio psicofisiologico consenta la valutazione della prestazione nelle attività CAD.

I risultati emersi dalla revisione della letteratura sono stati corroborati dal primo studio empirico, nel quale 20 ingegneri progettisti hanno generato modelli CAD 3D a partire da disegni tecnici con proiezioni isometriche o ortografiche, mentre la prestazione in CAD è stata valutata mediante misure basate sia sull'output sia sul processo anche attraverso i dati da elettroencefalografia (EEG) registrati durante la modellazione. Lo studio è stato concepito come un primo passo empirico volto a verificare i prerequisiti per la successiva prescrizione del modello e del metodo:

- che un fattore di influenza rilevante dal punto di vista pratico, ossia il formato di rappresentazione, declinato secondo diversi metodi di proiezione, produca differenze misurabili negli esiti della prestazione CAD e nei processi di modellazione CAD;
- che la registrazione EEG durante una modellazione CAD realistica sia fattibile e sufficientemente sensibile da rilevare differenze nell'attività cerebrale legate alla condizione sperimentale; e

- che la modellazione della prestazione umana, di tipo cognitivo, sia fattibile e informativa per rappresentare la modellazione CAD come un processo strutturato attraverso la mappatura del comportamento CAD osservabile su operatori percettivi, cognitivi e motori (P, C, M).

I risultati hanno confermato il formato di rappresentazione come fattore di influenza, con effetti visibili sia negli esiti sia nei processi di modellazione attraverso molteplici livelli gerarchici di obiettivi e compiti. Hanno, inoltre, mostrato che l'attività cerebrale è sensibile al tipo di proiezione, suggerendo che il formato di rappresentazione possa modulare il CL. Lo studio ha inoltre stabilito una decomposizione gerarchica del compito supportata da visualizzazioni temporali, fornendo così una base per l'analisi basata sul processo e per la successiva modellazione teorica.

Sulla base di queste evidenze teoriche ed empiriche, la tesi ha sviluppato modelli teorici complementari della prestazione CAD e del CL, e ha prescritto un metodo multimodale che integra la valutazione soggettiva del CL (attraverso NASA TLX), la misurazione psicofisiologica (attraverso EEG) e la modellazione della prestazione umana basata sull'utilizzazione delle risorse, con input di dati e flussi informativi definiti per l'analisi a livello di compito e di segmento. Il modello di prestazione CAD collega l'interazione osservabile con un sistema CAD agli operatori P, C e M, e supporta un confronto sistematico tra progettisti ingegneristici e compiti CAD. Il modello di CL è collegato direttamente al modello di prestazione CAD. Esso

- quantifica il CL attraverso le utilizzazioni degli operatori P e C e mediante equazioni esplicite che collegano tali utilizzazioni al carico di lavoro (Workload, WL) e al CL; inoltre,
- introduce una pesatura specifica per segmento per rappresentare le differenze nelle richieste di elaborazione delle informazioni tra le fasi dei compiti unitari e i sottocompiti intermedi.

Il metodo ha caratterizzato la modellazione CAD decomponendo il processo in segmenti gerarchici inerenti le attività svolte, mappando il comportamento CAD osservabile sugli operatori P, C e M, calcolando indicatori basati sull'utilizzo degli stessi operatori e stime dinamiche del CL, e mettendo in relazione gli indicatori di CL con misure della prestazione CAD basate sia sul processo sia sull'output.

Un secondo studio sperimentale con 24 progettisti ingegneristici ha quindi applicato e testato il metodo integrato in condizioni di diversa complessità interna del compito. Lo studio ha identificato la Task-Related Power (TRP) nella banda alfa nelle regioni parietale e occipitale come un indicatore EEG prominente del CL nel contesto CAD. Gli indicatori basati sull'utilizzo degli operatori PCM hanno consentito di discriminare compiti unitari e i sottocompiti intermedi a livello di processo, mentre l'allineamento dinamico tra le serie temporali di CL derivate dall'EEG e quelle derivate dall'utilizzazione ha indicato che tali metriche catturano aspetti sovrapposti ma distinti del CL. Le analisi di regressione che collegano le stime del CL basate sull'utilizzo di operatori PCM alla prestazione complessiva nella modellazione CAD hanno fornito evidenze a supporto dell'ipotesi di ricerca: un CL stimato più elevato è risultato costantemente associato a una prestazione inferiore, e le utilizzazioni pesate hanno mostrato la relazione più chiara e robusta, indicando un valore esplicativo aggiuntivo quando l'utilizzo degli operatori PCM viene combinato con pesi di elaborazione delle informazioni derivati dall'EEG. Sulla base dei risultati del secondo studio empirico, la tesi ha inoltre proposto una variante semplificata del metodo basato sull'utilizzo di operatori PCM, che preserva la capacità di tracciare i cambiamenti del CL lungo la sequenza di modellazione con requisiti ridotti in termini di dati sperimentali e di analisi. Pertanto, la tesi contribuisce con due modelli teorici validati, relativi rispettivamente alla prestazione di ingegneri progettisti e al CL nelle attività CAD, e con un metodo multimodale operativo per misurare e analizzare il CL nella modellazione CAD.

Parole chiave:

Computer-Aided Design; Carico cognitivo; Elaborazione delle informazioni; Modellazione della prestazione umana; Elettroencefalografia.

This page intentionally left blank

PROŠIRENI SAŽETAK

CAD modeliranje jedna je od temeljnih aktivnosti suvremenog konstruiranja tijekom koje konstruktori kontinuirano međudjeluju s CAD sustavima. Ovi sustavi obuhvaćaju CAD okruženje i alate za interakciju čovjeka i računala, a konstruktori ih koriste kako bi stvarali, mijenjali, pregledavali i vrednovali digitalne prikaze tehničkih sustava. Karakteristike CAD sustava utječu na dostupnost informacija, način njihova prikaza, slijed mogućih radnji i odluke koje konstruktor mora donositi tijekom procesa modeliranja. Zbog toga CAD sustavi imaju važnu ulogu u suvremenom razvoju proizvoda te izravno utječu na uspješnost rada konstruktora, a time i na kvalitetu, troškove i trajanje razvojnih procesa.

Unatoč važnosti CAD sustava i aktivnosti u inženjerskoj praksi, konstruktorova izvedba u CAD okruženju još uvijek nije formalizirana na način koji bi jasno prikazao njezine ključne elemente, odnose među tim elementima i čimbenike koji utječu na nju. Postojeća istraživanja često su usmjerena na kvalitetu konačnoga CAD modela ili druge izlazne pokazatelje, dok se manja pozornost posvećuje samom procesu kojim se do toga rezultata dolazi, Time ostaje nedovoljno objašnjeno kako se CAD modeliranje odvija tijekom vremena, koje perceptivne, kognitivne i motoričke zahtjeve ono postavlja pred konstruktore te kako se ti zahtjevi povezuju s ishodom CAD aktivnosti. Promatranje CAD modeliranja iz perspektive obrade informacija te interakcije čovjeka i računala omogućuje prikaz ponašanja konstruktora kao strukturiranoga procesa, a tumačenje uspješnosti njegove izvedbe kao korištenje ograničenih kognitivnih resursa.

U ovom se istraživanju konstruktorova izvedba u CAD aktivnostima promatra kao korištenje perceptivnih, kognitivnih i motoričkih resursa za transformaciju ulaznih informacija (zahtjeva i opisa konstrukcijskoga problema) u izlazne informacije (digitalne prikaze tehničkoga sustava). Takav pogled temelji se na pretpostavci da se CAD modeliranje ne može u potpunosti razumjeti samo analizom konačnoga modela, nego je potrebno obuhvatiti i proces međudjelovanja s CAD sustavom. Tijekom modeliranja konstruktor percipira informacije prikazane u tehničkim crtežima, CAD sučelju i 3D modelu, kognitivno ih obrađuje, donosi odluke o daljnjim koracima te provodi motoričke radnje pomoću alata za međudjelovanje s CAD okruženjem. Stoga je za razumijevanje konstruktorove izvedbe u CAD aktivnostima važno modelirati način na koji se ti resursi koriste i raspoređuju tijekom rješavanja zadatka.

Potreba za takvim pristupom posebno je izražena zbog vizualne, virtualne i kognitivno zahtjevne prirode CAD aktivnosti. Konstruktori tijekom CAD modeliranja istodobno prate više izvora informacija, tumače geometrijske odnose, koriste sposobnosti prostorne vizualizacije i mentalne rotacije te upravljaju nizom naredbi i mogućih postupaka u CAD sustavu. Budući da je isti cilj u CAD okruženju često moguće postići na različite načine, konstruktor mora neprestano procjenjivati mogućnosti, donositi odluke i prilagođavati svoj postupak. Takve aktivnosti zahtijevaju značajne ljudske informacijsko-procesne resurse koji su ograničeni, zbog čega kognitivno opterećenje postaje važan čimbenik za razumijevanje konstruktorove izvedbe u CAD aktivnostima.

Kognitivno opterećenje u ovom se istraživanju promatra kao trošak korištenja dostupnih perceptivnih i kognitivnih resursa pri ostvarivanju ciljeva CAD aktivnosti. Riječ je o višedimenzionalnome konstruktu koji proizlazi iz ograničenoga kapaciteta ljudskoga kognitivnoga sustava za obradu informacija. U kontekstu konstruiranja, razumijevanje kognitivnoga opterećenja važno je jer može pridonijeti razvoju metoda i alata koji podržavaju kognitivne procese konstruktora, smanjuju nepotrebno procesuiranje te poboljšavaju učinkovitost međudjelovanja s CAD sustavima. U tom smislu, bolje razumijevanje kognitivnog opterećenja može omogućiti razvoj naprednijih CAD okruženja, primjerice pravovremenim pružanjem potrebnih informacija, prilagodbom prikaza, automatizacijom pojedinih koraka ili preraspodjelom zadataka između čovjeka i računala.

Iako se kognitivno opterećenje često spominje kao moguće objašnjenje razlika u težini CAD zadataka ili uspješnosti njihove izvedbe, još uvijek nedostaju pokazatelji prilagođeni CAD-u koji bi omogućili pouzdano i kontinuirano mjerenje promjena kognitivnog opterećenja tijekom izvedbe realističnih CAD aktivnosti. Poseban izazov proizlazi iz toga što je kognitivno opterećenje konstrukt koji se ne može izravno opažati te njegovo mjerenje tijekom složenih i dugotrajnih aktivnosti, poput CAD modeliranja, zahtijeva metode koje ne ometaju izvedbu zadatka. Dosadašnja istraživanja najčešće su se oslanjala na subjektivne metode procjene, primjerice upitnike kao što je NASA TLX, dok su fiziološke i psihofiziološke metode, uključujući elektroencefalografiju (EEG), rjeđe primjenjivane u realističnim CAD uvjetima. Istodobno, istraživanja CAD izvedbe i istraživanja kognitivnog opterećenja često su se razvijala odvojeno, zbog čega je odnos između kognitivnog opterećenja i uspješnosti konstruktorove izvedbe u CAD aktivnostima i dalje nedovoljno istražen.

Polazeći od navedenih nedostataka, cilj ovoga doktorskoga rada jest razviti i validirati teorijske modele i metodu koji omogućuju razumijevanje, mjerenje i analizu kognitivnog opterećenja u CAD aktivnostima te njegovo povezivanje s uspješnošću konstruktorove izvedbe. Istraživanje je usmjereno na tri međusobno povezana cilja. Prvi je cilj razviti i validirati teorijski model konstruktorove izvedbe u CAD aktivnostima. Drugi je cilj razviti i validirati teorijski model kognitivnog opterećenja u CAD aktivnostima. Treći je cilj razviti i validirati metodu za mjerenje i analizu kognitivnog opterećenja u CAD kontekstu. Istraživanje se pritom temelji na hipotezi da metoda mjerenja i analize kognitivnog opterećenja, zasnovana na praćenju psihofizioloških promjena, omogućuje procjenu razine uspješnosti konstruktorove izvedbe u CAD aktivnostima. Svrha doktorskoga rada jest razviti podršku u obliku metode koja omogućuje bolje razumijevanje i procjenu uspješnosti konstruktorove izvedbe u CAD aktivnostima modeliranjem, mjerenjem i analizom kognitivnog opterećenja. Budući da je postojeće razumijevanje kognitivnog opterećenja i konstruktorove izvedbe u ovom specifičnom kontekstu još uvijek ograničeno, razvoj metode zahtijevao je prethodnu teorijsku razradu ključnih konstrukata, odnosa među njima i čimbenika koji na njih utječu. Istraživanje je metodološki utemeljeno na metodologiji *Design Research Methodology* (DRM), uz primjenu načela eksperimentalnog istraživanja u konstruiranju (eng. *Experimental Design Research*) te validacijskog okvira *Validation Square*. Istraživačke aktivnosti organizirane su u nekoliko glavnih faza: razjašnjavanje istraživanja, deskriptivno istraživanje I, preskriptivno istraživanje, deskriptivno istraživanje II te završna validacija i predlaganje pojednostavljene inačice metode.

U fazi razjašnjavanja istraživanja provedeni su pregledi literature o konstruktorovoj izvedbi i kognitivnom opterećenju u CAD aktivnostima. Cilj ove faze bio je uspostaviti teorijsku osnovu za razumijevanje promatranih pojava, identificirati istraživačke praznine te definirati ciljeve, hipotezu, svrhu i očekivane doprinose istraživanja. Pregled literature obuhvatio je relevantne teorije, modele i pojmove iz područja CAD modeliranja, ljudske izvedbe, kognitivne arhitekture, interakcije čovjeka i računala te mjerenja kognitivnog opterećenja. Na temelju toga pregleda određene su dimenzije kojima se konstruktorova izvedba i kognitivno opterećenje mogu konceptualizirati, kvantificirati i analizirati u CAD aktivnostima.
Zaključno, pregled literature pokazao je da su konstruktorova izvedba i kognitivno opterećenje međusobno povezani, ali su dosad najčešće istraživani odvojeno. Konstruktorova izvedba u CAD aktivnostima često se mjeri ocjenom kvalitete konačnoga modela i kumulativnim procesnim pokazateljima, dok se kognitivno opterećenje najčešće procjenjuje subjektivno ili psihofiziološki, ali u zadacima koji nisu izravno povezani s realističnim CAD modeliranjem. Nedostaju istraživanja koja istodobno promatraju CAD proces, CAD ishod i vremenski promjenjivo kognitivno opterećenje. Također nedostaju formalni modeli koji bi povezali opažljive CAD radnje sa zahtjevima na perceptivne, kognitivne i motoričke resurse te omogućili tumačenje promjena kognitivnog opterećenja tijekom modeliranja. Na temelju pregleda literature definirana je istraživačka potreba za razvojem teorijskog modela konstruktorove izvedbe u CAD aktivnostima, teorijskog modela kognitivnog opterećenja u CAD aktivnostima i metode koja će omogućiti njihovu zajedničku operacionalizaciju.

Deskriptivno istraživanje I proširilo je spoznaje dobivene pregledom literature te dodatno razradilo ključne elemente konstruktorove CAD izvedbe i kognitivnog opterećenja, njihove karakteristike i međusobne odnose. Pregled literature istaknuo je format prikaza inženjerskih informacija kao relevantan čimbenik koji može utjecati na ishode i procese CAD modeliranja, ali i na kognitivno opterećenje načinom na koji usmjerava percepciju, vizualnu pretragu, odlučivanje i izbor CAD radnji. Stoga je prva empirijska studija, provedena u ovoj fazi, usmjerena na ispitivanje utjecaja formata prikaza, operacionaliziranoga tipom projekcije u tehničkom crtežu, na konstruktorovu izvedbu i njegovu moždanu aktivnost tijekom modeliranja. Ovom su studijom provjereni važni preduvjeti za kasniji razvoj modela i metode. U studiji je ispitano utječe li format prikaza inženjerskih informacija, odnosno tip projekcije u tehničkom crtežu, na ishode i procese CAD modeliranja. Također je ispitana prikladnost elektroencefalografije (EEG-a) za kontinuirano bilježenje moždane aktivnosti tijekom realističnog CAD modeliranja te prikladnost metode modeliranja ljudske izvedbe (eng. *Human Performance Modelling*) za prikaz CAD aktivnosti perceptivnim, kognitivnim i motoričkim operatorima. Rezultati prve empirijske studije potvrdili su relevantnost formata reprezentacije kao utjecajnog čimbenika, pokazali izvedivost praćenja moždane aktivnosti tijekom CAD modeliranja korištenjem EEG-a i omogućili izradu početnih modela konstruktorove izvedbe u CAD aktivnostima temeljenih na empirijskim podacima.

Na temelju teorijskih i empirijskih spoznaja u preskriptivnoj su studiji razvijeni teorijski model konstruktorove izvedbe i teorijski model kognitivnog opterećenja u CAD aktivnostima. Model konstruktorove izvedbe u CAD aktivnostima povezuje opažljivo međudjelovanje konstruktora i CAD sustava s perceptivnim, kognitivnim i motoričkim operatorima te omogućuje sustavnu usporedbu procesa modeliranja između konstruktora, zadataka i CAD sustava. Model kognitivnog opterećenja nadovezuje se na model konstruktorove izvedbe te opisuje kako se zahtjevi za obradu informacija mogu prikazati korištenjem perceptivnih i kognitivnih resursa. U okviru preskriptivnog dijela istraživanja razvijena je i metoda za mjerenje i analizu kognitivnog opterećenja u CAD kontekstu, koja povezuje subjektivnu procjenu opterećenja, psihofiziološko mjerenje i modeliranje ljudske izvedbe temeljeno na iskorištenosti perceptivnih, kognitivnih i motoričkih operatora.

Deskriptivno istraživanje II obuhvatilo je drugu empirijsku studiju kojom je propisana metoda operacionalizirana i testirana u CAD kontekstu. Eksperimentalni nacrt uključivao je subjektivnu procjenu kognitivnog opterećenja pomoću NASA TLX upitnika, psihofiziološko mjerenje EEG-om, mjerenje uspješnosti konstruktorove CAD izvedbe pomoću procesnih i izlaznih pokazatelja te modeliranje konstruktorove izvedbe perceptivnim, kognitivnim i motoričkim operatorima. U ovom je eksperimentu ispitivan utjecaj unutarnje složenosti CAD zadatka kao glavne varijable. Studijom su identificirani EEG pokazatelji kognitivnog opterećenja prikladni za CAD kontekst, istaknuti su EEG pokazatelji zatim uspoređeni s rezultatima analize iskorištenosti operatora i subjektivnim procjenama. Dinamičke promjene kognitivnog opterećenja povezane su sa segmentima procesa CAD modeliranja i s pokazateljima uspješnosti konstruktorove izvedbe temeljenima na analizi konačnoga 3D CAD modela.

Rezultati druge empirijske studije istaknuli su parijetalnu i okcipitalnu snagu (moždanu aktivnost) alfa valova povezanu sa zadatkom (eng. *Task Related Power*; TRP) kao EEG pokazatelje kognitivnog opterećenja u CAD kontekstu. Također je potvrđen utjecaj unutarnje složenosti zadatka na izvedbu konstruktora i kognitivno opterećenje tijekom CAD modeliranja. Pokazatelji kognitivnog opterećenja temeljeni na iskorištenosti kognitivnih i perceptivnih operatora omogućili su razlikovanje kognitivnog opterećenja između pojedinih faza i segmenata zadataka, čime je potvrđena njihova korisnost za analizu procesa CAD modeliranja. Međutim, dinamičko podudaranje između EEG pokazatelja i pokazatelja temeljenih na iskorištenosti operatora bilo je ograničeno, što upućuje na to da ovi pokazatelji zahvaćaju povezane, ali ne potpuno jednake aspekte kognitivnog opterećenja. Regresijske analize otkrile su poveznicu između više procijenjene razine kognitivnog opterećenja i niže razine uspješnosti CAD izvedbe, pri čemu je iskorištenost kognitivnih i perceptivnih operatora imala najstabilniji odnos s pokazateljima uspješnosti izvedbe. Na temelju tih rezultata predložena je i pojednostavljena inačica metode koja smanjuje zahtjeve za prikupljanjem i obradom podataka, a zadržava mogućnost praćenja promjena kognitivnog opterećenja tijekom CAD modeliranja.

Validacija razvijenih modela i metode provedena je prema okviru *Validation Square*, koji omogućuje procjenu teorijske strukturne valjanosti, empirijske strukturne valjanosti, empirijske uporabne valjanosti i teorijske uporabne valjanosti. Teorijski modeli vrednovani su prvenstveno teorijski; s obzirom na to predstavljaju li njihovi gradivni elementi i odnosi među njima prikladnu osnovu za razumijevanje izvedbe i kognitivnog opterećenja. Metoda je vrednovana prema svim koracima validacije, pri čemu je provjerena prikladnost propisanih ulaznih podataka, postupaka obrade i izlaznih podataka, mogućnost primjene na realističnim CAD zadacima te korisnost rezultata za povezivanje kognitivnog opterećenja s uspješnošću konstruktorove izvedbe. Kriteriji vrednovanja definirani su prema svrsi i ciljevima teorijskih modela i metode.

Znanstveni doprinos doktorskoga rada obuhvaća tri međusobno povezana aspekta. Prvi doprinos jest teorijski model kognitivnog opterećenja u CAD aktivnostima, koji formalizira način na koji se zahtjevi za obradu informacija mogu prikazati u odnosu na proces CAD modeliranja i njegove segmente. Drugi doprinos jest teorijski model konstruktorove izvedbe u CAD aktivnostima, koji formalizira izvedbu iz perspektive obrade informacija i utjelovljene kognicije (eng. *Embodied Cognition*) te povezuje procesne i izlazne pokazatelje CAD modeliranja. Treći doprinos jest metoda za mjerenje i analizu kognitivnog opterećenja u CAD aktivnostima, koja integrira modeliranje ljudske izvedbe, EEG pokazatelje i subjektivnu procjenu kognitivnog opterećenja, omogućujući dinamičku analizu kognitivnog opterećenja na razini segmenata CAD procesa i njegovo povezivanje s uspješnošću konstruktorove izvedbe u CAD aktivnostima.

Razvijeni modeli i metodu pritom je preporučeno tumačiti kao istraživački okvir za razumijevanje procesne perspektive konstruktorove izvedbe u CAD aktivnostima, a ne kao izravni preskriptivni alat za automatsko poboljšanje CAD izvedbe. Njihova je primjena u ovom doktorskom radu vezana uz individualno modeliranje pomoću značajki u parametarskim CAD sustavima, ponajprije u kontekstu oblikovanja i detaljne konstrukcijske razrade. Stoga je za širu primjenu u industrijskom kontekstu potrebno dodatno ispitati primjenjivost modela i metode u drugim CAD aktivnostima, CAD paradigmama, ranijim i kasnijim fazama razvojnoga procesa i složenijim industrijskim zadatcima. Buduća istraživanja trebala bi istražiti mogućnosti automatizacije prikupljanja procesnih podataka iz CAD zapisa, makronaredbi i zapisa međudjelovanja između čovjeka i računala kako bi se pojednostavilo korištenje metode. Nadalje, bilo bi korisno provesti dodatnu kalibraciju prepoznatih pokazatelja kognitivnog opterećenja te istražiti primjenu spoznaja dobivenih korištenjem metode u razvoju CAD sustava, obrazovanju konstruktora i oblikovanju procesa i aktivnosti konstruiranja kako bi se smanjilo nepotrebno kognitivno opterećenje.

Ključne riječi:

Računalno potpomognuto konstruiranje; Kognitivno opterećenje; Procesuiranje informacija; Modeliranje ljudske izvedbe; Elektroencefalografija

LIST OF FIGURES

Figure 1 Building bricks of human performance modelling 30
Figure 2 Memories and processors of the MHP, adopted from Card et al. [7]........................ 32
Figure 3 Knowledge levels and modelling strategies, adopted from Bhavnani and John [98] 47
Figure 4 Model of CL, suggested by Paas and Van Merriënboer [107] 51
Figure 5 Effective-ineffective CL in information processing, suggested by Xie and Salvendy [102] 53
Figure 6 Model for predicting CL, suggested by Xie and Salvendy [102] 54
Figure 7 Model of CL including potential load, suggested by Lavrsen and Daalhuizen [110] 56
Figure 8 ES#1: Participants' frequency of CAD modelling and creating or reviewing technical documentation 84
Figure 9 ES#1: Experimental procedure 84
Figure 10 ES#1: Steps of the introductory CAD modelling task 85
Figure 11 ES#1: Isometric projection of the part 1 (left) and orthographic projection of the part 2 (right) 86
Figure 12 ES#1: Experimental setup 87
Figure 13 ES#1: Completeness of CAD modelling outcomes 91
Figure 14 ES#1: Total time (panel a) and edit time (panel b) 91
Figure 15 Sketch entities: a) Average number; b) Total number 92
Figure 16 Sketch relations: a) Average number; b) Total number; c) The most used relations 93
Figure 17 CAD features: a) Total number; b) Distribution per type 93
Figure 18 Process layers 99
Figure 19 Average performer: Descriptive model of CAD performance - Isometric condition 108
Figure 20 Average performer: Descriptive model of CAD performance - Orthographic condition 108
Figure 21 Average performer: Visualisation of unit task phases - Isometric condition 110
Figure 22 Average performer: Visualisation of unit task phases - Orthographic condition ... 110
Figure 23 Average performer: Distribution of subtask clusters across unit tasks - Isometric condition 111
Figure 24 Average performer: Distribution of subtask clusters across unit tasks - Orthographic condition 111
Figure 25 Average performer: Distribution of subtasks 112

Figure 26 Average performer: Average utilisation of operators in each unit task – Isometric condition 114
Figure 27 Average performer: Average utilisation of operators in each unit task – Orthographic condition 114
Figure 28 ES#1: EEG data pre-processing pipeline 127
Figure 29 CAD task segmentation example 128
Figure 30 Duration of the a) entire CAD task, b) first CAD task segment, and c) second CAD task segment 131
Figure 31 a) Theta TRP, b) Alpha TRP, and c) Beta TRP over the cortex when interpreting the projection 131
Figure 32 a) Theta TRP, b) Alpha TRP, and c) Beta TRP over the cortical hemispheres when interpreting the projections 132
Figure 33 a) Theta TRP, b) Alpha TRP, and c) Beta TRP over the cortical areas when interpreting the projections 133
Figure 34 a) Theta TRP, b) Alpha TRP, and c) Beta TRP over the cortex when generating the 3D CAD models 133
Figure 35 a) Theta TRP, b) Alpha TRP, and c) Beta TRP over the cortical hemispheres when generating the 3D CAD models 134
Figure 36 a) Theta TRP, b) Alpha TRP, and c) Beta TRP over the cortical areas when generating the 3D CAD models 136
Figure 37 Theta TRP at the individual electrodes when generating the CAD models 136
Figure 38 Key elements of CAD performance and CL in CAD 146
Figure 39 Visualisation of the theoretical model of CAD performance, depicting hierarchical levels of CAD tasks/goals, process layers, and model information layers 152
Figure 40 High-level subtasks/subgoals with corresponding unit task operations and (S+F) model information layer 154
Figure 41 Intermediate-level subtasks/subgoals with corresponding CAD operations and (GE+R) model information layer 155
Figure 42 Low-level subtasks/subgoals with corresponding CAD actions 157
Figure 43 Visualisation of unit task phases within the high-level subtasks/subgoals 158
Figure 44 Mapping the operators to CAD actions and unit task phases 163
Figure 45 Steps to build a model of CAD performance 177
Figure 46 An example of a decomposition 179
Figure 47 An example of mapping the operators 182

Figure 48 An example of the model of CAD performance 183
Figure 49 Steps to build a model of CL in CAD 184
Figure 50 EEG data pre-processing pipeline 187
Figure 51 EEG data analysis pipeline 188
Figure 52 Selection of an inferential test.. 190
Figure 53 Steps for calculating information-processing weights 192
Figure 54 An example of the model of CL at the level of moving windows (10s, 2s) 196
Figure 55 An example of the model of CL at the level of unit task phases........................... 196
Figure 56 An example of the model of CL faceted across the subtask clusters 197
Figure 57 ES#2: Experimental procedure 204
Figure 58 ES#2: Experimental setup 205
Figure 59 Low-complexity component (left) and high-complexity component (right)......... 206
Figure 60 Differences in EEG features between the HC and the LC task 214
Figure 61 Repeated-measures correlations between NASA TLX scores and EEG features.. 216
Figure 62 Theil-Sen regression models between EEG features and mental demand (NASA TLX) 218
Figure 63 Median parietal alpha TRP across unit task phases and subtasks 225
Figure 64 Overall CL values calculated from utilisation-based indicators 232
Figure 65 Visualisation of Theil-Sen regression models between utilisation-based indicators and NASA TLX scores... 233
Figure 66 Differences in CL assigned to subtasks when assessed by different indicators..... 241
Figure 67 Differences in CL assigned to subtasks when assessed by different indicators and considering time share... 244
Figure 68 Subtasks' contribution to the overall CL 245
Figure 69 P#4: Dynamic changes in CL across moving windows (10s, 2s) – High Performer 250
Figure 70 P#16: Dynamic changes in CL across moving windows (10s, 2s) – Average Performer 251
Figure 71 P#8: Dynamic changes in CL across moving windows (10s, 2s) – Low Performer 251
Figure 72 Theil-Sen regression models: Utilisation-based CL and Overall performance 257
Figure 73 P#4: Dynamic changes in CL indicated by global values of weighted utilisations – High Performer.. 286

Figure 74 P#16: Dynamic changes in CL indicated by global values of weighted utilisations – Average Performer.. 287
Figure 75 P#8: Dynamic changes in CL indicated by global values of weighted utilisations – Low Performer 287
Figure 76 High performer: Descriptive model of CAD performance - Isometric condition . 318
Figure 77 High performer: Descriptive model of CAD performance - Orthographic condition 319
Figure 78 High performer: Visualisation of unit task phases - Isometric condition 319
Figure 79 Visualisation of unit task phases - Orthographic condition 320
Figure 80 High performer: Distribution of subtask clusters across unit tasks - Isometric condition 321
Figure 81 High performer: Distribution of subtask clusters across unit tasks - Orthographic condition 321
Figure 82 High performer: Distribution of subtasks 322
Figure 83 High performer: Average utilisation of operators across unit tasks – Isometric condition 323
Figure 84 High performer: Average utilisation of operators across unit tasks – Orthographic condition 324
Figure 85 Low performer: Descriptive model of CAD performance - Isometric condition .. 326
Figure 86 low performer: Descriptive model of CAD - Orthographic condition 326
Figure 87 Low performer: Visualisation of unit task phases - Isometric condition 328
Figure 88 Low performer: Visualisation of unit task phases - Orthographic condition......... 328
Figure 89 Low performer: Distribution of subtask clusters across unit tasks - Isometric condition 329
Figure 90 Low performer: Distribution of subtask clusters across unit tasks - Orthographic condition 330
Figure 91 Low performer: Distribution of subtasks 331
Figure 92 Low performer: Average utilisation of operators in each unit task – Isometric condition 333
Figure 93 Low performer: Average utilisation of operators in each unit task – Orthographic condition 334
Figure 94 P#1: Dynamic changes in CL across moving windows (10s, 2s) 336
Figure 95 P#1: Dynamic changes in CL indicated by global values of weighted utilisations 337
Figure 96 P#3: Dynamic changes in CL across moving windows (10s, 2s) 338

Figure 97 P#3: Dynamic changes in CL indicated by global values of weighted utilisations 338
Figure 98 P#4: Dynamic changes in CL across moving windows (10s, 2s) 339
Figure 99 P#4: Dynamic changes in CL indicated by global values of weighted utilisations 340
Figure 100 P#5: Dynamic changes in CL across moving windows (10s, 2s) 341
Figure 101 P#5: Dynamic changes in CL indicated by global values of weighted utilisations 341
Figure 102 P#6: Dynamic changes in CL across moving windows (10s, 2s) 342
Figure 103 P#6: Dynamic changes in CL indicated by global values of weighted utilisations 343
Figure 104 P#7: Dynamic changes in CL across moving windows (10s, 2s) 344
Figure 105 P#7: Dynamic changes in CL indicated by global values of weighted utilisations 344
Figure 106 P#8: Dynamic changes in CL across moving windows (10s, 2s) 345
Figure 107 P#8: Dynamic changes in CL indicated by global values of weighted utilisations 346
Figure 108 P#10: Dynamic changes in CL across moving windows (10s, 2s) 347
Figure 109 P#10: Dynamic changes in CL indicated by global values of weighted utilisations 347
Figure 110 P#11: Dynamic changes in CL across moving windows (10s, 2s) 348
Figure 111 P#11: Dynamic changes in CL indicated by global values of weighted utilisations 349
Figure 112 P#12: Dynamic changes in CL across moving windows (10s, 2s) 350
Figure 113 P#12: Dynamic changes in CL indicated by global values of weighted utilisations 350
Figure 114: P#13: Dynamic changes in CL across moving windows (10s, 2s) 351
Figure 115 P#13: Dynamic changes in CL indicated by global values of weighted utilisations 352
Figure 116 P#14: Dynamic changes in CL across moving windows (10s, 2s) 353
Figure 117 P#14: Dynamic changes in CL indicated by global values of weighted utilisations 353
Figure 118 P#15: Dynamic changes in CL across moving windows (10s, 2s) 354
Figure 119 P#15: Dynamic changes in CL indicated by global values of weighted utilisations 355
Figure 120 P#16: Dynamic changes in CL across moving windows (10s, 2s) 356

Figure 121 P#16: Dynamic changes in CL indicated by global values of weighted utilisations 356
Figure 122 P#17: Dynamic changes in CL across moving windows (10s, 2s) 357
Figure 123 P#17: Dynamic changes in CL indicated by global values of weighted utilisations 358
Figure 124 P#18: Dynamic changes in CL across moving windows (10s, 2s) 359
Figure 125 P#18: Dynamic changes in CL indicated by global values of weighted utilisations 359
Figure 126 P#19: Dynamic changes in CL across moving windows (10s, 2s) 360
Figure 127 P#19: Dynamic changes in CL indicated by global values of weighted utilisations 361
Figure 128 P#20: Dynamic changes in CL across moving windows (10s, 2s) 362
Figure 129 P#20: Dynamic changes in CL indicated by global values of weighted utilisations 362
Figure 130 P#21: Dynamic changes in CL across moving windows (10s, 2s) 363
Figure 131 P#21: Dynamic changes in CL indicated by global values of weighted utilisations 364
Figure 132 P#22: Dynamic changes in CL across moving windows (10s, 2s) 365
Figure 133 P#22: Dynamic changes in CL indicated by global values of weighted utilisations 365
Figure 134 P#23 Dynamic changes in CL across moving windows (10s, 2s) 366
Figure 135 P#23: Dynamic changes in CL indicated by global values of weighted utilisations 367
Figure 136 P#24: Dynamic changes in CL across moving windows (10s, 2s) 368
Figure 137 P#24: Dynamic changes in CL indicated by global values of weighted utilisations 368

LIST OF TABLES

Table 1 Research stages 6
Table 2 Objectives for the theoretical models, according to Wu et al. [23] 10
Table 3 Objectives for the outputs of the method and the process by which they are generated 11
Table 4 List of studies on engineering designers' CAD performance 18
Table 5 The MHP's principles of operation, adopted from Card et al. [7] 35
Table 6 Overview of task analysis frameworks 39
Table 7 List of KLM's operators, adopted from Card et al. [92] 40
Table 8 KLM's heuristic rules for placing the M operator, adopted from Card et al. [92] 41
Table 9 Cognitive architectures and WL/CL quantification 58
Table 10 Mathematical equations for calculating CL/WL from cognitive models, adopted from Wu and Liu [116] 59
Table 11 Linear regression equations for relating WL from cognitive models and NASA TLX, adopted from Wu and Liu [116] 59
Table 12 Linear regression model relating WL from cognitive models and NASA TLX, adopted from Cao and Liu [120], [121] 60
Table 13 Mathematical equations for calculating CL/WL from cognitive models, adopted from Jo et al. [112] 60
Table 14 Linear regression model for relating WL from cognitive models and NASA TLX, adopted from Jo et al. [112] 61
Table 15 Mathematical equations for calculating CL/WL from cognitive models, adopted from Park and Myung [119] 61
Table 16 Linear regression model for relating WL from cognitive models and NASA TLX, adopted from Park and Myung [119],[118] 62
Table 17 Mathematical equations for calculating CL/WL from cognitive models, adopted from Oh et al. [117] 62
Table 18 List of studies on cognitive load in (engineering) design tasks 64
Table 19 Methods for measuring CL 70
Table 20 EEG-based CL indicators in (computer-aided) design or similar HCI tasks 75
Table 21 ES#1: Demographics, education, and prior experience 83
Table 22 ES#1: Description of collected experimental data 87
Table 23 CAD modelling outcomes: Completeness 90

Table 24 CAD modelling process: Duration ... 91
Table 25 CAD modelling process: Sketch entities ... 92
Table 26 CAD modelling process: Sketch relations ... 92
Table 27 CAD modelling process: CAD features ... 93
Table 28 Data of three participants whose video screen captures were manually coded ... 98
Table 29 List of CAD actions ... 100
Table 30 List of subtasks/subgoals ... 101
Table 31 Unit task phases ... 103
Table 32 List of actions, operators, and associated systems ... 104
Table 33 Coding rules for assigning operators ... 105
Table 34 Average performer: Types, number of occurrences, and average duration of unit tasks ... 109
Table 35 Average performer: Number of occurrences, average duration, and distribution of unit task phases ... 109
Table 36 Average performer: Number of occurrences and average duration of subtasks at the level of the entire task ... 112
Table 37 Average performer: Number of occurrences and average duration of CAD actions at the level of the entire task ... 113
Table 38 Average performer's average operator utilisation across subtasks ... 115
Table 39 Number of occurrences, and relative duration of unit task phases across conditions and performance profiles ... 119
Table 40 Relative duration of subtask clusters (CAD operations) across task conditions and performance profiles ... 120
Table 41 Occurrences of subtask clusters (CAD operations) across task conditions and performance profiles ... 121
Table 42 Number of occurrences (N) and mean (M) duration of CAD actions across task conditions and performance profiles ... 122
Table 43 Comparison of the TRPs over the cortex when interpreting the projections ... 131
Table 44 The main effect of the cortical hemisphere ... 132
Table 45 Significant differences in TRP regarding the cortical hemispheres and projections 132
Table 46 Comparison of the TRPs over the cortex when generating the 3D CAD models ... 133
Table 47 The main effect of the cortical hemisphere ... 134
Table 48 Difference between the LH and the RH in theta TRP over the individual electrodes ... 135

Table 49 Difference between the LH and the RH in alpha TRP over the individual electrodes 135
Table 50 Simple main effects on theta TRP 136
Table 51 Theta TRP at and among the individual electrodes 137
Table 52 List of operators, operations, and associated systems 150
Table 53 List of subtask clusters and CAD operation examples 156
Table 54 List of CAD actions 157
Table 55 Parts of unit tasks/goals 160
Table 56 Proposed information-processing intensity rank of segments, based on utilisations 166
Table 57 Input data requirements to the method 175
Table 58 List of CAD events 177
Table 59 An example of a HCI script 178
Table 60 Heuristics for assigning P, C, M, and R operators 181
Table 61 List of EEG features emerging from literature 185
Table 62 ES#2: Demographics, prior experience, and frequency of conducting CAD activities 203
Table 63 Description of collected experimental data in the ES#2 205
Table 64 Descriptions of NASA TLX components, adopted from [181] 207
Table 65 CAD performance outcome metrics 208
Table 66 Differences in NASA TLX scores between the HC and the LC task 209
Table 67 List of considered EEG features 210
Table 68 Differences in EEG features between the HC and the LC task 215
Table 69 Repeated-measures correlation results between mental demand and EEG features 215
Table 70 Theil-Sen regressions 217
Table 71 Descriptive statistics of parietal alpha TRP values for unit task phases and subtasks 225
Table 72 Differences in parietal alpha TRP between acquisition, inspection, and subtasks .. 226
Table 73 Information-processing weights assigned to acquisition, inspection, and subtasks 227
Table 74 Global utilisation values and information-processing weights 231
Table 75 Summary of overall CL values calculated from utilisations 232
Table 76 Theil-Sen regression models: Utilisation-based indicators and Mental Demand (NASA TLX) 233

Table 77 Friedman test: Differences in CL associated with acquisition, inspection, and subtasks 239
Table 78 Significant pairwise differences in CL associated with acquisition, inspection, and subtasks 240
Table 79 Friedman test: Differences in CLC of acquisition, inspection, and unit task phases to the overall CL 242
Table 80 Significant pairwise differences in CLC of subtasks to the overall CL 242
Table 81 Summary of alignment metrics by indicator pairs (Med; IQR) 253
Table 82 Comparison of alignment between utilisation-based indicators and parietal alpha TRP 254
Table 83 Theil-Sen regression models: Utilisation-based indicators and Overall performance 256
Table 84 High performer: Types, number of occurrences, and average duration of unit tasks 319
Table 85 High performer: Number of occurrences, average duration, and distribution of unit task phases 320
Table 86 High performer: Number of occurrences and average duration of subtasks at the level of the entire task 322
Table 87 High performer: Number of occurrences and average duration of CAD actions at the level of the entire task 322
Table 88 High performer's average operator utilisation across subtasks 324
Table 89 Low performer: Types, number of occurrences, and average duration of unit tasks327
Table 90 Low performer: Number of occurrences, average duration, and distribution of unit task phases 327
Table 91 Low performer: Number of occurrences and average duration of subtasks at the level of the entire task 331
Table 92 Low performer: Number of occurrences and average duration of CAD actions at the level of the entire task 332
Table 93 Low performer's Average operator utilisation across subtasks 334
Table 94 P#1: characteristics and CAD performance outputs 336
Table 95 P#3: characteristics and CAD performance outputs 337
Table 96 P#4: characteristics and CAD performance outputs 339
Table 97 P#5: characteristics and CAD performance outputs 340
Table 98 P#6: characteristics and CAD performance outputs 342

Table 99 P#7: characteristics and CAD performance outputs ... 343
Table 100 P#8: characteristics and CAD performance outputs ... 345
Table 101 P#10: characteristics and CAD performance outputs ... 346
Table 102 P#11: characteristics and CAD performance outputs ... 348
Table 103 P#12: characteristics and CAD performance outputs ... 349
Table 104 P#13: characteristics and CAD performance outputs ... 351
Table 105 P#14: characteristics and CAD performance outputs ... 352
Table 106 P#15: characteristics and CAD performance outputs ... 354
Table 107 P#16: characteristics and CAD performance outputs ... 355
Table 108 P#17: characteristics and CAD performance outputs ... 357
Table 109 P#18: characteristics and CAD performance outputs ... 358
Table 110 P#19: characteristics and CAD performance outputs ... 360
Table 111 P#20: characteristics and CAD performance outputs ... 361
Table 112 P#21: characteristics and CAD performance outputs ... 363
Table 113 P#22: characteristics and CAD performance outputs ... 364
Table 114 P#23: characteristics and CAD performance outputs ... 366
Table 115 P#24: characteristics and CAD performance outputs ... 367
Table 116 Metrics for quantifying alignment in dynamic changes of CL indicated by parietal alpha TRP and utilisations ... 369

LIST OF ABBREVIATIONS AND SYMBOLS

ABBREVIATIONS

ACT-R	-	Adaptive Control of Thought-Rational
ACT-R/PM	-	Adaptive Control of Thought-Rational/Perceptual Motor
AEI	-	Acquisition, Execution, Inspection
AIS	-	Auditory Image Store
C	-	Cognitive
CAD	-	Computer-Aided Design
CL	-	Cognitive Load
CPM	-	Cognitive Perceptual Motor; Critical Path Method
DRM	-	Design Research Methodology
EDR	-	Engineering Design Research
EEG	-	Electroencephalography
GOMS	-	Goals, Operators, Methods, Selection rules
HC	-	High Complexity
HCI	-	Human-Computer Interaction
KLM	-	Keystroke Level Model
LC	-	Low Complexity
LTM	-	Long-Term Memory
M	-	Motor; Mean
MAD	-	Median Absolute Deviation
Med	-	Median
MHP	-	Model Human Processor
P	-	Perceptual
POW	-	Frequency band POWer
QN	-	Queueing Network
R	-	System Response
SD	-	Standard Deviation
T	-	Time

TLX	-	TaskLoad Index
TRP	-	Task Related Power
VIS	-	Visual Image Store
WL	-	WorkLoad
WM	-	Working Memory

SYMBOLS

ω	-	Information-processing weight
w	-	Base operator weight
T	-	Time

This page intentionally left blank

1 INTRODUCTION

The first chapter introduces the research motivation and scope, defines the research aims and hypothesis that guide the work, and provides the conceptual context for the subsequent chapters. The chapter then summarises the adopted research methodology, states the purposes and objectives used for developing and validating the proposed models and the method, outlines the expected scientific contributions, and concludes with an overview of the thesis structure.

1.1 Research motivation, aims, and hypothesis

In Computer-Aided Design (CAD) engineering designers continuously interact with CAD systems, consisting of CAD software and Human-Computer Interaction (HCI) tools, while solving challenging design tasks [1]. CAD systems are regularly employed in the contemporary engineering design process for creating, recreating, reviewing, and modifying digital design representations (e.g. CAD representations) [2]. CAD system characteristics shape what information is available, how it is presented, and what actions are required to progress within and across the design activities they support (i.e. CAD activities) [1], [3]. Consequently, CAD systems play a fundamental role in many design activities and thus affect the engineering designer's performance – one of the essential factors determining the quality, cost, and timelines of the contemporary product development process [2], [4], [5].

Despite its importance for the success of the engineering design process, the definition of CAD performance has not yet been formalised in a way that makes its key elements, relationships among them, and influencing factors explicit. This lack of formalism is accompanied by a paucity of support for its consistent measurement and analysis across studies and tasks. A major reason is that existing approaches have often emphasised CAD output quality, while providing less explicit representation of how performance unfolds over time and which perceptual, cognitive, and motor demands are imposed by interacting with the CAD system. An information-processing and HCI perspective may address this limitation by enabling CAD performance to be represented as a structured process and interpreted in terms of underlying resource engagement, thereby complementing outcome-based assessments and supporting systematic comparison across engineering designers, CAD activities, and CAD systems [6]. *To address this gap, the first research aim of the thesis is to develop and validate a theoretical model of engineering designer's performance in CAD activities*.

Within the thesis, CAD performance is observed from the information-processing and HCI perspectives; it is seen as a usage of engineering designers' perceptual, cognitive, and motor resources to transform requirements and descriptions of a design problem as an input information into the design representations as an output information. In line with embodied cognition, information processing emerges from continuous perception–action coupling with the CAD system, where interaction with HCI tools and external representations (e.g. technical drawings, 3D CAD models) supports and structures cognition [7]. Modelling the allocation of engineering designers' information-processing resources during CAD activities would thus enhance understanding of how CAD activities are executed throughout the process to reach the outcome, namely a design representation. Explaining the basic physical, perceptual, and cognitive operations engineering designers use, in relation to psychological evidence, has been suggested as one way of improving their effectiveness and efficiency [8]. Importantly, this perspective highlights the need to account for the cost of information processing in CAD activities, in terms of the demands placed on limited cognitive resources.

This need becomes especially clear when considering that formalising, measuring, and assessing the engineering designers' performance is aggravated by the visual (information is presented visually), virtual (they are situated in virtual environments), and cognitive nature of CAD activities [1], [6]. The cognitive complexity of CAD activities arises from the characteristics of design tasks as well as from the use of CAD systems [5], [9]. In particular, engineering designers simultaneously attend to, perceive, and interpret multiple information streams when conducting CAD activities [10]. Furthermore, they employ various cognitive abilities (e.g. spatial visualisation and mental rotation) and motor skills (related to HCI tools) to reach the high number of subgoals that the CAD activities consist of (e.g. drawing, dimensioning, manipulating views, etc.) [11]. The CAD systems offer the possibility to reach these subgoals in alternative ways, thus continuously making the engineering designers evaluate the possible options and decide on how to proceed [12]. It is assumed that the execution of CAD activities, due to their cognitive complexity, imposes high demands (i.e. cognitive load; CL) on the engineering designers' cognitive system [13]. Since engineering designers' information-processing cognitive resources are unavoidably limited, CAD performance depends on how these resources are allocated across CAD activity demands and HCI requirements; excessive CL can constrain effective and efficient performance in the virtual (CAD) environment [1]. Accordingly, understanding CAD performance benefits from formalising and measuring CL as part of the same resource-allocation problem.

From this viewpoint, the prerequisite for enhancing CAD performance (reflected in high-quality outputs and/or faster, less resource-intensive processes) is understanding, formalising, and measuring CL as costs of accomplishing CAD tasks using available perceptual and cognitive resources. CL is a multidimensional construct emerged from the observation that the human cognitive system has a limited capacity for information processing when performing cognitive tasks [14]. One of the main motivations guiding the research on CL within engineering design is its potential to facilitate the development of methods and tools that support engineering designers' cognitive processing and expand their information-processing capabilities [1], [5]. Given the prevalent use of computers in contemporary engineering design, and their significant contribution to engineering designers' performance, human-computer interaction (HCI) becomes a key area where CL should be closely examined. Enhanced HCI, better suited to limited human cognitive resources, could be achieved by redesigning CAD systems in a way that optimizes the flow and exchange of design information between engineering designers and virtual design environments [5]. For example, this could involve triggering information prompts, disseminating relevant information, or re-allocating certain tasks to the computer, which then may result in faster, less resource-intensive processes and high-quality outputs. Such advancements in HCI and CAD performance rely on the thorough understanding, and continuous measurement and observation of the CL experienced by engineering designers while solving CAD tasks. However, in the CAD context, CL has most often been treated as a general explanation for task difficulty or performance differences, while CAD-specific characterisations of when CL changes during CAD activities, which CAD tasks and interaction elements drive it, and how it can be measured and interpreted continuously remain limited and inconsistent across studies. This motivates theoretical formalisation regarding the key elements, relationships among them, and influencing factors. *To address this gap, the second research aim of the thesis is to develop and validate a theoretical model of CL in CAD activities*. Moreover, because CL is covert, measuring it in a reliable, objective and continuous manner presents challenges, especially during complex and lengthy tasks like those in CAD. *Accordingly, the third research aim of the thesis is to develop and validate a method for measuring and analysing CL in CAD activities*.

Finally, it is necessary to consider CL not only as a standalone construct, but also in terms of how it shapes CAD performance outcomes and processes. Previous studies within engineering design have revealed varying relationships between CL and design performance [15]–[18]. Scholars have used diverse methods and metrics to assess both CL and design performance. Subjective methods, particularly questionnaires like NASA Task Load IndeX (TLX), dominate

CL measurement, though physiological methods such as heart rate variability and electroencephalography (EEG) have also been explored. In parallel, CAD performance has typically been evaluated through outcome-based metrics (quality or accuracy of the final CAD models), with fewer studies providing structured, temporally informed descriptions of CAD modelling processes. However, these two research streams have largely progressed separately; CAD performance assessments are rarely linked to concurrent CL measurement, and CL evidence in CAD activities remains limited. As a result, the relationship between CL experienced by engineering designers and their CAD performance remains underexplored. Therefore, the reviewed literature underscores the need for studies that directly connect CL measures to CAD performance metrics, providing a foundation for understanding relationships between CL and performance in CAD activities. *Building on this need, the research hypothesises that the prescribed method for measuring and analysing CL, based on the monitoring of psychophysiological responses, enables the assessment of the engineering designer's level of CAD activity performance*.

In this thesis, CAD activities are retained as the broader term for design activities supported by CAD systems [2], while CAD modelling is treated as a specific and central type of CAD activity. CAD modelling refers to the activity in which engineering designers create, recreate, modify, inspect, and refine CAD representations, most commonly three-dimensional (3D) CAD models. This distinction is important because CAD activities may also include other computer-supported design activities, such as simulation, design review, topology optimisation, functional decomposition, conceptual modelling, or virtual product disassembly. However, the theoretical and empirical focus of this thesis is on CAD modelling activities, particularly the construction of 3D CAD models in a CAD system, through which the structure and physical attributes of a technical system are represented [2].

CAD modelling was selected as the focus because it represents one of the most common and central CAD activities in contemporary engineering design, especially during embodiment and detail design [2]. From an information-processing and HCI perspective, CAD modelling also provides a suitable core activity for formalisation because it involves continuous perception-action coupling with the CAD system; engineering designers perceive and interpret external representations, select and execute commands, manipulate views and geometry, inspect intermediate outcomes, and progressively transform available design information into CAD representations. Therefore, CAD modelling offers a clearly observable setting for analysing how perceptual, cognitive, and motor resources are allocated during interaction with CAD systems. Once this core CAD activity is formalised, the underlying logic may be extended to other CAD activities, provided that additional model elements and parameters (goal structures, operation and action classes, and segmentation schemes) are defined for those activities. Consequently, engineering design is considered in this thesis from the specific perspective of CAD modelling as a transformation of available design information with predefined characteristics and properties, rather than as the full breadth of engineering design process. Broader design activities, such as problem framing, concept generation, early-stage exploration, and the iterative co-evolution of problem and solution, are not primary focus of the thesis. The findings should therefore be interpreted as primarily applicable to CAD modelling and related HCI processes, while their generalisation to wider design activities requires further extension and validation.

1.2 Research methodology

The purpose of the PhD thesis is to develop a support in the form of the method that enables understanding and assessing the engineering designers' performance in CAD activities. It is intended to provide such support by modelling, measuring, and analysing the CL in CAD activities. The prerequisite for the method development is a thorough understanding of the CL, engineering designers' performance, and the means to measure and analyse them in CAD activities. The current understanding of these phenomena and means is unsatisfactory due to the limited number of studies investigating them in the CAD context, which asks for theoretical development prior to the prescription of the method. The guidelines for developing the understanding and support are found within the Design Research Methodology (DRM) [19], used as the main research framework. The guidelines provided by the DRM are complemented by the principles of Experimental Design Research (EDR) [20]. Additionally, validation of the developed models and the method follows the Validation Square framework [21]. The research activities are organised in four main stages: Research clarification, Descriptive study I (that embraces literature review and the first empirical study), Prescriptive study (that proposes the theoretical models and the method) and Descriptive study II (that validates the developed theoretical models and the method). The organisation of research activities in these research stages is available in Table 1 and further explained in the following sections.

Table 1 Research stages

Stage	Type	Method and Aim	Deliverables
Research clarification	Review-based	Literature review	• Research aims and hypothesis • Purpose and objectives of the method and the models
Descriptive study I	Comprehensive	Literature review • Describing the existing situation • Theoretical structural validation basis	• An overview of available (theoretical) models of CAD performance and CL in CAD • An overview of key elements, their characteristics, relationships between elements, and influencing factors (for similar tasks) • An overview of tools/methods and metrics that have been used for modelling and measuring CAD performance • An overview of tools/methods and metrics that have been used for modelling and measuring CL
			• Research questions for the first empirical study
		Experimental study I • Describing the existing situation • Empirical structural and performance validation	• Confirmed a representation type as an influencing factor • Confirmed suitability of EEG for capturing brain activity in CAD • Confirmed suitability of human performance modelling for modelling engineering designer's CAD performance • Built models of CAD performance for three participants • Modelled dynamic changes in CL indicated by utilisations of operators
			• Research questions for the second empirical study
Prescriptive study	Comprehensive	• Prescribing the models and the method	• Theoretical model of CAD performance • Theoretical model of CL in CAD • The initial method • Detailed requirements for inputs to the method
Descriptive study II	Comprehensive	Experimental study II • Empirical structural and performance validation	• Validated the theoretical models and the method • Confirmed internal task complexity as an influencing factor • Identified EEG-based CL indicators for the CAD context • Assigned information-processing weights to task segments • Compared results implied by EEG-based indicators, utilisation-based indicators, and self-reported subjective scores • Modelled dynamic changes in CL in relation to CAD performance • Related CL indicated by weighted utilisation-based indicators to CAD performance (outcome) assessments
Prescriptive study	Initial	• Suggesting the simplified method based on the results of the second empirical study	• Simplified version of the method

1.2.1 Research clarification

The Research clarification stage incorporated a narrative literature review on engineering designers' performance and CL in CAD activities to establish the state of the art and identify research gaps. The review did not follow a structured literature review approach (e.g. Preferred Reporting Items for Systematic reviews and Meta-Analyses; PRISMA), but it was organised as a comprehensive summary and critical analysis of existing research on the phenomena the thesis focuses on. The review first examined relevant theories, models, and concepts to build the theoretical basis for interpreting these phenomena in the CAD context. It also identified key dimensions through which engineering designers' performance and CL can be conceptualised, quantified, and categorised. Based on this synthesis, the review clarified the purpose of the method and the theoretical models and derived their general objectives to support their later validation. Finally, it surveyed available methods and tools for measuring and analysing CL and CAD performance across the identified dimensions, including existing experimental frameworks, protocols, and guidelines from engineering, psychology, and neuroscience.

1.2.2 Descriptive study I

The Descriptive study I built on the knowledge gained through the initial literature reviews conducted in Research clarification stage. The body of knowledge was here augmented with reviews of the more specific literature. In particular, the key elements of CAD performance and CL, their characteristics, and the relationships among the elements were identified. Furthermore, gathered insights and results were presented in the form of the overview of the methods and tools for modelling and measuring CL and CAD performance to present the state of art. The overview was accompanied by the advantages and disadvantages of methods and tools, their suitability for the given CAD context, and the challenges of their application for measuring and analysing the CL in CAD activities.

In addition to the literature review, the first empirical study was conducted in this stage. Its findings complemented and extended the initial understanding provided by the literature review, particularly by confirming the engineering information format (representation type) as an influencing factor and by highlighting the need for reliable CL indicators for the CAD context. The study also tested and confirmed the suitability of electroencephalography (EEG) for capturing brain activity during CAD modelling. In parallel, the study confirmed the suitability of human performance modelling for representing engineering designers' CAD performance, and it resulted in the development of initial CAD performance models from empirical data. Finally, the study demonstrated how dynamic changes in CL can be modelled through utilisation-based indicators derived from the activity of perceptual and cognitive operators.

1.2.3 Prescriptive study

Insights gathered through the literature reviews and the first empirical study allowed the development of the theoretical model of CL and the theoretical model of engineering designers' performance in CAD activities as a part of the first Prescriptive study. These models formalised the key elements of CAD modelling, the relationships among them, and the role of influencing factors, including the hypothesised relationships between CL and CAD performance. In parallel, this research stage proposed the method for measuring and analysing CL in the CAD context and specified detailed requirements for the data inputs needed to apply the method. The simplified version of the prescribed method was subsequently suggested in the second Prescriptive study based on findings from Descriptive study II.

1.2.4 Descriptive study II

The understanding gained through Descriptive study I was used to design the second, detailed empirical study. The experimental design (including the setup, procedure, and the methods and tools for data gathering, processing, and analysis) was guided by the findings of the first empirical study and relevant frameworks identified in earlier stages through literature reviews, and it built upon multimodal approaches to studying CAD activities and engineering designers' behaviour. Given the multidimensional character of CL, qualitative and quantitative, as well as analytical and empirical, methods were combined, including subjective assessment (NASA TLX questionnaire), psychophysiological measurement (EEG), performance measurement (e.g. errors, task completion time, CAD model completeness), and human performance (cognitive) modelling. Importantly, the first empirical study had already tested and confirmed the suitability of EEG for capturing brain activity during CAD modelling and the suitability of utilisation-based human performance modelling for representing CAD performance and CL-related changes. These methods were therefore integrated in the second empirical study within the single experimental framework that had been prescribed in the first Prescriptive study.

The results of the second empirical study were then used to validate the proposed theoretical models and the method, and confirm internal task complexity as an additional influencing factor. In addition, the second empirical study compared CL changes implied by EEG-based indicators, utilisation-based indicators, and self-reported subjective scores. Moreover, dynamic changes in CL were modelled in relation to the segments of the CAD modelling process. Finally, CL indicated by weighted utilisations that represent integrated EEG-based and utilisation-based metrics was related to CAD performance assessments to enable confirming/rejecting the research hypothesis.

1.2.5 Validation procedure

Validation of the research outcomes was conducted in accordance with the Validation Square framework [21], [22]. Although the Validation Square is primarily intended for engineering design methods, it can also be used to evaluate research outcomes in general, provided that research results can be assessed in terms of the likelihood of fulfilling their intended application (theoretical structural validity), tests of research usefulness can be assessed for appropriateness (empirical structural validity), research results can be assessed in terms of usefulness for particular applications (empirical performance validity), and inferences can be made about general usefulness of research results (theoretical performance validity) [22]. Since the method and theoretical models proposed in this thesis fulfil these conditions, the Validation Square was applied. Specifically, the research outcomes are defined with an explicit CAD-related purpose, which enables assessment of whether their constructs and structure are likely to support that purpose (theoretical structural validity). The thesis also includes empirical studies that provide concrete datasets and example problems, enabling both evaluation of the suitability of the conducted tests (empirical structural validity) and examination of usefulness in the tested CAD applications (empirical performance validity). Finally, the outcomes are formulated in a way that supports reasoned inferences about broader usefulness beyond the specific experimental tasks by specifying their scope and relating the findings to the state of the art (theoretical performance validity).

Accordingly, validation followed the four-step structure: 1) theoretical structural validation, 2) empirical structural validation, 3) empirical performance validation, and 4) theoretical performance validation [21]. Within this framework, validation is understood as the justification of knowledge claims and as building confidence in a method's or models' usefulness with respect to its purpose. Usefulness is judged through effectiveness and efficiency; effectiveness concerns whether the method/model provides outputs correctly (structural validity), whereas efficiency concerns whether it produces correct outcomes with acceptable operational performance developed and realised with less cost and/or in less time (performance validity). In this thesis, the Validation Square is applied to the prescribed method and to the theoretical models it integrates. For the theoretical models, validation is conducted primarily in terms of theoretical structural validity; where relevant, evidence from the empirical studies is used to demonstrate that the model constructs can be instantiated and used on real datasets (operational/empirical usability). For the prescribed method, validation proceeds through all four Validation Square steps to evaluate the prescribed workflow and inputs, the suitability of the selected empirical tests, the usefulness of the produced outputs for the intended application, and the plausibility of usefulness extending beyond the specific example problems. The evaluation criteria and metrics used across these validation steps are grounded in the purpose, application domain (i.e. the characteristics of design tasks for which the method/model is intended), and objectives [21], which were defined in the Research clarification stage and are specified in the following section.

1.3 Model and method purpose and objectives

The purpose of the theoretical models is to formalise CAD performance and CL by identifying, describing, and representing their key elements, the characteristics of these elements, and the relationships among them. In this thesis, the CAD performance model captures the mechanisms of human performance in CAD modelling as an HCI activity, while the CL model specifies how changes in information-processing demands can be represented and related to the CAD modelling process. The objectives that the theoretical models should meet are summarised in Table 2 and provide the basis for evaluating them within the validation process. Specifically, the models are expected to 1) address relevant performance mechanisms, here referring to the perceptual, cognitive, and motor systems of embodied cognition that underlie interaction with CAD systems; 2) demonstrate usefulness by supporting the investigation of CAD modelling in realistic settings and enabling insights that can contribute to improved understanding of engineering designers' performance; 3) achieve robustness and generality by describing and explaining observed outcomes without over- or under-fitting to a specific task instance; and 4) maintain simplicity, meaning that the models should remain as simple as possible while still providing a useful and robust representation of the targeted mechanisms.

Table 2 Objectives for the theoretical models, according to Wu et al. [23]

Mechanisms	The model addresses the mechanisms of human performance; in this particular case these refer to the mechanisms of perceptual, cognitive, and motor systems of the embodied cognition
Usefulness	The model is useful in real-world CAD system design to improve engineering designers' performance.
Robustness and Generality	The model describes/predicts results without over- or under- fitting.
Simplicity	The model uses the minimum number of parameters necessary to yield a useful and robust description/prediction grounded in human performance mechanisms.

Given this purpose and these objectives, the theoretical models serve as the conceptual foundation for prescribing the method and for interpreting its outputs. The method is intended to measure and analyse CL in CAD modelling tasks in a way that enables its systematic relation to CAD performance, explicitly considering performance both as a process (how designers interact and progress through task segments) and as outputs (the resulting CAD models and their quality-related characteristics). To operationalise this purpose, a set of objectives is specified in Table 3 and organised around functionality, scope, usability, and reliability [24]–[26]. In essence, these objectives state that the method should 1) capture dynamic changes in CL across the CAD modelling sequence and relate them to task segments; 2) rely on a minimal yet robust and sensitive set of metrics and clearly prescribed steps; 3) provide both process-based and output-based performance descriptors without interfering with task execution; and 4) remain applicable to real-world CAD setups while being understandable and reproducible for engineering design researchers. Accordingly, Table 3 provides the baseline against which the method is evaluated throughout the validation process in Chapter 8.

Table 3 Objectives for the outputs of the method and the process by which they are generated

Functionality	The method enables measurement of dynamic changes (in time/sequence) in CL in CAD modelling tasks.
	The method proposes metric(s) for measuring dynamic changes (in time/sequence) in CL in CAD modelling tasks.
	The method proposes metrics for measuring (qualitatively and/or quantitatively) engineering designers' performance in CAD.
	The method prescribes steps for building a model of engineering designers' performance in a CAD modelling task.
	The method prescribes steps for building a model of CL in a CAD modelling task.
	The method prescribes steps for relating CL and CAD performance. The relationship between CL and CAD performance (perceptual, cognitive, motor aspects) should be made in time/sequence when considering the process as one performance aspect.
	A model of CAD performance (developed using the method) describes and integrates perceptual, cognitive, and motor processes (aspects of behaviour/CAD performance) through which engineering designers reach CAD task goals.
	A model of CAD performance describes dynamic changes (in time/sequence) of perceptual, cognitive, and motor processes to represent performance throughout the execution of CAD tasks.
	A model of CL (developed using the method) represents dynamic changes (in time/sequence) in CL throughout the execution of CAD tasks.
	A model of CL represents when the change in CL occurs, in relation to CAD task segments.
	A model of CL represents a magnitude of changes in CL.
	A model of CL consistently varies with influencing factors of interest.
	Measurement of CL must not interfere with engineering designers' performance (execution of CAD tasks).
Scope	The method is suitable for modelling and measuring CL in CAD modelling tasks with predefined characteristics and properties of design components (technical systems).
	The method is suitable for usage with various CAD systems.
Usability	The method is usable by individual engineering design researchers.
	The steps are described at a level of detail suitable and sufficient to be understood by engineering design researchers without prior experience in researching these phenomena (not cognitive psychologists or neuroscientists).
	The method is easy to learn and use. The method opts for the minimum number of metrics through which changes in CL can be reliably measured. The method opts for the simplest metrics through which changes in CL can be reliably measured.
	The method should be usable in and applicable for the real-world CAD setups and tasks.
Reliability	The method is reliable: different users applying the same prescribed steps should obtain same or similar results.

1.4 Scientific contribution

To address the research aims and test the hypothesis, relevant literature findings were reviewed, and synthesised so that key constructs can be adapted to the CAD context and then examined through two empirical studies. This integration and testing resulted in three main scientific contributions of the thesis:

1) **A theoretical model of cognitive load (CL) in CAD activities**, which formalises how information-processing demands can be represented in relation to the CAD modelling process and its segmentation, and how an influence of task-related factors central to CAD can be considered when modelling CL.

2) **A theoretical model of engineering designer's performance in CAD activities**, which formalises CAD performance through information-processing and embodied-cognition lenses, specifying key elements, a hierarchical structure for multi-level process and CAD model analysis, and relationships between performance descriptors of CAD modelling processes and outcomes.
3) **A method for measuring and analysing CL in CAD activities**, which operationalises the two theoretical models as an end-to-end, multimodal workflow combining cognitive performance modelling (operator utilisations), psychophysiological monitoring (EEG-based CL indicators), and subjective assessment (NASA TLX), enabling dynamic, segment-level CL estimation and analysis of a relationship between CL and performance in realistic CAD modelling settings.

1.5 Thesis structure

The thesis is divided into eight chapters that mostly follow the stages of the adopted research methodology presented in Table 1 and described in Section 1.2. After introducing the motivation and research aims in Chapter 1, the thesis progresses from establishing the state of the art, through model and method development, to empirical testing, refinement, and validation.

Chapter 2 presents the research background and related work underpinning the thesis. It reviews how CAD performance has been conceptualised and measured, and summarises key factors influencing engineering designers' CAD performance. The chapter then introduces modelling of human performance in complex HCI tasks by reviewing relevant cognitive architectures and task-analysis frameworks, including cognitive models proposed for CAD. Next, it reviews theoretical models of CL, approaches for measuring CL, and factors affecting cognitive demands in CAD tasks. Finally, it synthesises current knowledge on relationships between CL and performance in engineering design and CAD, identifying gaps that motivate the research questions and subsequent work.

Chapter 3 reports the first empirical study (as a part of Descriptive Study I), which provides additional evidence relevant to the thesis assumptions, corroborates key insights from the literature review and addresses the feasibility questions raised therein. In particular, it serves to test the suitability of the individual methods and approaches later integrated in the prescribed method (e.g. feasibility of the experimental setup, suitability of candidate metrics for capturing dynamic changes in CL in CAD, and practicality of analysing CAD activity through the proposed modelling and segmentation logic). In this way, the study supports early methodological decisions that informed the theoretical models and the method development.

This chapter marks the end of Descriptive study I; the subsequent chapters follow an iterative development cycle in which models and the method are proposed, operationalised in the second empirical study, and then refined and validated.

Chapter 4 introduces the initial theoretical models of engineering designers' CAD performance and CL in CAD activities. Both constructs are described as functions of four core elements (engineering designer, CAD task, CAD system, and environment), grounded in prior literature on cognitive architectures, human performance models, and CL theory. The chapter formalises the model elements and relationships which provide the basis for the step-by-step method prescribed in the following chapter.

Chapter 5 prescribes the method for measuring and analysing CL in CAD activities. It specifies the required inputs, processing steps, and outputs. It integrates 1) cognitive modelling (operator utilisations) to represent CAD performance processes, 2) EEG-based indicators to reflect information-processing intensity, and 3) NASA TLX as a complementary subjective reference. Together, the suggested steps provide explicit inputs, processing logic, and outputs that can be executed end-to-end and later operationalised and evaluated in the second empirical study.

Chapter 6 reports the second empirical study (as part of Descriptive Study II), which implements the prescribed method in a CAD context and evaluates its key outputs. In particular, the study supports identification of EEG-based CL indicators suitable for CAD tasks, validates the practical implementation of the CAD performance model and utilisation calculations, and examines how EEG-based, utilisation-based, and NASA TLX measures relate to each other and to CAD task segments. Finally, it investigates whether the resulting CL indicators are meaningfully associated with CAD performance outcomes, providing evidence relevant to testing the research hypothesis.

Chapter 7 integrates validation and discussion of the research outcomes. Building on the validation criteria defined in Section 1.3 (objectives and main purposes), it evaluates the theoretical models and the prescribed method through theoretical and empirical structural and performance validity, following the Validation Square framework [21]. The chapter also verifies the research hypothesis, reflects on the extent to which the research objectives are achieved, discusses implications, and outlines limitations and future work.

Chapter 8 concludes the thesis by summarising the main contributions, consolidating key findings across models, method, and empirical studies, and proposing venues for further work.

This page intentionally left blank

2 RESEARCH BACKGROUND AND RELATED WORK

The second chapter provides the research background and related work that underpin the thesis. It first reviews how CAD performance has been conceptualised, explored, and measured, and summarises the key factors that influence engineering designers' performance in CAD. The chapter then turns to modelling human performance in complex HCI tasks, introducing relevant cognitive architectures and task-analysis frameworks, and outlining existing cognitive models that have been proposed specifically for the CAD context. Next, it reviews theoretical models of CL, approaches for measuring CL, and factors that influence engineering designers' cognitive demands during CAD tasks. Finally, the chapter synthesises current knowledge on the relationship between CL and CAD performance outcomes, highlighting how these constructs have been linked in prior studies. Together, these sections establish the theoretical foundation for the thesis and reveal gaps in the literature that motivate the subsequent research.

2.1 Literature review process

The literature search was conducted using scientific databased including Scopus, Web of Science, and Google Scholar. The main search was conducted during the Research Clarification stage between 2020 and 2022, with additional targeted searched performed during thesis writing in 2025 to include relevant recent publications and to refine specific theoretical and methodological aspects. The search included peer-reviewed journal papers, conference papers, books, and book chapters. Search terms included combinations of keywords related to cognitive load, cognitive workload, mental workload, cognitive ergonomics, CAD performance, design performance, engineering designers' performance, human performance in computer-aided design, cognitive load measurement, performance measurement, and psychophysiological assessment. Examples of search strings included "cognitive load" OR "mental workload" OR "cognitive workload"; "cognitive load" AND "theory" OR "model"; "cognitive load" AND "measurement" OR "metric" OR "analysis"; "cognitive ergonomics" AND "engineering"; "CAD performance" OR "computer-aided design performance"; "design performance" OR "engineering designer performance"; and "human performance" AND "computer-aided design".

Publications were included when they addressed at least one of the following topics: CAD performance, engineering design performance, human performance in HCI tasks, cognitive load theory, cognitive ergonomics, or methods for measuring and analysing CL or performance in design-related contexts. Priority was given to peer-reviewed sources, highly cited theoretical contributions, and studies directly related to CAD, engineering design, HCI, or psychophysiological measurement. Publications were excluded when they did not address human performance, CL, CAD, engineering design, or measurement and analysis methods relevant to the thesis scope. Sources focused exclusively on CAD software functionality or design outcomes without a human performance or CL perspective were excluded unless they provided relevant contextual information. To ensure broader coverage, backward and forward citation tracing was applied to highly relevant publications. The review was not intended as a formal systematic review, but as a structured narrative review supporting model development and method formulation.

2.2 CAD performance

Engineering designers perform CAD through the process divided into activities during which requirements and descriptions of a design problem (input information) are progressively transformed into output information in the form of design representations [8]. These representations are information objects that encode design characteristics (e.g. form, position, and dimensions) and design properties (e.g. function) [2], [8]. CAD performance, therefore, can be viewed through a process lens as the effectiveness and efficiency with which engineering designers carry out such information transformation across activities [27].

To describe how engineering designers move from input information to successive CAD representations, prior work has proposed a limited number of process-oriented design models that can be, at least partially, reused for modelling CAD performance. In line with the process perspective adopted in this thesis, Ishino and Jin proposed a *three-layer design process model* and associated method based on data captured by a CAD system [28]. The model distinguishes: 1) an *Event Layer* capturing primitive-level design events generated by engineering designers using a CAD system; 2) an *Operation Layer* representing higher-level design operations as clusters of events that reflect design actions; and 3) a *Product Model Layer* representing the design alternatives generated through sequences of design operations [28]. The accompanying method aims to automatically identify operations and their relationships with product models from primitive design events, while highlighting correct event-to-operation clustering as the key challenge due to the large number of possible event combinations [28].

Building on this foundation, Sadeghi et al. augmented the proposed three-layer view into a *multi-layered process model* that explicitly separates *Product Layer* and *Process Layer* information [29]. The *Product Layer* is decomposed into the *Object Layer*, *Product* *forms/features Layer*, and *Product Model Layer*, while the *Process Layer* distinguishes activities occurring inside versus outside the CAD system [29]. The *Process inside the CAD* is further structured into a *Software interaction operation layer*, a *Low-level design operation layer*, and a *High-level design operation layer* [29]. The authors complemented this framework with *Product Data* and *Process Data Models* intended to structure the data required for analysis [29].

These frameworks pave the way for translating the concept of how design information is progressively transformed throughout the CAD modelling process into analysable structures by linking what engineering designers do in CAD activities (HCI events and modelling operations) to what they produce (evolving design representations) and enabling process descriptions across levels of abstraction. In this sense, they make an important contribution by proposing a general decomposition of CAD modelling activity into layers; however, they do not provide a fixed, reusable list of actions or operations within each layer, as these are treated as dependent on the application scenario [29]. Finally, none of these frameworks was developed as a dedicated CAD performance model, which motivates the following overview of CAD performance studies. The next sections review how CAD performance has been explored in prior work; Section 2.2.1 outlines the main foci covered by CAD performance studies, while Section 2.2.2 summarises how CAD performance has been described, measured, and visualised through different metrics.

2.2.1 Studies on CAD performance

The reviewed literature includes a range of studies that investigated at least one aspect of engineering designers' CAD performance, considering both performance outputs and processes through they were reached. Their key methodological details are summarised in Table 4, which reports the experimental task, the independent variables examined (where applicable), and the metrics used to quantify CAD performance in each study. Overall, the studies illustrate how different factors can shape CAD performance, although most focus on either outputs or processes rather than both. To support synthesis, the findings discussed in the sections following Table 4 are organised according to whether the independent experimental variables relate to 1) engineering designers' characteristics, 2) methods and tools, or 3) task characteristics.

Table 4 List of studies on engineering designers' CAD performance

Reference	Experimental task	Independent variable(s)	Metrics
[30], [31]	2D part modelling	Course on modelling strategies	Number of strategies implemented, command usage, comparison time, errors
[32]	Solid part modelling	Strategy and spatial visualisation lessons	Planning time, working time, N of models, N of expert strategies used
[33]	Design	Engineering experience	Time series, cumulative time plots of segments, progress time plots, quality of solutions compared to requirements
[34]	Solid part modelling	Technical fundamentals	N of changes
[35]	Solid part modelling	-	N of features, N of sketches per features, duration, errors (accuracy)
[36]	Solid part modelling	Learning style	Duration, Speed, N of features, rate of features
[37]	2D part modelling	Cognitive style	Creativity and technical quality judged by experts, model correctness and completeness
[38]	Solid part modelling	Graphical capability	Sketch plane selection, origin selection, model orientation, symmetry planes, sketch relations, fully defined sketches, feature sequence, feature relations, part design intent
[39]	Solid part modelling	Spatial visualisation ability	Quality rubrics: base feature, sketch plane orientation, origin, sketch simplicity, sketch definition, design intent, feature end conditions, symmetry/duplication, accuracy/completeness
[40]	Solid part modelling	Alternative goals (ease of alterability, speed)	Model quality (N of features and entities, feature sequences, weak dimensions), duration, feature organisation, alterability
[41]	Solid part modelling	Alternative goals (ease of alterability, speed)	Alternation duration, alternation procedure, list of attributes (e.g. sketch plane, origin, feature sequence, N of features, reference geometry)
[42]	Solid part modelling	Spatial visualisation ability	Approach, structure, accuracy, robustness, creativity
[43]	Design	Gender	Time series of segments
[44]	Solid part and assembly modelling	Internal task complexity	Model quality dimensions (rubrics)
[45]	Solid part modelling	-	N and proportion of segments, segment sequency, duration, frequency, transitions
[29]	Solid part modelling	-	Duration, N of segments (operations), N of features
[46]	Assembly modelling	Engineering information format	N of errors, duration, rates of rework, direct and indirect work
[47]	Solid assembly modelling	Multi-user vs single-user Spatial visualisation ability	Duration, communication, individual effectiveness/time, value added per time unit
[48], [49]	Design	-	Sequences of segments
[50]	Design	Two different CAD systems	N of solutions, duration, N and distribution of approaches taken to solving problems
[51]	Solid part modelling	Individual vs teamwork	Expert judgments on quality, communication patterns, keystrokes/mouse clicks timeline
[52]	Solid part modelling	Training approach for strategic knowledge	Model quality dimensions (rubrics): complete, valid, consistent, concise, conveys design intent
[53]	Solid assembly modelling	-	Assembly quality dimensions (rubrics): valid, complete, consistent, concise, clear, conveys design intent
[54]	Solid part modelling	Collaboration with shared vs. parallel modelling control	Model quality dimensions (rubrics): complete, concise, consistent
[55]	Solid part modelling	Spatial visualisation ability	Model quality dimensions (rubrics)
[56]	Solid part modelling	-	Segment sequences, sequence patterns
[57]	Solid part modelling	-	Completeness, N of features, feature sequence
[58]	Solid part modelling	CAD experience	Model completeness and correctness based on dimensions, segment timelines and transitions
[59]	Design	-	Segments (CAD actions), Contribution of team members to each segment
[60]	Solid part modelling	Working styles: individual, pairs sharing control, pairs working simultaneously (individual control)	Model quality dimensions (rubrics), speed, communication patterns, user interaction (cursor activity), satisfaction

[12]	Solid part modelling	Internal task complexity	Model and process variability, file size, feature types, N and proportion of segments, segment transitions
[61]	Solid part modelling and optimisation	Modelling techniques Internal task complexity	CAD model quality and reusability
[62]	Solid assembly modelling	Team composition Gender, Experience	Building/Revision ratio, Segments from [45], Distribution of segments
[63]	Design	High vs low performing teams	Segment distributions, transitions
[64]	Design	Design review Immersiveness Internal task complexity	N and proportion of actions (creation, revision)
[65]	Design, Solid part modelling	ChatGPT vs no ChatGPT	Completeness

2.2.1.1 Engineering designers' characteristics

An influence of engineering designer's characteristics has been examined most often through **spatial visualisation ability**, with multiple studies reporting a significant, positive relationship with CAD performance; designers who scored higher on spatial tests (e.g., Purdue Spatial Visualisation Test [66]) tended to perform better in CAD modelling tasks [39], [42], [47], [55]. Beyond visuospatial ability, scholars linked CAD performance to **technical and graphical fundamentals**; mathematics, computer science and engineering, CAD methodologies, graphics, and mechanical design fundamentals showed correlations with at least some CAD performance metrics (most strongly for mathematical foundation), while graphical capability in plane and solid geometry was also positively associated with CAD performance [34], [38]. Other work related individual differences in **learning and cognitive styles** with CAD knowledge acquisition and CAD modelling performance (e.g. [36], [37]). In addition, studies comparing performance based on **level of expertise** (e.g. between novices and experts) highlighted differences in the CAD modelling processes (e.g. experts allocated more time to problem definition and information gathering [33]; higher creation/revision ratios were found among experts than novices [58]), even when output-based metrics such as number of features used did not significantly differ. Finally, research on **gender**-related differences reported contrasts in CAD modelling tendencies; males producing more complex models that did not necessarily meet the design specifications, and females adhering more closely to specifications and spending more time on revisions [43], while higher creation/revision rates were observed among the latter, although these implications remained tentative due to limited sample size [62].

2.2.1.2 Methods and tools

Within the stream of studies focusing on methods and tools, scholars examined how design representations, CAD system capabilities, CAD modelling strategies**,** methodical interventions (including design reviews and lessons on CAD modelling strategies), and AI support influence

CAD performance. Studies related to **design representations** indicated that the medium used to communicate engineering information can significantly affect CAD modelling performance. For instance, Sweany et al. compared modelling in 2D CAD, 3D CAD, and using 3D-printed physical models, reporting the best performance when working from the latter and the worst in the 2D CAD setup [46]. Other work focused on **CAD system capabilities** by comparing performance in multi-user and single-user CAD, generally reporting benefits for collaboration (e.g. increased awareness and communication between users), but mixed evidence for quality of performance outcomes; some studies found no differences [51], others reported slightly worse performance in team-based multi-user settings [47], [60], whereas Arshad et al. reported significantly better overall performance in the multi-user setup [54]. Beyond collaboration mode, scholars compared parametric and freeform modelling CAD systems and showed that engineering design students employ different CAD modelling behaviours depending on the system type [50]. Studies addressing **modelling strategies** further indicated that strategy effectiveness depends on task demands; explicit reference and resilient strategies were robust for simple parts but decreased in effectiveness as CAD model complexity increased, while horizontal modelling was associated with significantly lower performance [61]. In addition, research on **methodical interventions** suggested that structured changes to the CAD modelling workflow can shape subsequent CAD activity. For example, design reviews conducted in virtual environments of different immersion levels and generated by different media (traditional desktop interface vs. virtual reality) were associated with measurable differences in subsequent CAD modelling behaviour (e.g. slightly more CAD actions and a higher proportion of creation actions after reviews in virtual reality) [64], while structured lessons on CAD modelling strategies consistently enhanced CAD performance across multiple studies [30]–[32], [52]. Finally, recent studies on **AI support** indicated that ChatGPT may hinder performance in collaborative CAD-specific tasks reliant on CAD knowledge and does not necessarily improve the quality of the CAD modelling outputs [55].

2.2.1.3 Task characteristics

Finally, research addressing task characteristics has mainly examined how task goals and task complexity shape CAD performance. Studies comparing alternative **task goals** (e.g. modelling for ease of editability vs. modelling as fast as possible) reported similar CAD model attributes (e.g. sketch planes and CAD features) and CAD modelling processes across task conditions [40], [41]. However, when editability was evaluated directly, models produced under the editability goal tended to be easier to modify, even if the CAD modelling duration, output, and process differences were small [41]. Beyond task goals, several studies noted that the **internal complexity of CAD tasks** likely affected observed results and conclusions, yet complexity was not treated as the primary manipulated variable [44], [61], [64], [67]; accordingly, its influence remains insufficiently explored and asks for future systematic research to provide comprehensive results.

In summary, prior work indicates that CAD performance is shaped by factors related to the engineering designer, the employed methods and tools, and the task itself, but the strength and maturity of evidence differ across these streams. The most consistent relationships are reported for variables related to an engineering designer (especially spatial visualisation ability), whereas studies related to methods and tools provide broader but more context-dependent findings, including evidence that design representation format can affect CAD modelling outcomes. Task-related research remains comparatively limited, and although task complexity is frequently suggested as an important driver of performance differences, it has rarely been treated as a primary, systematically manipulated factor in prior experimental studies.

The following section shifts attention from what has been studied to how CAD performance has been operationalised in the literature. Specifically, it reviews how CAD performance has been described, measured, and visualised, and summarises the metrics used to quantify and compare both performance outputs and processes in previous studies.

2.2.2 Describing, measuring, and visualising CAD performance

Performance in engineering design is commonly assessed through the concepts of performance effectiveness (referring to the extent to which desired outcomes are achieved) and efficiency (referring to the resources invested in the process of producing those outcomes) [27]. Consistent with this distinction, CAD performance has typically been observed considering two complementary aspects: 1) the quality of resulting outputs (the design representation) and 2) the characteristics of the CAD modelling process through which these outputs are produced, reflecting how (well) the activities are carried out. The following sections therefore outline the metrics used in prior studies to quantify each of these aspects of CAD performance.

2.2.2.1 Metrics for quantifying CAD modelling outputs

The output of CAD modelling is a 3D CAD model; a digital representation of a design (its characteristics and properties), built as a combination of sketch entities and CAD features, constrained dimensionally and geometrically, and positioned relative to the origin of the virtual

3D environment [2], [68]–[70]. In CAD performance research, outcomes are typically evaluated through the concept of CAD model quality. However, while CAD model quality is frequently mentioned and examined, its definition as well as measurement are implicitly considered and under-discussed in prior work.

To provide a structured definition and operationalise CAD model quality, Company et al. proposed six quality dimensions for primarily solid part models: validity, completeness, consistency, conciseness, simplicity, and effectiveness and efficiency in conveying design intent [44]. Building on this work, Otey et al. proposed a comparable list of quality dimensions (rubrics) tailored to assembly models [53]. In both versions, the identified dimensions were introduced mainly for educational settings, framed as rubrics that students can use to create high-quality CAD models [44]. Several empirical studies have adopted these rubrics or metrics aligned with them. For example, Phadnis et al. adapted the rubrics to their experimental tasks and used them to compare CAD model quality in individual versus team-based setups [60]. Another example is the study of Šklebar et al., who compared completeness of CAD models generated by teams of engineering designers in collaborative design with and without ChatGPT support [65]. More broadly, many studies use output metrics that map onto these dimensions even when the rubrics are not explicitly cited. Most commonly, completeness is quantified as the accuracy in replicating a target component's size and shape without a clear reference to completeness as defined by Company et al. [44] (e.g. [30], [34], [42]). For example, Bhavnani et al. quantified the model quality through the number of errors related to style (e.g. incorrect line thickness, font style or font size), dimension (e.g. an element not drawn to scale), and form (an element does not meet or line-up with another, a line that is not straight, or an arch which is not correctly drawn) [30]. With the similar goal of providing a structured and reliable way to assess CAD performance, Branoff et al. proposed eight rubrics covering aspects such as identifying the base/core feature, selecting an appropriate sketch plane orientation and origin, producing simple and fully constrained sketches, applying appropriate features and symmetry/duplication, achieving accuracy/completeness, and CAD modelling strategy efficiency [71].

In most cases, CAD model quality is assessed manually, but there have been efforts to automate its evaluation (e.g. [72]–[74]). For instance, Kirstukas developed a computer program that detects issues and reports CAD model quality scores across twelve categories, including units, unconstrained sketches, auto dimensions, missing/repeated/unwanted dimensions, banned constraints, unused sketches, non-united bodies, shape and size, orientation, absent solids, negative dimensions, and missing drawing dimensions [73]. Similarly, Jaakma and Kiviluoma proposed an auto-assessment tool that reports 1) geometry assessment by analysing the shape and comparing it with the expected outcome and 2) design intent assessment by changing the parameters and observing how introduced changes affect CAD models [74].

These available tools were primarily developed for educational contexts to support CAD teachers in evaluating large volumes of student work. In practice, many metrics related to evaluating the CAD outcome (e.g. volume or shape) require a "gold standard" reference model for comparison, along with user-defined grading weights and acceptable discrepancy thresholds. As currently implemented, such tools mainly support evaluation of quality dimensions that are comparatively straightforward to operationalise, especially validity and completeness (e.g. via size and shape), and in some cases editability, assessed as the model's ability to remain correct after parameter changes [73], [74]. The assessment becomes substantially more challenging for dimensions that are inherently less direct to quantify and compare with the "gold standard", such as effectiveness and efficiency in conveying design intent [44]. Overall, existing tools facilitate faster extraction and comparison of CAD outcome data, but remain limited in their ability to compute comprehensive, transferable CAD performance metrics across the full set of model quality dimensions. Outside educational settings, a central challenge is defining metrics and acceptable values that reflect the expected performance level for a given CAD modelling scenario. Given the high variability of CAD model types and CAD modelling goals, such definitions are unlikely to be universal and typically need to be tailored to the application context.

To conclude, metrics related to CAD modelling outputs are widely used as they are intuitive and can be extracted from a tangible deliverable of CAD modelling – the final CAD model. However, they provide only a snapshot of the end state and offer limited insight into how engineering designers arrived there, how efficiently they progressed, or where difficulties and revisions occurred during CAD modelling. As a result, output metrics alone can miss important differences in CAD modelling behaviour and are often insufficient for characterising CAD performance as a dynamic process. For this reason, prior research has increasingly complemented outcome measures with process-oriented metrics that describe how CAD models are created; these are reviewed in the next section.

2.2.2.2 Metrics for quantifying and describing CAD modelling processes

A typical CAD modelling process consists of several overt steps in which the (imagined or interpreted) geometry is first sketched in 2D on the selected plane using sketching entities (e.g. line, circle, etc.) that are mutually connected by sketch relations (dimensional or geometric) and then transformed into volumes by CAD features (e.g. extrude boss or cut, revolute, chamfer, etc.). The range of possible combinations of sketches and CAD features, resulting in visually equal models, is virtually unlimited [38], [52], [61]. The variability in CAD modelling process has been related to prior knowledge of users, the way in which they were taught about best practices, and experienced best practices [12]. The unlimited range of possibilities makes it challenging, in most cases, to propose one universally superior process that leads to the best performance outcomes [29]. Consequently, structured measurement and assessment of the CAD modelling processes is challenging as well.

2.2.2.2.1 Output-based metrics for describing and comparing CAD modelling processes

Scholars have considered several cumulative metrics (e.g. CAD modelling duration, number, type, and order of CAD features) to quantitatively describe CAD modelling processes, most often by extracting this information from the final feature tree available in CAD software. These metrics have then been used to relate CAD modelling processes to CAD modelling outcomes (described in Section 2.2.2.1) and to test the effects of various factors (described in Section 2.2.1), typically motivated by the aim of prescribing structured ways of building high-quality CAD models. For example, Rosso et al. explored CAD modelling process variability by analysing the type and order of CAD features selected by participants, and related this variability to CAD model editability (ability to alter the geometry easily) and reusability (ability to reuse existing geometry in other contexts) [12]. Similarly, Otto and Mandorli linked the number and frequency of CAD features to CAD model completeness [57], while Chen used the number of CAD features to distinguish between intermediate and expert CAD users [58]. Bhavnani focused on CAD modelling duration as a comparison metric and associated better performance with cases in which fewer features were used, regardless of their complexity, arguing that such models can be produced faster [31]. In contrast, Rynne and Gaughran [35] as well as Chester [32] emphasised CAD modelling approaches based on simpler sketches to support easier editing.

Beyond metrics based on CAD features, some studies have examined CAD performance at the level of sketch entities and relations. For instance, Company et al. analysed the CAD modelling process at the sketch level by investigating associations between sketch relations and CAD model reusability [75]. However, compared with CAD features, sketch entities and relations have been examined less frequently, likely because they are not directly available from the feature tree. Their extraction and interpretation therefore require a more fine-grained description of CAD modelling activity than the cumulative summaries typically provided by the feature tree. Accordingly, the following section introduces segment-based approaches for analysing CAD modelling processes, which may offer the resolution needed to capture and interpret entities and relations at the sketch level.

2.2.2.2.2 Segment-based metrics for describing and comparing CAD modelling processes

Studies reviewed in the previous section indicate that CAD performance has often been quantified using output-oriented, cumulative metrics (e.g. CAD modelling duration, number and type of features). While useful, such measures still largely reflect the final stored CAD model and its feature tree, and therefore provide limited access to the CAD action history through which the model was created. To examine CAD performance through the process lens, alternative approaches included segmenting the CAD activity using coding schemes and then analysing the extracted segments.

Across this line of work, researchers have used three main approaches to capture and analyse large amounts of CAD data related to CAD modelling processes: 1) audio/video recording with subsequent manual coding (e.g. [33], [50]); 2) custom data logging or dedicated software for CAD research (e.g. [43], [76]) cloud-based CAD systems with built-in functions for recording and storing CAD data related to process (e.g. [47], [60], [62], [63]).

When segmentation relied on manual coding, segments were typically defined at a relatively high level of abstraction. For instance, Atman et al. decomposed the task into problem definition, gathering information, gathering ideas, modelling, feasibility analysis, evaluation, decision making, and communication [33]. At the similar level, Buckley et al. divided the task into identifying, planning, implementing, and evaluation, and analysed these four phases to distinguish between different approaches to solving design problems (persist, explore, alternative, abandon, seek assistance, assistance given) [50].

In contrast, studies that developed custom scripts for CAD data logging purposes typically provided larger volumes of more fine-grained data [43]. Consequently, this enabled them to investigate processes at the lower, more detailed level that divided tasks into shorter segments, often referred to as CAD actions, events, commands or operations, and used interchangeably across the studies. For example, Gopsill et al. investigated the potential of analysing CAD logs to provide insights into the engineering designers' behaviour in design activities by analysing CAD command use (e.g. creating, editing, constraining), proportions of command types, and transitions between commands [45]. Bhavnani et al. grouped commands into setup, translation, modify, and creation categories and analysed metrics such as the number of completed commands [30]. Rosso et al. defined ten event types (transition, viewing, assembly, deleting, reversing, editing 3D, creating 3D, editing 2D, creating 2D, and constraining) and compared cumulative counts and proportions to capture process variability [12]. Kitajima et al. analysed CAD logs, with focus on command type count (e.g. insert sketch, edit sketch, add component, etc.) and time gaps between them to extract information about user actions and relate them to CAD system know-how [56]. At the similar level of detail, Chen focused on audit trail actions; they were divided into 1) actions with derivable time durations (open and close a drawing, add a new sketch, add a new part studio feature, edit an existing sketch, edit an existing part studio feature, cancel editing a sketch/feature) and 2) actions without time durations (delete a sketch/part studio feature, suppress a feature, rename a sketch/part studio feature, show/hide a sketch/feature, move feature/rollback bar, suppress/unsupress feature, create folder) [58]. Other studies similarly operationalised CAD process at the command/action level through command/action groupings that extend beyond CAD modelling commands to include behaviours related to interaction and collaboration (e.g. communication, help menu access) in synchronous cloud-based CAD tasks [60], [77].

The emergence of cloud-based CAD environments further enabled continuous and non-intrusive (to the process) collection of CAD data [78]. This supported analyses at particularly high granularity in realistic settings. For example, Leonardo and Olechowski analysed industry data from a product development project involving professional designers, grouping logs into 79 unique CAD actions (e.g. copy-paste sketch, rename folder, restore document from history, delete workspace, etc.) and clustering them into 14 action groups (Sub-Part, Part Studio, Assembly, Review, Overhead Change, View/Scan, Create/Edit, Element, Delete, Upgrade, Export, Comment, Add Parts, and Branch) [59]. In a similar direction, Celjak et al. used cloud-based CAD logs to examine proportions of creating (part/assembly), editing (part, assembly, non-geometry), deleting (part/assembly), reversing, viewing, and organizing (design, support design process) in collaborative tasks and related these to team performance [63]. Chen analysed audit trail actions, separating those with derivable durations (open and close a drawing, add a new sketch, add a new part studio feature, edit an existing sketch, edit an existing part studio feature, cancel editing a sketch/feature) from those without (delete a sketch/part studio feature, suppress a feature, rename a sketch/part studio feature, show/hide a sketch/feature, move feature/rollback bar, suppress/unsupress feature, create folder) [58].

In some cases, segmentation was tailored to the specific CAD system or experimental task. McComb et al., for example, analysed sequences of predefined design operations available in their CAD environment (adding/removing a joint, adding/removing a number, changing the size of a single member/all members, moving a joint) to support simulation of team performance [48], [49].

Xie et al. segmented CAD actions into building (e.g. build a foundation, build a wall), revising (e.g. revise a foundation, revise a wall), and switching (open another design or template) categories to study engagement and gender differences in design behaviours [43].

Beyond quantitative comparisons of frequencies and distributions, several authors emphasised that process understanding benefits from visualising what designer did and when they did it [33], [58]. Here, extracted task segments are often represented using timeline visualisations, where time progresses from left to right and segment occurrences are displayed as intervals on separate horizontal tracks. Atman et al. used timelines to qualitatively contrast the CAD modelling processes of low-, average-, and high-performing students [33]. Building on this approach, Chen used similar "event plots" to visualise audit trail data and to search for patterns indicative of CAD modelling behaviour (e.g. forward planning), signs of inexperience in CAD modelling, and methodical vs. disorganised work [58]. Similarly, Rosso et al. visualised transitions between segments in which they divided CAD modelling processes (viewing, assembly, deleting, reversing, editing 3D, creating 3D, editing 2D, creating 2D, and constraining) to compare them among the participants [12]. In that sense, visualisation complements summary statistics by revealing the temporal organisation of CAD modelling activity and making process differences more transparent and intuitive to examine.

2.2.3 Conclusions of the literature review on CAD performance

The reviewed literature provides valuable but fragmented foundations for understanding engineering designers' CAD performance. Research attention has been unevenly distributed across factor groups; variables related to an engineering designer (especially spatial visualisation ability) have been examined most frequently and yield the most consistent evidence, whereas studies on methods and tools include a wide range of variables and therefore produce findings that are often context-specific and difficult to compare. Task-related factors have been addressed least, and despite repeated acknowledgement that task complexity shapes CAD performance, it has rarely been explored as a primary experimental variable. Taken together, these findings guide the thesis later decision to focus on design representation format (in the first empirical study) and task complexity (in the second empirical study) as practically relevant determinants that remain insufficiently consolidated in prior work, particularly when CAD performance is interpreted through a process lens.

A second key conclusion concerns how CAD performance has been operationalised. The literature employs a wide range of metrics across different levels of analysis, reflecting diverse study foci but also indicating limited consistency in defining CAD performance. Metric selection is strongly influenced by data availability; early work largely relied on outcomes and on cumulative process descriptors derived from what is readily accessible in CAD systems, most notably the resulting 3D geometry and the feature tree. While outcome metrics are essential for assessing effectiveness, they provide limited insight into how performance unfolds over time. In contrast, fewer studies have decomposed CAD modelling into lower-level segments and visualised CAD activity temporally (e.g. via timelines), even though such approaches are crucial for revealing the sequence, structure, and evolution of CAD modelling.

The review highlights methodological challenges that directly affect the comparability and reusability of prior findings. Studies vary widely in segmentation granularity and often use terms such as commands, events, actions, and operations interchangeably, making analyses difficult to replicate and results difficult to integrate. This points to the need for a clearer and more transferable process structure that can support consistent segmentation, mapping, and interpretation across studies.

Finally, the literature review showed that CAD performance can be understood as a function of 1) the engineering designer, 2) the CAD task, and 3) the CAD system employed. Yet most CAD performance studies consider only one or two of these elements at a time, limiting explanatory power and synthesis [78]. This motivates the next section of the literature review, which introduces human performance modelling as a methodological approach capable of integrating these elements within a single framework and reviews related work that can inform CAD performance modelling.

2.3 Human performance modelling in complex HCI tasks

HCI research examines how people perceive information, make decisions, and act through interactive systems, and how the properties of the user, task, and interface jointly shape performance. In this field, human performance is typically treated not only as the quality of the final output, but also as the observable and explainable behaviour that produces it; i.e. the sequences of perceptual (P), cognitive (C), and motor (M) operations of embodied cognition through which users transform information into actions [7]. This perspective complements the CAD performance literature reviewed in the previous sections. CAD studies have often quantified performance through outputs (e.g. model completeness or editability) and through cumulative process proxies derived from readily available CAD data (e.g. modelling time, feature counts, feature-tree summaries). While valuable, such measures provide limited explanation for why CAD performance differs across engineering designers, tasks, and representation formats, and they often lack a consistent basis for comparing results across studies due to differences in segmentation level and terminology. Human performance modelling offers a way to address these limitations by providing a unifying framework that explicitly links observable CAD actions to underlying information-processing demands, thereby supporting more interpretable, comparable, and potentially predictive accounts of CAD performance, particularly at a finer temporal and behavioural granularity. Importantly for this thesis, it also provides a coherent basis for linking CAD performance to CL; by modelling performance as resource-limited information processing over time, changes in CAD modelling process structure and P, C, and M operator demands can be interpreted in terms of varying information-processing requirements, creating a bridge between what engineering designers do, how well they do it, and how demanding it is.

Performance models are abstractions used to characterise, understand, predict, and optimise human performance in various tasks [79], [80]. They either focus on outputs of human performance (i.e. models that produce the same outputs as humans) or on the processes involved in human performance (i.e. models that simulate the processes humans use to reach outputs) [79]. The further overview as well as the rest of the thesis focuses on the latter type of models because CAD performance unfolds over time as an interaction process; engineering designers perceive design information, interpret and manipulate it, and execute CAD actions that gradually construct a 3D model as a digital representation of a design [2], [6], [81]. Process-oriented performance modelling therefore not only supports richer explanations of CAD behaviour, but also enables a structured representation of information-processing demands that can be related to CL. Moreover, it enables analysis at a granularity that is typically finer than in most CAD performance studies; down to the sequences of P, C, and M operators, mapped to CAD actions, while still allowing aggregation to higher-level segments when needed.

Building the model of human performance requires 1) a cognitive architecture and 2) a task analysis framework [79], [82], as visualised in Figure 1. Cognitive architectures describe users' embodied cognition by specifying how information is processed and stored, whereas task analysis provides a systematic decomposition of user's goals and the methods used to achieve them in interaction with a computer system [83]. Together, they enable performance to be described as a structured mapping (through the set of operators) between what the user is trying to achieve and how the user's information-processing system can achieve it under constraints. In the context of this thesis, this mapping is especially valuable because it provides an explicit way to associate segments of CAD activity with their expected processing demands, which can then be examined alongside empirical measures of CL.

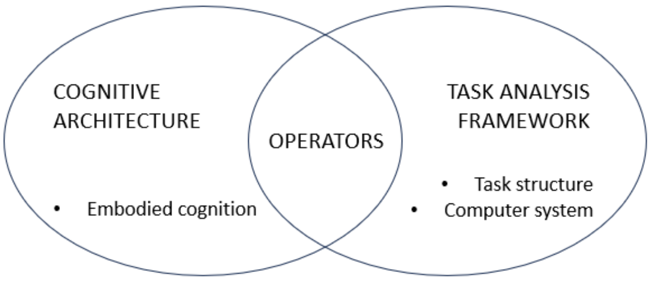

Figure 1 Building bricks of human performance modelling

Performance models are often framed as stage models of human information processing; they are based upon an assumption that human cognition and behaviour can be successfully analysed in terms of stages to which operators that transform inputs into outputs are mapped [84]. In the conventional formulation, external stimuli are first assumed to be processed by a perceptual (P) system; perceived information in then passed to a cognitive (C) system that interprets and manipulates it; finally, responses are executed through a motor (M) system. Cognitive architectures differ in how they organise these systems and in whether they assume strictly serial operation (implying largely sequential task execution) or allow conditionally parallel operation (enabling overlapping goals and concurrent processing). This distinction is particularly relevant for CAD, where engineering designers frequently interleave perception (e.g. perceiving geometry), cognition (e.g. planning, constraint reasoning), and motor actions (e.g. mouse movements, key presses) in a partially parallel manner. As such, adopting an HCI perspective on CAD modelling provides both a conceptual bridge and a methodological foundation for explaining CAD performance differences observed in prior studies, while enabling more granular descriptions of processes that are also well suited to relating CAD activity to CL.

2.3.1 Cognitive architectures for modelling human performance in HCI tasks

Cognitive architectures are employed to symbolically model user's embodied cognition (P, C, and M processes) by representing and explaining their information-processing system [84]. They assume modular organisation of human mind and a limited capacity of the information-processing system [84]. This way of framing cognition aligns with CAD modelling as a cognitively demanding HCI activity in which engineering designers repeatedly perceive rich visual stimuli, manipulate intermediate representations under working memory (WM) limits, and execute actions through mouse and keyboard input.

This section further focuses on Model Human Processor (MHP) cognitive architecture as one of the most used for human performance modelling in HCI [7]. Due to its broad application in HCI accompanied by its simplicity, the MHP is expected to be the most suitable for modelling engineering designers' CAD performance. In the context of this thesis, the MHP provides a suitable starting point for modelling CAD performance because it offers a clear and usable decomposition into the P, C, and M systems, together with parameters that formalise processing constraints (e.g. cycle times and memory characteristics) in a way that is interpretable and compatible with operator-based task analyses [7]. The thesis aims to connect the observable CAD actions to underlying information-processing demands without requiring a full theory of cognition. MHP is well matched to this goal; it is intentionally simplified, but sufficiently explicit to support structured modelling of HCI with CAD systems and to motivate explanatory links between CAD behaviour and information processing demands, which is central to this thesis' process-oriented view of CAD performance and its relationship to CL.

More detailed architectures such as Queueing Network MHP (QN-MHP) [83] and Adaptive Control of Thought–Rational/Perceptual-Motor (ACT-R/PM) [82] are introduced after the MHP to provide a reader a sense of additional detail but also associated complexity they bring. These two cognitive architectures thus demonstrate how modelling fidelity can be increased, particularly by representing concurrency, multi-goal processing, and richer memory or control mechanisms. However, their use typically requires substantially more assumptions, parameters and knowledge representations than are easily justified or validated in CAD studies. For this reason, MHP is suggested here a pragmatic starting point that can be extended once a stable CAD-specific task structure and operator mapping are established.

2.3.1.1 Model Human Processor (MHP)[1]

MHP, proposed by Card et al. is a cognitive architecture intended for building simplified models of human performance in HCI [7]. Due to its simplicity (compared to other cognitive architectures), it is not intended to (fully) describe the human mind and behaviour but to support understanding, calculating, and potentially predicting human performance relevant to HCI [7]. The MHP is described by 1) subsystems consisting of a set of memories and processors together with 2) a set of principles of operation [7]. Three interacting subsystems the MHP consists of are: 1) perceptual, 2) cognitive, and 3) motor system. Each of these systems has a processor responsible for approximating the information processing, while perceptual and cognitive systems also have memories as information storages (see Figure 2).

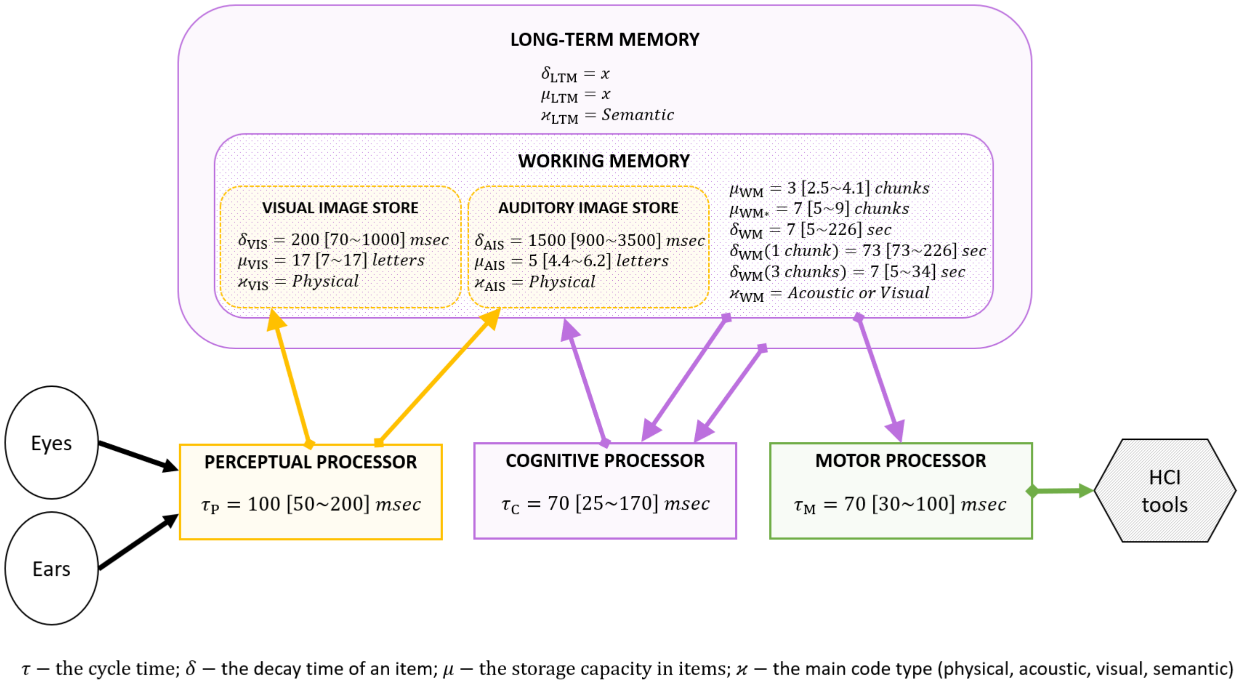

Figure 2 Memories and processors of the MHP, adopted from Card et al. [7]

The three processors MHP proposes work in series for some tasks (e.g. pressing a key in response to a light), while parallel operation of three processors is permitted and required for describing other tasks (e.g. typing or reading). In the latter case, all three processors work simultaneously. The processors are described by their cycle time $\tau_{P,C,M}$ as the most important parameter, referring to the average time required to finish an operation. In addition to the P processor, the P system consists of sensors (eyes and ears) and two associated buffer memories: Visual Image Store (VIS) and Auditory Image Store (AIS) to hold the output of the sensory system while it is being symbolically coded. The C system receives symbolically coded

[1] The description is adopted from Card et al. (1983) The psychology of human-computer interaction [7]

information from the buffer memories in its Working Memory (WM) and uses previously stored information in Long-Term Memory (LTM) to make decisions about how to respond to stimuli. The M system does not have a memory; it only carries out the response. The most important parameters of memories are: the storage capacity expressed in items ($\mu_{\mathrm{VIS,AIS,WM,LTM}}$), the decay time of an item ($\delta_{\mathrm{VIS,AIS,WM,LTM}}$), and the main code type ($\varkappa_{\mathrm{VIS,AIS,WM,LTM}}$). The parameter definitions and values are described in following sections; they are adopted from Card et al. [7] unless stated otherwise.

2.3.1.1.1 The perceptual system

The P system translates sensations of the physical world detected by the body's sensory systems (visual and/or auditory) into internal representations of the mind. Visual information enters the system through the eyes, which are in continual movement in a sequence of saccades, each approximated to take 30 msec to jump from one to the new point of focus and stay fixed there 60 - 700 msec, summing in a total duration of eye movement of 230 msec on average. This value may range from 70 msec to 700 msec, depending on condition of measurement, task or subject variables. For the auditory stimulus, there is a corresponding AIS working in a similar way. These two sensory memories (VIS and AIS) hold information (visual or auditory) coded physically (as represented in physical world), in an analogue to the external stimulus ($\varkappa_{\mathrm{VIS}} = physical$; $\varkappa_{\mathrm{AIS}} = physical$). Shortly after a physical representation of a stimulus appears in one of the perceptual memories, a recognised symbolic, visually or acoustically coded representation of at least part of the perceptual memory contents occurs in WM. The perceptual memory trace may fade before all items get transferred to symbolic representation in WM. This scenario occurs in three main cases: 1) when the contents of perceptual memory (input information) are complex or numerous; 2) the stimulus is presented only fleetingly; 3) WM may already be filled to full capacity. The MHP accounts for the potential decay of VIS or AIM through the half-life parameter (δ), defined as the time after which the probability of information retrieval is less than 50% ($\delta_{\mathrm{VIS}} = 200[70{\sim}1000]$ msec; $\delta_{\mathrm{AIS}} = 1500[900{\sim}3500]$ msec). At the physical level, when visual information enters the system, the C processor can specify which portion of the perception memory is to be encoded (this is related to attention). The capacity of the VIS and AIS is roughly approximated as $\mu_{\mathrm{VIS}} = 17[7{\sim}17]$ letters; $\mu_{\mathrm{AIS}} = 5[4.4{\sim}6.2]$ letters). The cycle time of the P processor is approximated to $\tau_{\mathrm{P}} = 100[50{\sim}200]$ msec; it varies within this range according to conditions (e.g. it is shorter for more intense stimuli).

2.3.1.1.2 The cognitive system

The C system consists of two memories: a WM that holds the information under current consideration, and an LTM that stores knowledge for future use. WM holds information that is the intermediate product of thinking (including an activated subset of the elements in the LTM) and the representations produced by the P system. Information held in WM that arrived from the P system are symbolic, non-physical visual or acoustic codes of WM ($\varkappa_{\text{WM}} = acoustic\ or\ visual$), unaffected by physical parameters of the stimulus (e.g. intensity), as opposed to the non-symbolic, physical codes of the sensory image stores, which are affected by physical parameters of the stimulus. From the functional point of view, WM is where all mental operations obtain their operands and leave their outputs after processing.

Coded information held in WM consists of symbols called chunks, which can be related to other chunks and organised into larger chunk units. What constitutes a chunk is a function of the user and the task, and it depends on the contents of the user's LTM. When a chunk in LTM is activated, this activation spreads to related chunks and to chunks related to those. As the activation spreads to new chunks, the previously activated chunks become less accessible, because there is a limited amount of activation resources. The new chunks interfere with the old ones; the effect of this interference is that the chunk appears to fade from WM with time (unless reactivated). As a working value MHP assumes the following for the decay parameter of chunks in WM: $\delta_{\text{WM}} = 7[5{\sim}226]$ sec. The decay rate is also sensitive to the number of recalled chunks; the decay gets faster with the higher number of chunks in the WM ($\delta_{\text{WM}}(1\ \text{chunk}) = 73[73{\sim}226]$ sec; $\delta_{\text{WM}}(3\ \text{chunks}) = 7[5{\sim}34]$ sec). An important characteristic of WM is its limited capacity. A pure capacity of WM is $\mu_{\text{WM}} = 3[2.5{\sim}4.1]$ novel chunks. When this pure capacity is augmented by the use of LTM, the effective capacity of WM extends to the generally familiar number of 7±2 chunks: $\mu_{\text{WM}*} = 7[5{\sim}9]$ chunks. This capacity can be extended by structuring information into larger, meaningful chunks [85]. In that way, the number of total chunks stays within the capacity limits, but they carry more content. This ability depends on information stored in the LTM and the quality of built schemas, which usually enlarges with more knowledge and expertise [86], [87]. LTM holds the user's entire available knowledge; both facts and procedures. It consists of a network of related chunks organised in schemas, which are accessed associatively from the contents of the WM. Such creation of links between chunks is referred to as semantic coding of information. Consequently, the predominant code type of the LTM is $\varkappa_{\text{LTM}} = semantic$. The MHP assumes no erasure from the LTM ($\delta_{\text{LTM}} = \infty$) as well as no defined capacity limits: $\mu_{\text{LTM}} = \infty$. However, successful retrieval of a chunk from the LTM depends on whether associations to it can be found. There are two main reasons the attempt to retrieve a chunk might fail: 1) effective retrieval associations cannot be found, or 2) similar associations to several chunks interfere with the retrieval of the target chunk.

The basic representation of cognitive processing is the recognise-act cycle; the contents of WM initiate associatively linked actions in the LTM ("recognise"), which in turn modify the contents of WM ("act"), setting the stage for the next cycle. Procedures and other forms of extended organised behaviour are built up of an organised set of recognise-act cycles. The cycle time of the C processor is set to $\tau_{\mathrm{P}} = 70[25{\sim}170]$ msec. As for the P processor, the cycle time is not constant but variable; it can be shortened by practice, task pacing, greater effort, or reduced accuracy. The C system is parallel in its recognising phase and serial in its action phase. Therefore, the C system can be aware of many things but can do only one thing at a time. This seriality occurs on top of the parallel activities of the P and M systems.

2.3.1.1.3 The motor system

Thought is finally translated into action by activating muscles. Movement is not continuous, but consists of a series of discrete micromovements, each requiring about one cycle time of the M processor, defined as $\tau_{\mathrm{M}} = 70[30{\sim}100]$ msec. For computer users, the two most important sets of effectors are the arm-hand-finger system and the head-eye system.

2.3.1.1.4 Principles of operation

Based on the description of the C, P, and M systems, the MHP proposes ten principles of operations; they are summarised in Table 5.

Table 5 The MHP's principles of operation, adopted from Card et al. [7]

Principle	Explanation
Recognise-Act cycle of the C processor	On each cycle of the C processor, the contents of WM initiate actions associatively linked to them in LTM; these actions in turn modify the contents of WM.
Variable P processor rate	The P processor cycle time τ_{P} varies inversely with stimulus intensity.
Encoding specificity	Specific encoding operations performed on what is perceived determine what is stored, and what is stored determines what retrieval cues are effective in providing access to what is stored.
Discrimination	The difficulty of memory retrieval is determined by the candidates that exist in the memory, relative to the retrieval clues.
Variable C processor rate	The C processor cycle time τ_{C} varies inversely with stimulus intensity.
Fitts' law	The time T_{pos} to move the hand to a target of size S which lies a distance D away is given by: $T_{pos} = l_M log_2(D\ /\ S + 0.5)$, where $l_M = [70 \sim 120]\ msec/bit$.
Power law of practice	The time T_n to perform a task on the n-th trial follows a power law: $T_n = T_1 n^{-\alpha}$, where $\alpha = 0.4[0.2{\sim}0.6]$, where T_1 is performance time on the first trial.
Uncertainty	Decision time T increases with uncertainty about the judgment or decision to be made: $T = l_C H$, where H is the information-theoretic entropy of the decision and $l_C = 150[0 \sim 157]\ msec/bit$. For *n* equally probable alternatives (called Hick's law); H= $\log_2(n + 1)$. For *n* alternatives with different probabilities p_i, of occurrence; H= $\sum_i p_i \log_2(1\ /\ p_i + 1)$.
Rationality	A person acts so as to attain her goals through rational action, given the structure of the task and her inputs of information and bounded by limitation on her knowledge and processing ability:

	Goals + Task + Operators + Inputs + Knowledge + Process-limits → Behaviour.
Problem space	The rational activity in which people engage to solve a problem can be described in terms of 1) a set of states of knowledge, 2) operators for changing one state into another, 3) constraints on applying operators, and 4) control knowledge for deciding which operator to apply next.

2.3.1.2 Other cognitive architectures

Two extensions of the MHP illustrate how modelling fidelity can be increased beyond the baseline decomposition into P, C, and M systems, but also highlight an important trade-off; richer psychological detail typically comes with additional assumptions, parameters, and implementation effort, which can complicate model specification and validation.

Queueing Network-MHP (QN-MHP) integrates queueing networks with the MHP by representing the P, C, and M processors as interconnected subnetworks of servers [83]. Processing cycle times from MHP are used to define server times, enabling both mathematical analysis and real-time simulation of performance. Its key advantage over MHP is the ability to represent concurrency and multitask behaviour, where multiple goals and information streams can be processed in parallel and WL emerges from network dynamics (multiple streams of information flowing through a QN) [83]. QN-MHP typically distinguishes visual and auditory perception in the P subnetwork, separates WM functions (e.g. visuospatial sketchpad, phonological loop, and central executor, following the suggestions from Baddeley [88]), and goal execution in the C subnetwork, and represent motor processing through multiple servers (e.g. for sensorimotor integration, movement tuning, motor element recall, and motor sequencing or programming) and actuators (the right hand, the left hand, eyes, feet, and mouth) in the M subnetwork [89]. However, this added representational power increases complexity as the modeller must specify network structure, routing, and timing assumptions in more detail, and the resulting models can become difficult to parameterise and validate in applied contexts where fine-graine empirical evidence about concurrent goal streams is limited.

Adaptive Control of Thought-Rational/Perceptual-Motor (ACT-R/PM) augments the ACT-R cognitive architecture [80] with a perceptual-motor system to support modelling of HCI performance by coordinating P, C, and M activity within a modular framework [82]. ACT-R is commonly positioned as a comprehensive architecture for describing and predicting human performance in primarily cognitive domains [80]. ACT-R/PM is built from the modules (serial processors specialised for P, C, or M operations) and buffers (interfaces that hold one currently available chunk and mediate communication between modules). Central cognition is commonly represented through declarative memory (knowledge retrieval), procedural memory (production rules controlling behaviour), and goal management, optionally supported by imaginal or spatial modules. Perceptual-motor modules provide interfaces for visual and auditory processing and for motor control (e.g. mouse and keyboard actions), enabling interaction with the task and computer system [82], [80]. These modules communicate with central cognition modules in two forms: 1) production actions can request that a particular module perform a certain command, and 2) modules can modify declarative memory, primarily by creating declarative chunks [82]. The main advantage of ACT-R/PM is its capacity to capture knowledge-driven behaviour and control in a coherent theoretical framework (e.g. how memory retrieval and production rules shape action selection). At the same time, this advantage is also its main limitation in applied CAD contexts; meaningful ACT-R/PM models often require detailed specification of domain knowledge, production rules, and retrieval structures, which substantially increases modelling effort and makes the model sensitive to assumptions that are difficult to justify or verify without extensive additional data.

Together, QN-MHP and ACT-R/PM demonstrate how performance modelling can capture richer mechanisms such as parallel processing, multi-goal control, and more detailed memory structures, but they also increase modelling complexity relative to MHP, which is an important consideration for applied CAD performance modelling. This trade-off is particularly salient in CAD, where tasks are already complex and highly variable across CAD modelling goals and strategies. Adding a highly detailed cognitive architecture on top of this task complexity can make model development and validation disproportionately demanding; consequently, such architectures are still more often practical for domains with more constrained and well-defined tasks. For these reasons, this thesis adopts MHP as a pragmatic starting point and relies on the task analysis framework to capture the structure of CAD tasks at an appropriate level of granularity. Accordingly, the next section introduces task analysis approaches used to specify users' goals, methods, and operator sequences in HCI.

2.3.2 Task analysis frameworks

Cognitive architectures provide explanations and assumptions on how human embodied cognition operates (e.g. P, C, and M processing and their limitations). However, they do not describe what the user is trying to achieve, how the task is structured, or how interaction with a specific HCI tool unfolds. Task analysis frameworks serve this purpose in human performance modelling by describing task structure through description of goals, actions, and decision logic while explicitly considering a HCI system a user interacts with (e.g. application interface and system responses to user inputs). In that sense, these frameworks operationalise the fundamental

Rationality principle from MHP, which links behaviour to goals, task structure, available operators, and limited processing capacity [7]:

A person acts so as to attain his goals through rational action, given the structure of the task and his inputs of information, and bounded by limitations on his knowledge and processing ability: Goals + Tasks + Operators + Inputs + Knowledge + Process-limits → Behaviour.

Additionally, frameworks either implicitly or explicitly relate to assumptions on human information processing described within cognitive architectures [84], [90].

Commonly used task analysis frameworks for the HCI context belong to the GOMS family (Goals, Operators, Methods, and Selection rules), originally introduced by Card et al. [7] and elaborated upon by John and Kieras [84]. GOMS frameworks have remained influential because they are 1) usefully approximate for many HCI tasks, 2) learnable and usable by both practitioners and researchers, and 3) flexible enough to cover a broad range of behaviour involved in routine HCI tasks [84].

GOMS proposes a structure of the user's task performance consisting of 1) the user's goals, 2) basic operators that the user could employ to achieve those goals, 3) methods as combinations of the basic operations, and 4) selection rules for choosing between alternative methods [91]. When GOMS framework is used to analyse a task, these four components give a complete dynamic description of a task performance [84]. In standard GOMS terminology, goals specify the symbolic structures that define what is to be achieved and constraints the set of possible methods that could achieve the state. Operators represent the elementary operations of the P, M, and C system available to the user; the execution of an operation represented by the operator is the only way to change any aspect of the user's knowledge state or act on the task environment. Methods are already learned procedures (sequences of operators and subgoals) for accomplishing goals; they are available to the user at performance time and as such they do not account for plans that are created while performing the task. When more than one method for accomplishing a goal is available, selection rules must be used to choose between them. Sometimes the task environment dictates which method is appropriate; other times the user makes the rule. In skilled behaviour these selections are made smoothly and quickly. The difference between a goal and an operator in GOMS analysis is a matter of the level of detail chosen by the analyst; for a goal, the analysis provides a method that uses lower-level operators to specify the details of how it is to be accomplished, while operators are treated as the lowest-level actions and are not further broken down.

Importantly for this thesis, the GOMS family is well suited for the CAD context because CAD modelling, especially for trained practitioners, often involves goal-directed, repeatable HCI sequences (e.g. creating a sketch, constraining geometry, applying CAD features), where performance depends on both the task structure and the sequence and timing of CAD actions and associated operators. This enables consistent modelling of both what engineering designers do (goal/method structure) and how they do it (operator sequences), which supports comparing CAD modelling strategies, relating processes to outcome quality, and deriving interpretable measures linked to information-processing demands. At the same time, GOMS comes with limitations that are particularly relevant in the CAD context: it is best suited to procedural, routine behaviour (rather than open-ended, exploratory reasoning), and it requires an externally defined set of tasks/goals derived from observation, interviews, or analyst judgement. These constraints motivate careful scoping toward well-defined CAD modelling tasks and repeatable HCI sequences, while using complementary approaches to capture less routine and more exploratory aspect of CAD activity.

Different instances of GOMS-based task analysis frameworks have been introduced in HCI research to account for different levels of granularity and analytical goals of the modelled interactive behaviour. In this thesis, the focus is placed on three GOMS variants that appear most appropriate for capturing HCI in CAD at the required level of detail (Table 6): KLM for comparing efficiency via execution time estimates for low-level CAD actions (e.g. mouse movements, clicks, menu selections); CMN-GOMS for representing hierarchical goal decomposition and alternative methods in CAD workflows where the same CAD modelling goal can be achieved through different sequences of operations; and CPM-GOMS for cases where modelling need to account for overlaps between P, C, and M activity (e.g. visually inspecting geometry while preparing the next command), noting that the value of this added detail depends on data availability and modelling scope.

Table 6 Overview of task analysis frameworks

Framework	Parallel or Serial	Task structure	Task decomposition	Advantage
KLM	Serial	Flat operator sequence	Only operators	Rapid estimate of task execution time
CMN-GOMS	Serial	Hierarchical (Goals → Methods → Operators)	Operators organised into goals and methods, hierarchy included (e.g. unit tasks)	Simplicity of KLM coupled with structured task segmentation
CPM-GOMS	Parallel	Hierarchical (Goals → Methods → Operators) Network of overlapping operators	Operators mapped in time, showing concurrency	Based directly on MHP

The remainder of this section therefore introduces these three task analysis frameworks from the GOMS family in more detail.

2.3.2.1 Keystroke-level model (KLM)

Within the GOMS family, the Keystroke-level model (KLM) is the simplest (in terms of number of parameters) and one of the most widely used frameworks for building predictive models for estimating the execution time of error-free performance by skilled users [92]. In practice, KLM analysis yields a sequence of low-level operators (e.g. keystrokes, pointing actions, homing movements) that represent how a user executes a specific task instance. The analysis constructs this operator sequence based on an assumed or observed method of interaction (i.e. the steps the user takes), and the model then predicts task execution time by summing the durations of the operators.

The original KLM includes six primitive operators: K to press a key or button, P to point with a mouse to a target on a display, H to home hands on the keyboard or other device, D to draw a line segment, M to mentally prepare to do an action or a closely-related series of primitive actions, and R to symbolize the system response time during which the user has to wait for the system [84]. Each operator is assigned as execution-time estimate (a constant, parameterised estimate, or a function such as Fitts' law for pointing). The operator definitions and typical time estimates used in KLM are summarised in Table 7 (adopted from Card et al. [92]).

Table 7 List of KLM's operators, adopted from Card et al. [92]

Operator	Operation	Duration
K	A keystroke or a button push.	Best typist (135 words/min): $t_K = 0.08$ Good typist (90 words/min): $t_K = 0.12$ Average skilled typist (55 words/min): $t_K = 0.20$ Average non-secretary typist (40 words/min): $t_K = 0.28$ Typing random letters: $t_K = 0.50$ Typing complex codes: $t_K = 0.75$ Worst typist (unfamiliar with keyboard): $t_K = 1.20$
P	Pointing to a target on a display with a mouse. The key press is not part of P; it is represented by a K following the P.	Varies as a function of the distance d to the target, and the size s of the target, according to Fitts' law: $t_P = 0.8 + 0.1 \cdot \log_2(d/s + 0.5)\ sec$ The fastest time according to this equation is $t_P = 0.8$ sec and the longest likely time (d/s = 128) is $t_P = 1.5$ sec, but Card et al. kept a constant time of $t_P = 1.1$ sec for a simplicity [92].
H	Homing the hand(s) on the keyboard or other device, including the fine positioning adjustment of the hand on the device.	Assume a constant of $t_H = 0.4$ sec for movement between any two devices.
D	Manually drawing a set of straight-line segments using the mouse. It is restricted to the computer mouse and assumes that the drawing system constraints the cursor to lie on a 0.56 cm grid.	t_D is a linear function of two parameters; the number of segments n_D and the total length of all segments l_D. The coefficients of this function can be different for different users. $t_D = 0.9n_D + 0.16l_D$
M	Mental preparation to execute physical operators described above.	Estimated to take $t_M = 1.35$ sec on average. The use of a single mental operator is a simplification.

R	System response time. The R times are counted only when they require the user to wait for the system. When an M operation follows an R, the response time is not counted unless it is over 1.35 seconds, since the expert user can completely overlap the M operation with the response time.	Response time in seconds. Response times are different for different systems, for different commands within a system, and for different contexts of a given command. The KLM does not embody a theory of system response time. The response times must be input to the model.

Since the KLM operator sequence is built explicitly by the analyst, KLM does not require any particular computational implementation; it can be applied "by hand" (or in a spreadsheet) as long as the operator sequence and timing assumptions are stated. Conceptually, KLM assumes serial execution of operators (a single operator stream), and it treats cognitive processing in a simplified way through the placement of M operators. To support consistent placement of cognitive processing time related to mental preparation, KLM provides a set of heuristic rules for inserting M operators [84]. These heuristics reflect the assumption that skilled users organise actions into cognitive chunks, preparing for the next chunk rather than each individual physical operation [92]. From that it follows that in executing methods (represented as sequences of the operators) the user is more likely to pause between chunks than within chunks. The rules are summarised in Table 8, summarised from Card et al. [92].

Table 8 KLM's heuristic rules for placing the M operator, adopted from Card et al. [92]

Rule	Rule explanation
Rule #0	Insert M operators in front of all K operators that are not part of argument strings proper (e.g. text strings or numbers). Place M operators in from of all P operators that select commands (not arguments).
Rule #1	If an operator following an M is fully anticipated in the operator just previous to M, then delete the M. It is assumed they belong in one chunk.
Rule #2	If a string of MK operators belongs to a cognitive unit (e.g. the name of a command), then delete all M operators but the first.
Rule #3	If a K is a redundant terminator (e.g. the terminator of a command immediately following the terminator of its argument), then delete the M in front of the K.
Rule #4	If a K terminates a constant string (e.g. a command name), then delete the M in front of the K; but if the k terminates a variable string (e.g. an argument string), then keep the M.

2.3.2.2 CMN-GOMS

CMN-GOMS builds on the operator-level foundations as KLM, but extends analysis by representing behaviour in a hierarchical task structure; goals are decomposed into subgoals and methods, and selection rules specify how alternative methos are chosen [84], [7]. In other words, CMN-GOMS supports modelling both low-level operator sequences and higher-level procedural organisation (Goals → Methods → Operators), which is useful when the same CAD outcome can be achieved through different workflows and decision points.

In CMN-GOMS, methods can be expressed in a structured notation resembling pseudocode, meaning they can include sub-methods (reusable method fragments) and conditional statements that capture method choice under different task of interface conditions [84]. This "program-like" representation is descriptive rather than implying that users literally execute code; it is simply a convenient way to specify hierarchical procedures and branching logic precisely. Compared with KLM, CMN-GOMS therefore supports richer explanations of why a particular operator sequence occurs (via goal/method structure), while retaining a sequential execution assumption. However, CMN sources do not provide an explicit step-by-step procedure for constructing CMN-GOMS models; instead, the framework is mainly communicated through worked examples that illustrate progressively expanding the goal hierarchy to the desired level of detail [84], [7].

2.3.2.3 CPM-GOMS

Cognitive-Perceptual-Motor GOMS (CPM-GOMS) framework is a GOMS variant that explicitly represents overlap between P, C, and M activity by mapping operators onto the corresponding processors and scheduling them over time [84]. Unlike KLM and CMN-GOMS, CPM-GOMS does not assume a single serial stream of operators; instead, operators assigned to different processors may run in parallel (e.g. visual encoding while the hand is moving), whereas operators on the same processor remain sequential. CPM-GOMS also respects logical dependencies; two operators cannot overlap if the output of one is required as the input to the other.

The "CPM" also reflects its use of the Critical Path Method; task execution time is determined by the longest dependency-constrained path through the scheduled operator network [84], [93]. In practice, CPM-GOMS is often built on a prior CMN-GOMS decomposition (to define goals and methods), after which operators are assigned to processors and scheduled to reflect concurrency.

To reduce the modelling effort and improve consistency, applying CPM-GOMS commonly leverages templates proposed by John et al. as predefined operator patterns for frequently occurring HCI units (e.g. point→click→verify sequences, visual search patterns, etc.) [94]. Templates function as reusable building blocks; instead of re-deriving a detailed P-C-M operator schedule from scratch for every task instance, the modeller selects an appropriate template and instantiates it with the task-specific details. This is particularly helpful when tasks contain repeated interaction segments, but it also introduces an important practical limitation; CPM-GOMS requires substantially more detailed assumptions (and often richer empirical grounding) to justify operator assignments, overlaps, and template selection that KLM or CMN-GOMS. For complex and variable domains such as CAD, this additional modelling burden can be significant, and the value of CPM-GOMS therefore depends strongly on data availability and modelling scope.

A central limitation shared by KLM, CMN-GOMS, and CPM-GOMS is that they are designed primarily to model error-free performance by skilled users, typically assuming that users already know the appropriate methods and execute them without mistakes, breakdowns, or exploratory detours [84], [92], [95]. In this context, error-free does not imply that users in practice never make errors; rather, it means that the model predicts performance for an idealised, routine execution of a task once it has been learned. As a consequence, models built with these task analysis frameworks do not natively capture behaviours that are common in complex, real-world CAD modelling, such as trial-and-error, backtracking, uncertainty-driven pauses, or strategy shifts unless the analyst explicitly encodes them as additional steps (e.g. including rework operators or alternative methods). This is particularly important in CAD because performance differences often emerge precisely through such non-routine episodes (e.g. iterations, revisions, and problem-solving), which can dominate both execution time and CL.
The implications of this assumption differ slightly across the three variants. KLM is most restrictive because it represents behaviour as a flat operator sequence and treats cognition through simplified placement of mental operators; it therefore best supports time estimates for well-practised, HCI-level routines. CMN-GOMS remains grounded in skilled, procedural execution, but it can better represent alternative learned methods and their choice via selection rules, which makes it useful for comparing established CAD workflows. CPM-GOMS further assumes skilled performance, but is particularly applicable when the task is practised enough that P, C, and M activities become reliably coordinated and partially overlapping; hence prior work noting that CPM-GOMS can be especially suitable for modelling error-free performance after extensive practice on a specific task [86]. With these assumptions and capabilities in mind, the next section reviews how introduced human performance modelling approaches have been applied to build cognitive models of CAD performance.

2.3.3 Existing (studies on) cognitive models of CAD performance

Human performance modelling has been applied in CAD research primarily as a method for evaluating interfaces and HCI procedures, especially in early HCI system design stages and in redesign context (e.g. [92], [96]). Rather than treating CAD performance only through output quality or cumulative summaries, these studies use GOMS-based modelling to make HCI procedures explicit and to reason about performance consequences (most often execution time) under particular task assumptions. The literature is still limited in volume, but it provides useful examples of 1) time prediction for well-defined HCI routines in CAD, 2) interface evaluation and redesign using procedural analyses, and 3) modelling of higher-level strategies and goal structures in CAD tasks.

2.3.3.1 KLM for execution time prediction in HCI tasks related to CAD

Card et al. introduced the KLM to predict one specific aspect of HCI performance – the execution time required for an expert user (given assumed motor parameters) to complete a specified graphic task using a specified method on a particular interactive computer system (graphic editor) [92]. To validate the resulting KLM models, users performed short, well-defined graphics-editing tasks (e.g. adding labels, reconnecting lines, deleting and copying objects) in three editors that differed in interaction style (e.g. how lines are defined and how commands are selected) [92]. Task methods were represented as serial sequences of primitive KLM operators (K, P, H, D, M, R), each associated with an estimated duration (see Table 7), with mental preparation time modelled through heuristic placement of M operators (according to Table 8) [84], [92]. Execution time was then predicted by summing the operator times (Equation **Error! Reference source not found.**), and total task time additionally considered acquisition time estimated from empirical data.

$$\begin{aligned} T_{execute} = {} & T_K + T_P + T_H + T_D + T_M + T_R \\ = {} & n_K t_k + n_P t_P + n_H t_H + n_D t_D + n_M t_M + n_R t_R \end{aligned} \tag{1}$$

Model predictions were compared with empirically observed times, yielding errors that varied across tasks (reported in the range of roughly 1% to 40%) [92]. Card et al. also examined several simplified variants (e.g. keystroke counting, using only physical operators with a prorated time for cognitive activity, or treating all operators as a single constant), showing that these simplifications generally reduced predictive accuracy but could still be useful when the goal is a rapid, coarse comparison of alternatives rather than precise prediction [92].

This work provides a foundational idea for CAD performance modelling; when CAD tasks can be defined as repeatable HCI routines, performance differences between interfaces designs or methods can be expressed and compared through differences in operator sequences and timings. It also highlights two constraints that are crucial in CAD contexts. First, KLM is best suited to short, well-specified, error-free routines; it does not naturally represent exploration, revisions, or strategy changes that are common in realistic CAD modelling unless these are explicitly encoded. Second, the comparison of simplified variants makes a practical point for CAD modelling; there is a trade-off between fidelity and effort; simplified models may be adequate for early-stage screening of interface or workflow alternatives, while more detailed operator representations are needed when the objective is to explain where time is spent and which interaction components drive performance differences.

2.3.3.2 CMN-GOMS for interface evaluation and redesign in CAD

Gong and Kieras applied CMN-GOMS task analysis framework to evaluate and redesign a Macintosh CAD interface, using the resulting GOMS model to identify usability problems embedded in the procedures by which users interact with the system and to estimate performance changes under redesign [96]. After eliciting high-level user goals through sessions with users, they constructed a model (not reported in full detail) and used it to diagnose six recurring procedural issues, including frequent goals supported by inefficient methods, frequent goals lacking a supported method, inconsistent procedures across similar goals, methods placing heavy demands on LTM or WM, and method requiring excessive mouse/hand movements [96]. These findings informed a revised interface intended to streamline interaction and reduce memory burdens, reported in the modelling results as eliminating over half of the mouse movements and substantially reducing the amount of information that needed to be maintained in WM during task execution [96].

A key contribution of this study is that it demonstrates how task analysis can guide CAD interface redesign and yield quantitative, testable predictions. Gong and Kieras validated their redesign by comparing predicted improvements with experimental data, reporting close agreement between model-based and observed gains (≈39–40% improvement in execution time and 46% reduction in learning time) [96]. Beyond the redesign outcome, the study also provides a practical modelling insight highly relevant to CAD; default KLM parameter values proposed by Card et al. [7] may not transfer well across interfaces. Their data suggested that standard estimates were reasonable for keystrokes, hand homing, and mouse-button clicks, but that the commonly used constant mouse movement time was consistently inaccurate in their CAD interface; they therefore refined pointing-time estimates using Fitts' law applied to the actual distances and target sizes in the interface, yielding smaller and more accurate values [96]. They further discussed limited cases where preparatory or verification elements could be effectively absorbed within well-practised sequences, and proposed rules for representing such overlap in modelling [96].

2.3.3.3 CMN-GOMS for modelling procedural knowledge and goal structure in CAD

Lang et al. focused less on time prediction and more on what CAD users need to know and what they do while performing CAD tasks, distinguishing declarative knowledge (e.g. knowledge of the object being designed and of CAD commands) from procedural knowledge (i.e. strategies and goal structures used to complete the CAD task) [97]. The purpose of distinguishing and observing these two types of knowledge in CAD tasks was to enhance and properly target CAD trainings towards one or the another. Their first experiment used behavioural measures such as command variety, errors, accuracy, and time to compare novices and experts; results suggested broadly similar declarative knowledge indicators (e.g. variety of different commands used, and the number of errors), while measures tied more closely to procedure (e.g. number of used commands) pointed to differences in how users executed the task [97]. Their second experiment then examined procedural knowledge more directly by reconstructing users' goal structures from HCI data (e.g. keystrokes or command sequences and pauses) available from the captured CAD modelling sessions of four participants, explicitly excluding errors and not decomposing behaviour into low-level actions because the aim was task-structure extraction rather than time prediction [97]. This study is therefore an example of using GOMS task analysis framework to represent what the user is trying to achieve and how they organise the work, rather than to estimate execution time.

Two findings from Lang et al. have important implications for cognitive modelling of CAD performance. First, they observed qualitatively different CAD modelling strategies; in a *function-related* strategy, high-level goals are organised around CAD command functions (e.g. drawing, cleaning, trimming, projecting) whereas lower-level subgoals are organised around the different parts of the object (e.g. the two circles, the keyway, the lines joining the circles). In contrast, an *object-related* strategy reflects a different structuring of goals and subgoals around the evolving object representation, where the high-level goals are organised around the command functions [97]. This implies that CAD performance models may need to accommodate multiple legitimate procedural organisations even when users produce the same final geometry. Second, Lang et al. questioned the common assumption that one generic GOMS model can explain performance of all users, with variation captured mainly through selection rules. They argued that when participants differ substantially in procedural knowledge and goal structure, a singe "baseline" model becomes inadequate, and modelling may require distinct method/goal representations for different user groups or strategy profiles, consistent with the large variability in observed task execution times associated with these fundamentally different procedural organisations [97].

2.3.3.4 GOMS for identifying efficient CAD strategies

Bhavnani and John focused on identifying efficient CAD strategies (non-obligatory goal-directed methods) since CAD systems typically allow for more than one way of achieving the same goal [98]. They argued that inefficient behaviour often persist for two practical: efficient strategies are not made explicit (and therefore remain unknown to many users), and efficiency is not prioritised because its performance consequences are not clearly demonstrated [98]. In an experiment on a 2D CAD task, they identified two contrasting strategies (sequence-by-operation and detail-aggregate-manipulate) and proposed a four-layer view of CAD knowledge (presented in Figure 3) to indicate where such strategies reside. The four layers defined through the goal decomposition are: 1) the *task layer* that describes the task performed by the user; 2) the *intermediate layer* that decomposes the task with knowledge of how to organize the different kinds of commands provided by the system; 3) the *command layer* that decomposes each of the stages in the intermediate layer into specific commands; and 4) the *keystroke layer* that specifies the operations needed to execute the commands [98]. A key implication is that strategically important differences may be encoded at the intermediate layer; in how users organise and combine command types, meaning they are not reliably detectable from command logs alone without a higher-level task decomposition and interpretation.

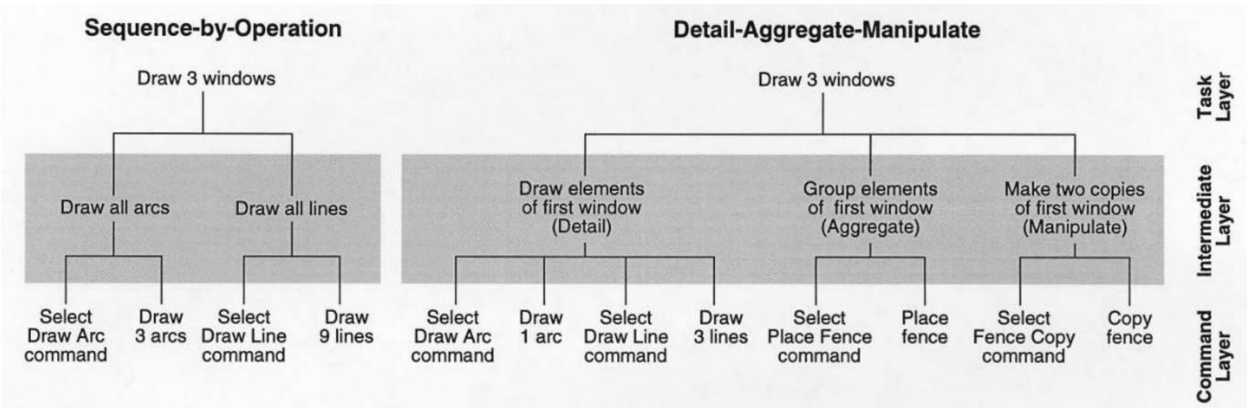

Figure 3 Knowledge levels and modelling strategies, adopted from Bhavnani and John [98]

They then used separate GOMS models (one per strategy) to compare predicted execution times for a panel clean-up task, showing that strategy choice can dominate performance outcomes; the detail–aggregate–manipulate strategy yielded a substantial reduction in predicted time (reported as 71%) [99], [100]. These results reinforce two points relevant for this thesis. First, modelling CAD performance benefits from representing alternative methods explicitly, because different strategies can produce the same output with markedly different efficiency. Second, the findings support the thesis' emphasis on the CAD modelling process and its structure; to understand and improve CAD performance (and later relate it to CL), analysis must go beyond quality of the resulting output and cumulative metrics to capture how CAD modelling process is organised across meaningful CAD task/goal levels. At the same time, the study also illustrates a recurring limitation in prior work: key modelling details (e.g. operator mappings and parameter choices) are often underreported, which constrains reproducibility and motivates clearer, more transparent modelling specifications in CAD performance modelling.

Across these studies, three findings are especially relevant for positioning the present thesis. First, error-free performance of skilled users is a pervasive assumption when the goal is time prediction; models typically represent routine execution of a specified method and do not natively capture exploration, breakdowns, or trial-and-error unless these are explicitly encoded as additional steps. Second, prior work often uses multiple models, not "the model"; different interface versions, alternative strategies, or distinct methods commonly require separate model encodings to make comparisons meaningful (e.g. baseline vs. redesigned interface; strategy A vs. strategy B). Third, much of the modelling emphasis is on predicting or comparing execution time because time estimates provide a practical, interpretable basis for evaluating alternative procedures and interfaces. However, CAD tasks are generally more complex and variable than the short, well-defined interaction tasks for which these frameworks were originally demonstrated; CAD modelling often involves iterative refinement, opportunistic planning, and revisions that challenge the routine, error-free assumption and complicate model specification and validation. This gap suggests that direct adoption of existing cognitive modelling practices in CAD requires careful scoping (e.g. focusing on well-defined procedures) and explicit statement of assumptions, while also motivating CAD-specific extensions that better accommodate iterations, goal interleaving, and variability of the task structure.

2.3.4 Conclusions of the literature review on human performance modelling

This review positions human performance modelling as a structured approach for studying CAD as an HCI process, because it offers explicit constructs for representing how observable interaction unfolds under limited information-processing capacity. In terms of cognitive architectures, the Model Human Processor (MHP) emerges as the most suitable starting point for this thesis. MHP is intentionally simplified; it is designed to support performance estimation and explanation in interactive tasks rather than to serve as a comprehensive theory of cognition [7]. This is an advantage in the CAD context, where modelling already faces substantial task complexity and limited access to the detailed data related to cognitive processes that richer architectures would require. In practice, MHP provides a clear and usable cognitive structure for linking CAD actions to resource-limited processing, without committing the thesis to assumptions that cannot be empirically justified.

More detailed architectures, such as QN-MHP and ACT-R/PM, illustrate what higher modelling fidelity can add; particularly the ability to represent concurrency, multi-goal control, and richer memory mechanisms [79], [83]. However, they also require substantially more parameters, stronger assumptions, and more detailed knowledge representations (e.g. production rules), shifting the modelling burden from interaction-level explanation toward cognitive specification that is difficult to validate for complex and variable CAD tasks. For this reason, the thesis adopts MHP as a pragmatic baseline and prioritises incremental extensions only where they are clearly warranted (e.g. limited parallelism or goal interleaving).

With respect to task analysis frameworks, the GOMS family provides the most practical modelling foundation for CAD because they support explicit representation of goals, methods, and operator sequences at a controllable level of granularity. The review of existing cognitive models of CAD performance highlights several implications that directly inform this thesis. First, time-predictive applications rely on the error-free skilled-user assumption, meaning that models typically represent routine execution of a specified method and do not capture exploratory behaviour or rework unless explicitly encoded. Second, meaningful comparisons commonly require multiple model instantiations (e.g. baseline vs. redesign, strategy A vs. strategy B), rather than a single universal model. Third, prior work is strongly time-oriented, using predicted or compared execution time as the primary performance outcome. Finally, CAD tasks are substantially more complex than the short, well-defined interaction tasks on which classic GOMS demonstrations were built, which makes careful scoping, transparent assumptions, and CAD-specific task decompositions essential in further work.

Taken together, these conclusions motivate the thesis' methodological direction; MHP is adopted as the foundation for describing engineering designers' embodied cognition and GOMS-derived models as the basis for defining the hierarchical structure of CAD tasks, supported by reproducible operator-mapping rules (heuristics) and segmentation coding schemes that are appropriate for the CAD context and its associated complexity. This approach also provides a direct bridge to CL as the thesis' second core construct, because the same models that describe what engineering designers do and when they do it can be used to reason about how information-processing demands vary across the CAD modelling process. Accordingly, the following section reviews the CL literature and the methods used to measure and analyse CL in CAD-related activities.

2.4 Cognitive load (CL)

CL is a multidimensional construct representing the load that performing a task imposes on the cognitive system of one who performs it [101]. It emerges from the observation that the human cognitive system has a limited capacity for information processing when performing cognitive tasks [14]. This perspective directly complements the preceding discussion of human performance modelling (Section 2.3.); while cognitive architectures and models based on task analyses describe what designers do and how HCI unfolds, CL construct help explain how demanding different segments of the CAD modelling process are under limited information-processing capacity. This is especially relevant in CAD, where engineering designers must continuously perceive rich visuospatial information, maintain and manipulate intermediate representations under WM constraints, and execute motor actions through complex CAD system interfaces, often while iterating and revising.

Still, it is important to highlight that definitions as well as measures of CL are not uniform across the literature, leading to the ongoing ambiguities in both [102], [103]. In fact, the main point of the agreement in researching CL across the decades revolves around the notion about the lack of its uniform definition. For example, Johannsen already in 1979 expressed the concern that *there exist too many conflicting ideas about the definition and measurement of workload. The term workload frequently is used even without any definition of what it is* [104]. One decade later, Linton et al. stated that *the simple fact of the matter is that nobody seems to know what workload is. Numerous definitions have been proposed, and many of them seem complete and intuitively "right". Nevertheless, current definitions of workload all fail to stand the test of widespread acceptance or quantitative validation* [105]. Moving fast forward, Salvendy and Karwowski more than 30 years later suggest no significant change in the status of the uniform CL definition: *Given this complex and multidisciplinary history, it is unsurprising that a single, universally accepted operational definition of mental workload remains elusive* [83].

Given the aims of this thesis, CL is observed from the viewpoint of resources as *the operator's allocation of limited processing capacity or resources to meet task demands; that is, the balance of internal resources and external demands* [83]. Additionally, the thesis uses the term *cognitive load*, while various scholars have referred to the same concept as *mental workload* or *cognitive workload*. For the literature review purposes, the thesis adopts original terms the scholars used in their papers, but the underlying meaning of these concepts stays the same as of *cognitive load* as they may be used interchangeably [83], [106].

The following sections first introduce theoretical models of CL, then discuss how CL can be measured, and finally summarise factors thus far investigated and proven to influence CL in the CAD context.

2.4.1 Theoretical models of CL

The overview groups theoretical models of CL into 1) generic conceptual models stemming from cognitive psychology, 2) conceptual models tailored for the engineering design context (beyond CAD, but informative for CAD as a design activity conducted through HCI), and 3) models that theorise about CL using cognitive architectures and associated task analysis frameworks (i.e. cognitive models).

2.4.1.1 Generic conceptual models of CL

Several foundational conceptual models were originally developed in contexts focused on learning. Their value in this thesis lies primarily in their structure; that is, in the way they decompose CL into interpretable components and relate CL to performance, rather than in the original learning setting in which they were introduced. To make them informative for the CAD context, the structure they propose must be interpreted and adapted to reflect how CAD activities are actually carried out; applying these conceptual models in CAD requires mapping their constructs and linking them to CAD process descriptions at a suitable level of granularity. Paas and Van Merriënboer observed and described CL of a learner through the causal and assessment factors, presented in Figure 4 [107].

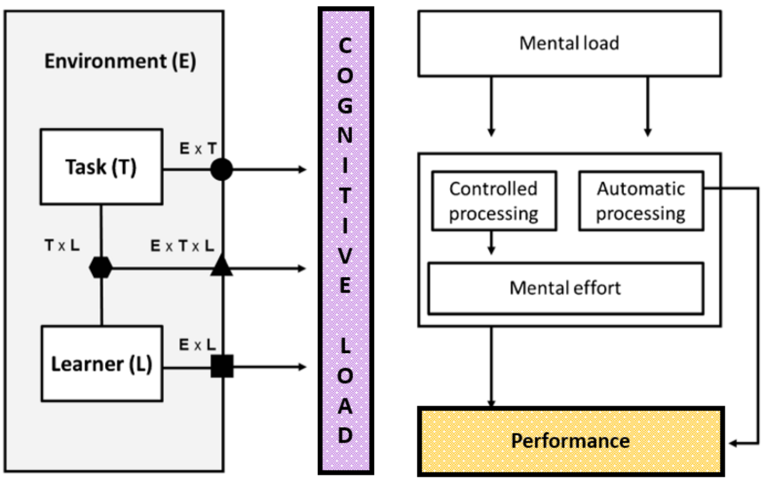

Figure 4 Model of CL, suggested by Paas and Van Merriënboer [107]

The left part of Figure 4 (causal factors) summarises sources of CL; they refer to the characteristics of a task (e.g. novelty or time pressure), learner (e.g. cognitive capabilities and prior knowledge), and environment as well as their interactions (e.g. motivation or state of arousal). The right part (assessment factors) summarises how CL is inferred, by observing three dimensions closely related to CL measurements: mental load, mental effort, and performance [107]. In this model, mental load is a task-centred dimension that represents demands imposed by the task and environment; it is independent of subject characteristics and constant for a given task in a given environment. Mental effort is a human-centred dimension that refers to the amount of resources actually allocated by the individual to accommodate the task demands; it reflects the amount of controlled processing the individual is engaged in. The final assessment factor is task performance, which reflects all three causal factors as the outcome of working under demands imposed by them.

Moreover, Sweller et al. in their model distinguish intrinsic and extraneous CL, depending on its function [13]. They explain intrinsic CL as a one imposed by the intrinsic nature of the information; in other words, by the basic structure of the information the learner needs to acquire for achieving learning goals irrespective of other influencing factors. Extraneous CL is imposed by the manner in which the information is presented or by the activities in which learners must engage. In that sense, this category of CL is extraneous to the learning goals. Consequently, extraneous CL should be reduced to minimum so that cognitive resources can be devoted to intrinsic CL and thus maximise learning performance.

In a closely related framing, Xie and Salvendy distinguish effective from ineffective CL, with the aim of highlighting the possibility to reduce ineffective CL by controlling factors that generate it [102]. Effective CL is the one that people must bear while working, even if they act efficiently and correctly. It is the minimum level of CL generated externally by the task requirements regardless of who performs the task (i.e. task load). In that sense, it is a constant for a given task and will not be affected by learning and training. In contrast, ineffective CL is the load that workers could avoid; it could be reduced by learning and training, and affected by the direction of effort. Ineffective CL does not directly contribute to completing the task, and it is generated internally by individuals; for example, from using a bad strategy, processing information irrelevant for the task, or searching some help from help menu. Different people experience different amounts of ineffective CL for the same task. Ineffective CL could be quantified as extra usage of resources for processing extra amount of information, higher effort than necessary, or extra time than needed. Xie and Salvendy illustrate these concepts across three information processing stages: information perception, decision/response selection, and response execution (Figure 5). These stages align well with CAD as an HCI activity that repeatedly alternated between interpreting visual information, selecting actions, and executing selected actions. The first two stages also overlap in their definition with the acquisition, while the third with the execution as information-processing stages in engineering design [69], [108].

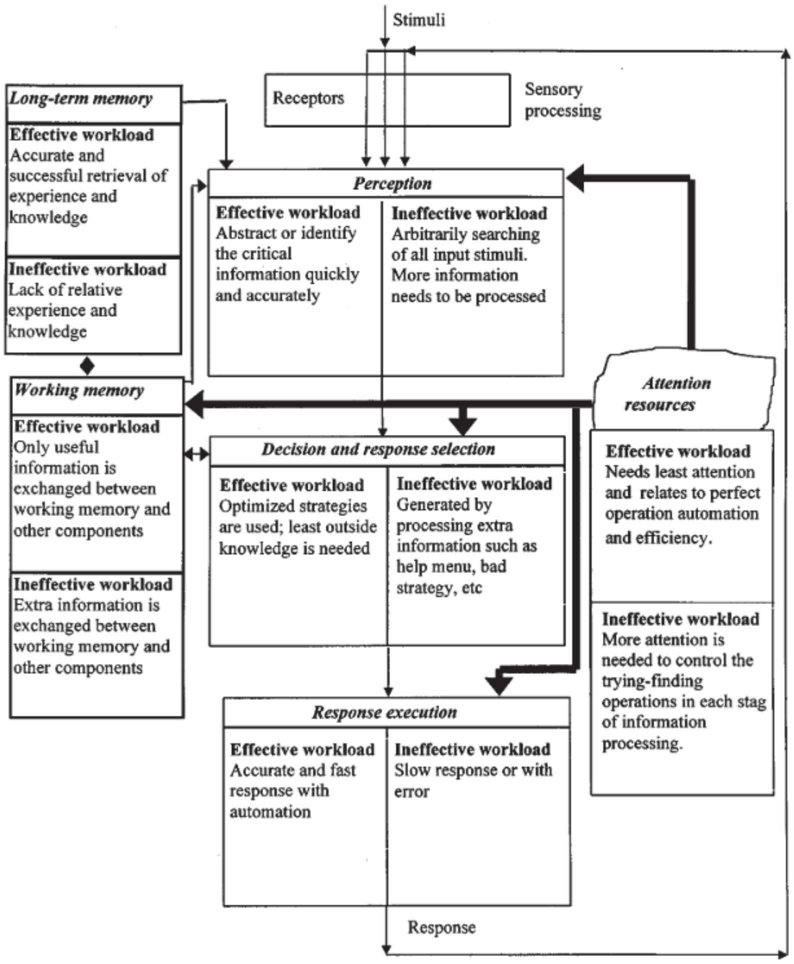

Figure 5 Effective-ineffective CL in information processing, suggested by Xie and Salvendy [102]

Furthermore, Xie and Salvendy proposed the conceptual model for predicting CL (Figure 6), considering effective and ineffective CL. These where quantified through factors related to task (external factors such as system-paced or operator-paced task type, task complexity, uncertainty, duration), individuals (internal factors such as stress, knowledge, fatigue, physical skill), and those factors that affect degrading factor K (such as task importance, attitude, motivation) [102]. In their framing, factor K is used to account for conditions that amplify or attenuate the consequences of CL, reflecting that the same nominal task demands can produce different observable outcomes depending on factors such as motivation, perceived task importance, or attitude. In other words, K does not introduce a new type of load; rather, it represents a moderating influence on how CL translates into performance degradation under particular contextual and individual states [102].

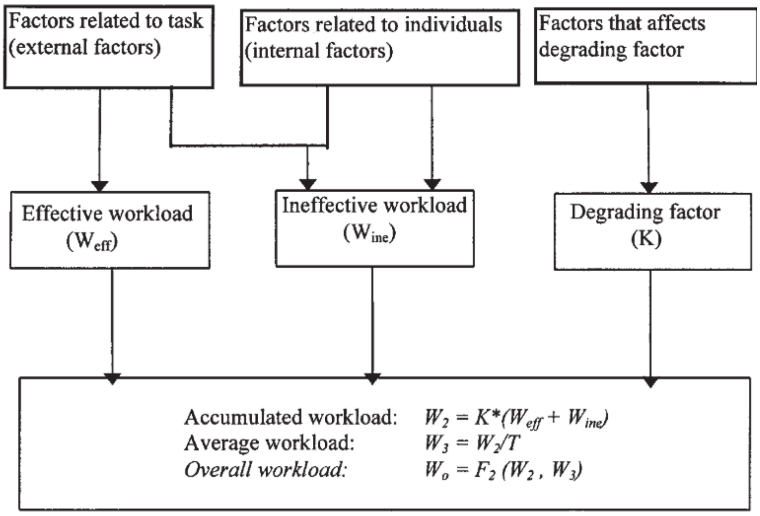

Figure 6 Model for predicting CL, suggested by Xie and Salvendy [102]

The common denominator to these three models is division of CL into two dimensions. One dimension stands for the CL that is task-centred, in a sense that it is imposed by the characteristics of the tasks and does not depend on the other factors. Paas and Van Merriënboer [107] refer to this dimension as the mental load, Xie and Salvendy [102] as the effective load, and Sweller at al. as the intrinsic load [13]. The second dimension depends on the other factors, primarily characteristics of the subject and the environment in which the task is performed. Paas and Merriënboer [107] cover this second dimension under the concept of mental effort, Xie and Salvendy [102] refer to it as ineffective load, and Sweller et al. [13] as extraneous load.

In CAD terms, the task-centred component (often aligned with intrinsic/effective load) is driven by the inherent geometric and CAD model complexity defining what must be modelled and reasoned [109], and the need to coordinate multiple information streams (e.g. interpreting technical drawings while tracking reference geometry, constraints, and feature dependencies). These demands persist even when the engineering designer is highly skilled and the interface of the CAD system is well designed, because they stem from the structure of the CAD modelling problem itself. By contrast, the avoidable component (often aligned with extraneous/ineffective load) can be amplified by suboptimal formats for representing design information, inefficient interface layouts, or workflows that force unnecessary visual search, added WM burden (e.g. by keeping multiple dimensions or constraint in mind due to limited external support), or avoidable actions (such as redundant view changes, repeated menu navigation) and rework. Moreover, because the same CAD modelling goal can be achieved via different strategies and interaction sequences (e.g. sketch-first vs. feature-first, constraint-driven vs. geometry-driven approaches), engineering designers may incur different amounts of avoidable CL even under identical task goals and internal complexity. In that sense, generic conceptual models of CL provide a valuable baseline for formalising CL in the CAD context; they guide separation of what is inherently demanding about the CAD modelling problem from what becomes demanding due to representation, interface, or strategy choices, which informs the interpretation of performance differences and motivates process-level analyses in later sections. Conceptual models presented in the following section, developed within engineering design literature, build on these generic conceptual models and distinctions they propose by accounting for mechanisms specific to engineering design, such as framing/reframing, management load, and potential load.

2.4.1.2 Conceptual models of CL in the engineering design context

Engineering design introduces characteristics relevant to CL that are less prominent in tightly specified HCI tasks. Two examples are especially instructive here: 1) engineering design often requires managing multiple concurrent concerns (requirements, constraints, feasibility checks, alternatives), and 2) engineering design process frequently depends on reframing (reinterpreting the problem and exploring the solution space) where added cognitive effort can sometimes be productive rather than wasteful. The two models below were developed for engineering design more broadly, rather than for CAD modelling tasks specifically, but they make these engineering-design-specific characteristics explicit and point to CL dimensions that may also surface in CAD activities, particularly when CAD is used to solve, coordinate, and revise open engineering design problems.

Lavrsen and Daalhuizen built upon the generic models of CL by proposing a conceptual framework mapping the relationship between CL and the process of framing and reframing in engineering design [110]. Their conceptual model, depicted in Figure 7 highlights the contribution of management load, related to managing, prioritising, and shifting between multiple tasks, as a part of the ineffective load. Additionally, they introduce the potential load; the CL actively contributing to the mental load by reframing to expand the design space and enhance the creative potential and quality of an output on a design task. Potential load stems from design problems being wicked and having multiple potential and appropriate solutions, allowing that the same problem can be framed in many ways. Lavrsen and Daalhuizen argue that this new type of load can neither be classified as extraneous load as removing it would result in a poorer design outcome nor does it fit neatly into the category of ineffective load since the added CL is not the result of ineffective strategies or lack of training [110].

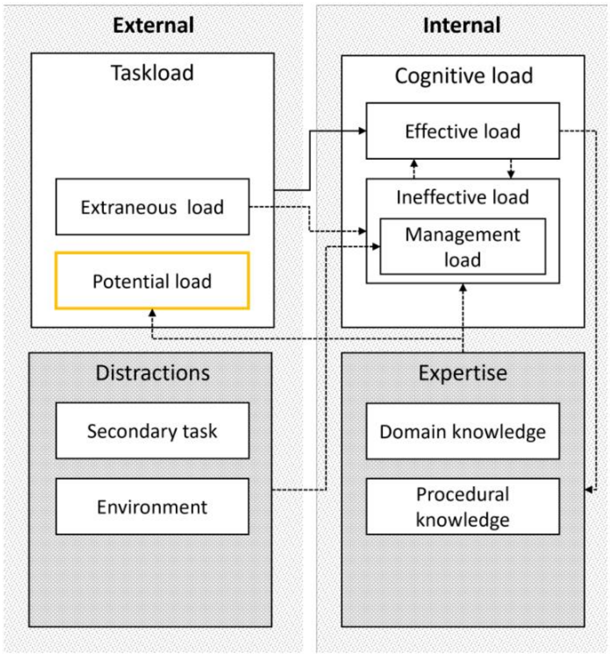

Figure 7 Model of CL including potential load, suggested by Lavrsen and Daalhuizen [110]

Moreover, Zhao et al. investigated CL in the context of design creativity, through the constructs of perceived WL, mental stress, and mental capacity [111]. In particular, they define mental stress in a mathematical expression (Equation (2)) as a ratio of perceived WL and mental capacity (calculated as a sum of knowledge and skills multiplied by affect), but keep it at the conceptual level rather than prescribing how to measure and calculate each construct.

$$Mental\ Stress = \frac{Perceived\ Workload}{(Knowledge + Skills) * Affect}, \quad \text{Affect} \in (0,1) \tag{2}$$

Therefore, scholars assume a positive relationship between mental stress and WL, and negative between mental stress and mental capacity. They define WL as an external load assigned to an engineering designer, whereas mental capacity is the engineering designer's ability to handle that external load. External WL is seen as the direct source of mental stress. According to Zhao et al., WL can be associated with the complexity of the problem. However, it is possible that the same WL triggers different mental stresses on different individuals and the same WL may trigger different mental stresses for the same individual under different circumstances. The effect depends on a designer's mental capacity, defined by knowledge, skills, and affect. Moreover, the lack of skills and knowledge for the current design problem may increase mental stress. Knowledge is influenced by 1) its structure and organisation, and 2) the availability of cognitive resources for retrieving information from the LTM. Skills refer to thinking styles, thinking strategies, and reasoning methods. Finally, engineering designers' affect determines how well their knowledge and skills will be used in the engineering design process. For example, uncertainty and unpredictability of engineering design tasks may increase mental stress on engineering designers if they perceive them as beyond their mental capacity.

Together, these two conceptual models suggest that a portion of CL observed in engineering design reflects 1) coordination and switching between parallel design concerns (management load) and 2) productive exploration that expands and restructures the solution space (potential load), while experienced CL is moderated by engineering designers' capacity related to knowledge, skill, and affect. At the same time, CAD modelling tasks often constrain problem solving and reduce reframing, meaning that CL specific to engineering design may appear only partially. This leads into the next section on cognitive models of CL, which moves beyond conceptual distinctions toward explicit formulations of how CL emerges and evolves during task execution, based on cognitive architectures and task analysis frameworks.

2.4.1.3 Cognitive models of CL

Building on Section 2.3, which introduced cognitive architectures and task analysis as foundations for modelling performance in HCI tasks, this section focuses on how these architectures have been used to quantify CL or WL during task execution. Although cognitive models built in accordance with cognitive architectures represent users' perceptual, cognitive, and motor processes and predict performance, they do not offer built-in functions to quantify CL or WL [112]. Consequently, special effort has been devoted in these studies to developing mathematical equations through which CL can be quantified from the activity of the modules represented in cognitive architectures. Such a quantification of CL from cognitive architectures builds upon the definition of CL as the ratio of the resources required to the resources available [113]. Similarly, Just et al. proposed that the capacity utilisation, which refers to the proportion of resources consumed in a given period, can serve as a representation of CL [114]. Hendy et al. applied the same logic to information processing and argued that the ratio of the time necessary to process the required information to the time allowable for making a decision determines WL and performance of the human information processor [115].

The reviewed studies span both continuous control contexts (e.g. driving [116], [117] or air traffic control [118]) and discrete HCI tasks (e.g. menu selection [112], [119] or judgement tasks [120], [121]). A cognitive architecture (e.g. ACT-R, QN-MHP) is first used to represent a task in terms of perceptual, cognitive, and motor processing. CL/WL is then computed from the model/s internal activity, most commonly using 1) capacity utilisation (how "busy" a module or subnetwork is over time) and/or 2) module activation time (how long a module remains engaged). Because many studies treat NASA TLX as a widely used criterion measure ("gold standard") for subjective WL, researchers then often compute WL derived from the cognitive model over the full task duration and then test whether it aligns with the reported NASA TLX outcomes. When strong alignment is found, some authors fir linear regression mappings that translate indices derived from the cognitive model into predicted NASA TLX scale values. Validation is in rare cases complemented by psychophysiological measures (e.g. respiration rate [117]). Table 9 summarises the cognitive architectures, task contexts, and validation approaches used in these studies.

Table 9 Cognitive architectures and WL/CL quantification

Reference	Task(s)	Cognitive architecture	Subjective	Physiological	Multiple sources of WL	Empirical validation	Additional factor
[116]	Driving	QN-MHP	NASA TLX	None	Yes	Yes	Age
[120]	Semantic judgment task	ACT-R	NASA TLX	None	Yes	Yes	-
[112]	Memorisation, visual-manual, menu selection task	ACT-R	NASA TLX	None	Yes	Yes	Weighted values for different modules, Error
[119]	Menu selection, visual-manual task	ACT-R	NASA TLX	None	Yes	Yes	Difficulty, time pressure
[121]	Cognitive judgment, target-shooting, arithmetic computation	ACT-R	NASA TLX; SAT	None	No	Yes	Time pressure
[118]	Air traffic control	ACT-R	NASA TLX	None	Yes	Yes	Time pressure
[117]	Driving takeover	ACT-R	Questionnaire	Respiration rate	No	Yes	Decay

Wu and Liu focused on extending the model of driver performance to cover mental WL in driving while accounting for age differences in driving performance and WL [116]. To achieve that goal, scholars proposed the average utilisation of a subnetwork in the cognitive model based on QN-MHP cognitive architecture as an index of CL [116]. In their notation, utilisation is expressed as ρ_i, where ρ_i is computed for subnetwork i as the proportion of the processing capacity that is consumed during the task (see Equation (3) in **Error! Reference source not found.**). Here, λ_i denotes the arrival rate of processing "jobs" into subnetwork i, Ci the number of QN servers in that subnetwork, and μ the server processing speed. To represent age effects, Wu and Liu introduce aging factor A that reduces effective processing speed, thereby increasing utilisation and predicted WL (see Equation (4) in **Error! Reference source not found.**). Their modelling approach proved to be successful in simulating driving performance, visualising dynamically driver mental WL, modelling age differences, and predicting NASA TLX rates.

Table 10 Mathematical equations for calculating CL/WL from cognitive models, adopted from Wu and Liu [116]

Weights/Factors	Equation for calculating WL/CL	
A is aging factor (A ≥ 1). λ_i is the arrival rate of the subnetwork *i*; C_i is the total number of servers in the subnetwork *i*; $\mu_{0,j}$ is the original processing speed of server *j* for the young adults.	$\rho_i = \frac{\left(\int_o^T \rho_i dt\right)}{T}, (0 \le \rho_i \le 1),$ where ρ_i is average utilisation of subnetwork *i* in task time T	(3)
	$\rho_i = \frac{\lambda_i}{\sum_{j=1}^{C_i} \left(\frac{1}{A}\right) \mu_{0,j}}$	(4)

Linear regression equations (Equation (5) to Equation (10)) relating the calculated CL scores with NASA TLX rates are available in Table 11. An equation is available for each NASA TLX dimension, representing an individual aspect of WL: Physical Demand (PD) in Equation (5), Temporal Demand (TD) in Equation (6), Effort (EF) in Equation (7), Performance in Equation (8), Frustration in Equation (9), and Mental Demand (MD) in Equation (10).

Table 11 Linear regression equations for relating WL from cognitive models and NASA TLX, adopted from Wu and Liu [116]

Parameters	Linear regression models	
Parameters a and b are constants in representing the direct proportional relation between the averaged utilisations and the subjective responses (a > 0). a and b were estimated only for the PD scale, the same value is used in estimating the subjective responses on the other five scales.	$PD = a\rho_m + b, (0 \le PD \le 100)$	(5)
	$TD = a\left(\sum_{All\ i} \rho_i\right)/4 + b, (0 \le TD \le 100)$	(6)
	$EF = a\left(\sum_{All\ i} \rho_i\right)/4 + b, (0 \le EF \le 100)$	(7)
	$PE = a\left(\sum_{All\ i} \rho_i\right)/4 + b, (0 \le PE \le 100)$	(8)
	$FR = a\left(\sum_{All\ i} \rho_i\right)/4 + b, (0 \le FR \le 100)$	(9)
	$MD = a\left(\sum_{ap,vp\ c} \rho_i\right)/3 + b, (0 \le MD \le 100)$	(10)

Cao and Liu modelled mental WL in semantic judgment category tasks using QN-ACTR cognitive architecture, relying on the similar idea the WL can be expressed as capacity utilisation (WL increases when the model requires more processing time relative to time available) [120], [121]. In its simplest form, this is expressed as an expected utilisation ratio, which increases under higher task demands and time stress (see Equation (11)).

$$Expected\ Utilisation = \frac{Time_{required}}{Time_{total}} \tag{11}$$

The developed model was validated with experimental data; it generated correct performance results (aligned modelled and experimentally observed timing) and NASA TLX rates similar to the corresponding human scores in the NASA TLX questionnaire (see Equation (12) and

Equation (13) in Table 12). Additionally, P, C, and M components of mental WL were visualised and showed the increase of mental WL with greater task demands. Moreover, in their later work Cao and Liu captured the difference between various tasks demand conditions in three different tasks (cognitive judgment, target-shooting, arithmetic computation) under time stress [121]. Their results indicate that the proposed WL modelling technique based on QN-ACTR cognitive architecture could support testing and evaluation of complex HCI system in various domains.

Table 12 Linear regression model relating WL from cognitive models and NASA TLX, adopted from Cao and Liu [120], [121]

Parameters	Linear regression model	
a = 570.4; b = -5.8	$Overall\ WL = -5.8 + 570.4 \cdot Overall\ U$	(12)
a = 482.9; b = -2.1	$Overall\ WL = -2.1 + 482.9 \cdot Overall\ U$	(13)

Jo et al. proposed a method to quantitatively predict mental WL (instantaneous IW, accumulated AW, and average W_{ave}) from ACT-R by expressing mental WL as a function of module activation time, $A_i(t)$, which indicates whether module i is active at time t (and how long it remains engaged) [112]. In that sense, they use the ratio of the total time spent for firing productions and exchanging chunks through the ACT-R modules to the available time for the given task as the index of mental WL. They define instantaneous WL (IW) as a weighted sum across modules (Equation (14) in Table 13), where W_i is a weight reflecting the assumed relative contribution of module i to WL. It increases from the lowest with perceptual-motor modules to the higher mental WL with central modules, and finally up to the highest WL with memory modules. $Ei(t)$ is an error-weighting factor that increases WL during error states. Integrating instantaneous WL over time yields accumulated WL, AW (Equation (15) and Equation (16) in Table 13). Dividing it by task duration T yields average WL, W_{ave} (Equation (17) in Table 13). Results revealed that the predicted values of mental WL achieved by the proposed method were highly correlated with the mean NASA TLX subjective ratings in three tasks: memorization, visual-manual, and menu selection task (see Equation (18) in Table 14 for linear regression model) [112].

Table 13 Mathematical equations for calculating CL/WL from cognitive models, adopted from Jo et al. [112]

Weights/Factors	Equation for calculating WL/CL	
Wi is weight assigned on module i $W_{pm} = 1$, $W_{central} = 2$, $W_{memory} = 4$. Ei(t) is the error weight function; Ei(t)=2 in error, Ei(t) =1 otherwise	Instantaneous WL: $$(\mathrm{IW})_i = W_i \cdot E_i(t) \cdot \left[\frac{(A_i(t)dt)}{dt}\right] = W_i \cdot E_i(t) \cdot A_i(t)$$ $$(\mathrm{IW})_i = \sum_i [W_i \cdot E_i(t) \cdot A_i(t)]$$	(14)
	Accumulated WL: $$(\mathrm{AW})_i = \int_0^T (\mathrm{IW})_i dt = \int_0^T W_i \cdot E_i(t) \cdot A_i(t) dt = W_i \cdot E_i(t) \int_0^T A_i(t) dt$$	(15)

	Accumulated WL (total): $$(\mathrm{AW})_{total} = \sum_i (\mathrm{AW})_i = \sum_i \left(W_i \cdot E_i(t) \int_0^T A_i(t)dt \right)$$	(16)
	Average WL: $$\mathrm{W}_{ave} = \frac{(\mathrm{AW})_{\mathrm{total}}}{T} = \sum_i ((\mathrm{AW}_i)/T) = \sum_i \left(\mathrm{W}_i \cdot \mathrm{E}_i(t) \left(\int_0^T \mathrm{A}_i(t)dt/T \right) \right)$$	(17)

Table 14 Linear regression model for relating WL from cognitive models and NASA TLX, adopted from Jo et al. [112]

Parameters	Linear regression model	
a = 12.08; b = 1.62	$(\text{NASA TLX})_{mean} = a \cdot \mathrm{W}_{ave} + b$	(18)

Park and Myung modelled performance in menu selection and visual-manual tasks varied by task difficulty and time pressure [119]. In their formulation (Equation (19) in Table 15) they proposed a mathematical representation of the modelled CL over time T_a with respect to an activated time of each module from ACT-R. Additionally, this mathematical representation accounts for the acceleration of information processing by time pressure through variables such as T_i (module time under pressure), $\mathrm{T}_{0,i}$ (baseline required time at each module *i* without time pressure), and C (processing rate representing human capacity).

Table 15 Mathematical equations for calculating CL/WL from cognitive models, adopted from Park and Myung [119]

Weights/Factors	Equation for calculating WL/CL	
Time pressure: $$\mathrm{T}_i = \left(\frac{1}{\mathrm{C}}\right) \cdot \mathrm{T}_{0,i}$$ C is processing rate representing human capacity. $\mathrm{T}_{0,i}$ is original required time at each module *i* without time pressure.	$\mathrm{W}_i = \mathrm{w}_i \cdot \mathrm{E}_i(t) \cdot (T_i/T_a)$	(19)

Their predicted values related to physical demand (PD; see Equation (20) in Table 16), mental demand (MD; see Equation (21) in Table 16), and temporal demand (TD; see Equation (22) in Table 16) were highly correlated with mean values of subjective ratings from subjects, implying that their method can reliably predict task-related properties as long as the ACT-R model can properly represent human performance in a given task. In a follow-up study, Park et al. extended this approach to air traffic control scenarios, using ACT-R to simulate how CL evolves with increasing traffic complexity [118]. They proposed six mathematical equations (see Equation (20) to Equation (25) in Table 16) to estimate all sub-scale dimensions of NASA TLX (PE, EF, FR in addition to previously mentioned PD, MD, and TD**Error! Reference source not found.**), respecting multiple sources of mental WL and the effects of time pressure. The proposed method accurately modelled human performance (based on time predictions) and highly correlated with the subjective ratings of each scale of NASA TLX from participants in diverse conditions, varied by the demanding resource and the time pressure [118].

Table 16 Linear regression model for relating WL from cognitive models and NASA TLX, adopted from Park and Myung [119],[118]

Parameters	Linear regression model	
$a_{PD} = 38.58;\ b_{PD} = 8.54$	$(\mathrm{PD})_{mean} = a \cdot \mathrm{W}_m + b,\ 0 \leq \mathrm{PD} \leq 100$	(20)
$a_{MD} = 22.04;\ b_{MD} = 15.12$	$(\mathrm{MD})_{mean} = a \cdot \left(\sum_{i=vp,ap,c} \mathrm{W}_i \right) + b, 0 \leq \mathrm{MD} \leq 100$	(21)
$a_{TD} = 22.55;\ b_{TD} = 2.77$	$(\mathrm{TD})_{mean} = a \cdot \left(\sum_{all\ i} \mathrm{W}_i \right) + b, 0 \leq \mathrm{TD} \leq 100$	(22)
Parameters a and b are the constants in representing the direct proportional relation between the mean value of scores on NASA TLX dimensions and the average WL of perceptual, central, and memory modules.	$PE = a \cdot \frac{T_r}{T_a} + b$	(23)
	$EF = a \cdot \left(\sum_{all\ i} (W_i) \right) + b$	(24)
	$FR = a \cdot \left(\sum_{i=p,c,m} (W_i) \right) + b$	(25)

Oh et al. introduced a method to computationally predict mental WL during takeover situation in automated driving, using the ACT-R cognitive architecture [117]. In their formulation, WL accumulated across subtasks *j* is discounted by a time factor t_j^{-d}, accounting for a decay parameter *d*, which down-weights earlier WL contributions over time (Equation (26) and Equation (27) in Table 17). The model predictions were validated against physiological data (respiration rate) and performance data from the experiment. Their model showed high accuracy (R^2=0.97), indicating a strong correlation between the predicted and “actual” (physiologically indicated) mental WL.

Table 17 Mathematical equations for calculating CL/WL from cognitive models, adopted from Oh et al. [117]

Weights/Factors	Equation for calculating WL/CL	
d is a WL decay parameter; *j* is a subtask	$\mathrm{AW}_{[\mathrm{T}_1,\mathrm{T}_2]} = \sum_i \left(\mathrm{W}_i \int_{\mathrm{T}_1}^{\mathrm{T}_2} \mathrm{A}_i(t)dt \right)$	(26)
	$\mathrm{MW\ (RT)} = \frac{\sum_j (\mathrm{AW})_j \cdot t_j^{-d}}{\mathrm{RT}}$	(27)

Taken together, these studies show how cognitive architectures can be used not only to simulate performance in terms of perceptual, cognitive, and motor processing but also to produce interpretable, temporally varying CL/WL indices grounded in module/subnetwork utilisation or activation. In this way, the models provide a basis for linking engineering designers' embodied cognition, task structure, and information-processing demands, which aligns with the thesis' process-oriented view of CAD performance. Importantly, the corresponding equations make explicit how additional factors can be accounted for in the CL estimates; either by modifying processing capacity (e.g. ageing factor A), re-weighting contributions of different processing components (e.g. module weights Wi), by penalising error states (e.g. through error factor $Ei(t)$), or by discounting earlier demands (e.g. with decay parameter d). However, because utilisation/activation primarily capture the temporal occupation of resources, they do not necessarily reflect differences in information-processing intensity per unit time; this motivates complementing time-based indices with psychophysiological measures, which are sensitive to changes in processing intensity (see Section 2.4.3.1).

At the same time, task contexts of the previous studies largely belong to domains with clearly specified procedures and measurable timing constraints (e.g. driving scenarios, menu selection, judgment tasks, air traffic control). This matters for CAD because these approaches to model CL are most straightforward when activity can be represented as a well-defined sequence of segments or repeatable HCI procedures. They are therefore easiest to apply to CAD work that can be segmented into recurring CAD modelling segments (e.g. sketch creation, constraint application, feature execution), and less direct for exploratory CAD modelling with frequent revisions, goal shifts, and opportunistic strategy changes unless these behaviours are explicitly scoped and encoded. Finally, the reviewed validation practice (most often using NASA TLX and occasionally psychophysiological methods) highlights that computational CL/WL indices ultimately depend on how WL is measured and interpreted in empirical CAD studies. For this reason, the next section reviews studies that have investigated CL in (engineering) design tasks, focusing on their task contexts, experimental variables, and measurement approaches to clarify what might complement CL/WL indices derived from cognitive models.

2.4.2 Studies on CL in (engineering) design tasks

The reviewed literature includes a diverse set of studies that examined CL in (engineering) design and related HCI tasks, spanning different design activities, engineering or design contexts, and measurement approaches. Their key methodological details are summarised in Table 18, which reports the experimental task, the independent variables investigated (where applicable), and the method used to assess CL in each study, while also indicating whether CL was explicitly related to performance outcomes. Collectively, these studies reinforce that CL is a multidimensional construct, proved to be influenced by many factors representing the characteristics and interactions between three core elements: the subject, the task, and the working environment in which the task is carried out [13]. To support synthesis, the findings discussed in the following sections are organised according to whether the experimental variables relate primarily to 1) engineering designers' characteristics, 2) methods and tools, or 3) task characteristics.

Table 18 List of studies on cognitive load in (engineering) design tasks

Reference	Experimental task	Independent variable(s)	Method	P*
[86]	Learning how to use CAD/CAM application	Instructional format Internal task complexity Medium	Dual tasks	-
[122]	Learning how to use CAD application	Instructional format Medium	Implied from the primary task performance	-
[123]	Designing a house	Method (sketching vs no sketching)	Protocol analysis	
[15]	Designing a teaching podium	Experienced designers vs novices	RSME	↑
[16]	Designing using different media	Medium (hand sketches, physical models, CAD software)	Dual tasks	-
[124]	Building the structure from different representations	Engineering information format (2D drawing, 3D CAD model, 3D printed model) Medium (paper, computer, physical world) Experience and expertise	NASA TLX	↓
[125]	Diagram comprehension	Diagram type Layout type	ET	
[126]	Solving three open-ended design problems: designing a house, a vehicle, and a desk	Mental stress	EEG	-
[127]	Arranging the furniture within a small office space	Interface type (graphical VR vs. tangible AR)	NASA TLX	
[128]	Playing a 2D computer game	Interface design	Questionnaire; Performance (duration)	
[129]	Solving six open-ended design tasks	CL measurement method	EEG; NASA TLX	
[130]	HCI task: visual and motor processing	Internal task complexity	EEG	
[131]	Observing two representations or architectural design	Design representation	ET	
[132]	Solving engineering problems	Task structuredness and complexity	EEG	-
[133]	Visuospatial task: origami	Virtual environment (2D vs. 3D) Medium (desktop vs. VR) Spatial visualisation ability	EEG	
[134]	HCI task: simulation of a tube rupture	Expertise (skilled vs. non-skilled) Method (with vs without support system)	EEG; NASA TLX	↓
[135]	Solving six open-ended design tasks	Type of conceptual design activity	EEG	
[136]	Manual task: engineering (assembly)	Instruction format (2D vs. 3D) Instruction medium (paper vs. AR)	EEG; NASA TLX	-
[137]	Drawing a table in AutoCAD	Inputs: keyboard/mouse vs. speech/gesture	EEG	
[138]	Solving listing and concept mapping tasks related to sustainability issues: climate	Method (concept mapping vs. listing)	EEG; NASA TLX	-

	change, food systems, renewable energy, and water availability			
[139]	Concept generation using brainstorming, morphological analysis, and TRIZ	Method (brainstorming, morphological analysis, TRIZ)	fNIRS	
[140]	Concept generation, concept selection, and physical modelling	Activity type (concept generation, concept selection, physical modelling)	NASA TLX	
[141]	Solving six open-ended design tasks	Conceptual design activity type	EEG; NASA TLX	
[17]	Analysis of technical drawing, functional analysis of assembly, considering manufacturing processes in design, failure analysis, load analysis in machine elements	Internal task complexity	NASA TLX	↓
[142]	Learning how to use additive manufacturing systems	AM processes (material extrusion, powder bed fusion)	Workload Profile Assessment	-
[143]	Designing in parametric design environments	-	Retrospective and concurrent protocol analysis, FBS ontology, linkography	
[144]	Redesigning for manufacturing (concept generation selection)	Experience (students vs. practitioners)	ET; NASA TLX	
[18]	Ideation and prototyping	Task (ideation vs. prototyping)	NASA TLX	↑
[145]	Solving a DfAM task	Design representation and subtasks (sketching, writing, CAD modelling, simulating build)	ET; NASA TLX	

*Column P stands for Performance; symbol in this column means that study related CL to performance; "↑" indicates positive significant relationship; "↓" indicates negative significant relationship; "-" indicates no significant relationship found.

2.4.2.1 Engineering designer's characteristics

Previous studies assumed that subjects may endure more CL than that the task actually imposes on them because of their lack of **knowledge or skill**, useless or redundant actions, or stress [102]. In particular, it seems to be possible that skilled subjects perform a complex and difficult task almost effortlessly. At the same time, a task that seems very simple and easy may be very difficult for subjects that lack required knowledge and skills [102]. Surprisingly, Sun and Yao found that experienced designers reported higher scores of mental effort (dimension through which they operationalise CL) when solving a design task [15]. They explained these results by more experienced designers' better understanding of the requirements and consequent employment of more reasoning methods. Similarly, Dadi et al. found that subjects with the least amount of engineering (construction) experience reported the lowest CL; in this case, scholars proposed familiarity of these subjects with virtual environments as a possible explanation of revealed results [124]. Moreover, Dadi et al. revealed statistically significant differences in CL related to demographic factors, such as occupation (**technical background**) and **CAD experience** levels [124]. More precisely, subject who use technical drawings daily felt less time pressure than subjects who never use them in their daily work. Similarly, perceived performance levels were affected by the level of CAD experience; higher levels of CAD experience led to subjects' more positive perception of their performance [124].

2.4.2.2 Methods and tools

Furthermore, scholars have varied the **methods and HCI tools** used to perform design tasks and examined how such variations affect CL [16], [123], [124], [127], [137], [139], [145]. For example, Shealy et al. reported that CL depends on **the concept generation technique** employed (brainstorming, morphological analysis, and TRIZ) [139]. Moreover, Barella et al. proved their assumption that making a concept mapping is more complex and induces higher CL than performing the listing in the same design task [138]. Bilda et al. explored the role of sketching for design process and concluded that CL is much lower when designers are allowed to sketch in comparison to processes in which there is no possibility to store information from WM to sketches since sketching is forbidden [123]. Choi et al. found reductions in CL and improvement in engineers' performance when applying the **support method**, with stronger effects for non-skilled than skilled users [134].

Baig and Kavakli found that drawing and manipulating 3D objects using speech and gesture is more challenging and imposes higher CL on expert users than a conventional **input setup** of **HCI tools** (keyboard and mouse); however, their findings were the opposite for novice CAD users [137]. Young et al. explored how altering the mapping of inputs and outputs in a 2D computer game affects CL [128]. The results of their study showed that the brain can cope with input/output changes when they occur at the same dimensions (for example swapping reversing the actions on the X or Y axis), but struggles to cope with input changes across different dimensions (for example swapping the actions of the X and Y axis) [128].

Studies also investigated **media** and virtual environments they mediate for recording and interacting with visual design representations [146]. For instance, Dan and Reiner revealed significant differences in CL depending on whether instructions were provided with 2D video presented on a desktop or in a stereoscopic 3D virtual environment [133]. Additional findings implied that benefits of 3D virtual environments were more obvious for participants with lower spatial visualisation abilities [133]. Similarly, Mohamed-Ahmed et al. searched for differences in CL and performance when using hand sketches, physical models, and CAD software in design tasks, but did not reveal significant differences when using these three types of media [16].

In addition, Cass and Prabhu suggested that CL changes with the **type of design representation** (e.g. sketches, written notes, CAD models, virtual simulations) engineering designers use during the process [145]. Similar results were reported by Kosch et al.; their finding indicate that in-situ projections of assembly instructions using augmented reality reduce CL compared to paper instructions [136]. On the contrary, some studies reported no CL differences across representation formats, such as computer-generated hidden-line perspective vs. digital photograph of the same space [131] or 2D drawings vs. 3D CAD interface vs. 3D printed model [124]. Martin-Michiellot and Mendelsohn reported no significant differences in CL or learning performance between different **formats of presentation of instructions**: manual plus a computer, a manual with juxtaposed screen images but no computer, a manual with integrated screen images but no computer [122].

2.4.2.3 Task characteristics

In addition to subjects' characteristics and the used methods/tools, previous studies suggest that **design activities and tasks** induce unique CL dependent on the underlying information processing related both to type and amount of information [17], [18], [132], [140], [141]. For instance, Jia et al. suggested that idea generation imposes higher CL than problem understanding, idea evaluation, and self-rating in open-ended design tasks [141]. Calpin and Menold reported differences in CL between idea generation and prototyping implying that idea generation is more cognitively demanding than prototyping [18]. Zimmerer and Matthiesen compared CL in several typical engineering tasks and concluded that failure analysis is more cognitively demanding then analysis of technical drawing, functional analysis, considering manufacturing processes in design, and load analysis in machine elements. In another study, Nolte and McComb reported findings that relate physical modelling with the highest CL, followed by concept generation and finally selection [140].

Task **internal complexity** is a frequently studied driver of CL, often explained via element interactivity; CL increases when elements must be processed simultaneously in WM because they are logically related [13]. Task complexity has also been quantified through 1) the number and nature of diverse set of subskills that are a part of the main skill involved in solving a task, and 2) the complexity of the goal hierarchies of the problems that define the task [107]. In the context of engineering design, Majdic et al. reported that CL arises with the difficulty of design tasks, where difficulty was defined by the complexity and structuredness of the task [132]. Schultze-Kraft et al. asked subjects to execute a task on a touch screen that required continuous effort of visual and motor processing with alternating difficulty, achieved by varying time between the stimuli and the level of randomness of presented content [130]. They found differences in brain activity measured with EEG, related to CL distinction between easy and difficult tasks [130].

2.4.2.4 Relationship between CL and CAD performance

Beyond identifying factors that shape CL, a subset of studies examined how CL related to engineering designers' performance; most often operationalising it through the quality of design outputs, such as the creativity of the creativity of design solutions [15]–[18]. Findings on this relationship have been mixed. For example, Calpin and Menold found a significant, moderately positive monotonic relationship between CL and the uniqueness, usefulness, and elegance of ideas [18], whereas Mohamed-Ahmed et al. found no relationship between CL and design outcomes [16]. In contrast, Zimmerer and Matthiesen reported a negative correlation between CL and performance in engineering tasks [17]. Related work also investigated how different types of provided instructions influence experienced CL and subjects' performance in assembly task (e.g. [136]) or during learning in engineering design (e.g. [86], [122], [132], [138], [142]). For instance, Chandler and Sweller [86], as well as Martin-Michiellot and Mendelsohn [122], related better success in learning how to use CAD applications to instructions designed in a way to reduce CL. Moreover, Nguyen and Zeng investigated how CL levels may lead to creative design solutions by redesigning methods and tools [126]. They proposed a curvilinear relationship between CL and design creativity, implying that design outputs improve with increasing CL up to an optimal level, beyond which excessive CL leads to performance decline. Notably, no prior studies have directly investigated relationships between CL and performance specifically in CAD tasks. However, related evidence from studies involving engineering information interpretation and transformation (often contained in CAD tasks) suggests that lower CL can be associated with higher productivity and accuracy. For example, Zimmerer and Matthiesen reported a weak but significant negative monotonic relationship between CL and task accuracy across six engineering design tasks (e.g. technical drawing analysis, functional analysis, design for manufacturing) [17], while Dadi et al. found that lower self-reported CL was associated with higher productivity (less percentage of non-effective work and rework) in reconstructing structures based on different engineering information formats (2D drawings, 3D CAD models, and 3D printed models [124].

In summary, prior work shows that CL in engineering design is shaped by interacting factors related to the engineering designer, the employed methods and tools, and the task itself, but the strength and maturity of evidence differ across these streams and across measurement approaches. Studies manipulating designer-related variables (e.g. experience, domain familiarity, demographics) indicate that CL is not determined by task demands alone and can vary counterintuitively with expertise, whereas research on methods and tools highlights that CL can be influenced by interaction modalities, support methods, and representation media, although findings are often context-dependent and sometimes conflicting across formats. Task-focused work provides comparatively clearer evidence that CL differs across design activity types and increases with internal complexity, yet these studies rarely resolve how CL fluctuates within longer, realistic design processes. Importantly, only a subset of the reviewed literature explicitly links CL to performance, and the reported relationships are mixed (positive, null, negative, and potentially curvilinear), with no direct evidence yet established for CL–performance relationships in CAD tasks. The overview further highlights a key gap for the CAD domain; only three studies have investigated CL directly during CAD activity ([86], [122], [137]), and only Baig and Kavakli [137] have done so in a way that supports continuous, temporally sensitive assessment that can be aligned with CAD process dynamics and performance. This gap motivates a closer examination of how CL can be measured in realistic design settings and which measurement approaches best support non-invasive, dynamic monitoring during CAD tasks. The next section therefore discusses methods for measuring CL, with particular emphasis on psychophysiological approaches and EEG-based indicators that enable capturing CL variations throughout the CAD modelling process.

2.4.3 Measuring cognitive load

CL is a covert construct, which cannot be measured directly but only indirectly through other variables [13], [102]. As a result, measuring CL in reliable, objective, and continuous manner remains a challenging task, especially in complex and lengthy task like those in CAD. One absolute measure for quantifying CL does not exist; instead, methods employed for studying CL measure variables that are thought to highly correlate with changes in CL [102]. Given its multi-dimensionality, combination of more than one method is typically needed to measure CL. Consequently, scholars have used various method to measure CL in HCI [112]. First, these methods of CL assessment are classified in Table 19 into two broad groups, depending on whether they are analytical or empirical [102], [105].

Table 19 Methods for measuring CL

Analytical	Empirical
TASK ANALYSIS METHODS Structural hierarchical decomposition of tasks and goals. • Task analysis frameworks	PERFORMANCE ASSESSMENT: Use operator behaviour to determine CL. • Primary and secondary tasks • Metrics for quantifying quality of outcomes
MATHEMATICAL MODELS Determine relationships between identified factors (variables) that influence CL to build predictive mathematical models that relate CL values with task performance. • Human performance modelling	SUBJECTIVE RATING: Use operator subjective opinion. • NASA-Task Load Index (NASA TLX) • Subjective Workload Assessment Technique (SWAT) • (Modified) Cooper-Harper Rating Scale • Subjective Workload Dominance (SWORD) • Bedford Workload Scale • Workload Profile
COMPUTER SIMULATION MODELS Use simulated subject instead of human user/operator to exclude influence of individual differences. • Computerised versions of cognitive architectures	PSYCHOPHYSIOLOGICAL ASSESSMENT: Measure changes in the operator physiology. • Brain activity: EEG, fMRI, fNIRS • Heart rate variability: electrocardiography • EMG amplitude: electromyography • Oxygen consumption • Eye behaviour: eye tracking • Skin conductance response

Analytical methods estimate CL without directly asking subjects about their experience or collecting data from them; they are based on techniques such as mathematical or simulation models and task analysis [101], [105]. Empirical methods, on the contrary, gather data directly from subjects; performance data is collected using primary and secondary task techniques, and quantifying performance outcomes; subjective data using questionnaires and rating scales, and psychophysiological data using psychophysiological techniques [101], [105]. Examples of the most popular techniques belonging to each category are provided in Table 19.

Furthermore, methods differ in their capability to calculate CL in terms of dynamic changes and granularity of time windows for which CL is measured. Considering the time dimension, CL can be calculated as: instantaneous, peak, accumulated, average, and overall load [101], [102]. Instantaneous CL reflects the dynamics of CL, which fluctuates every moment from the beginning to the end of performing a task. Peak CL represents the max value of instantaneous load while executing a task. Accumulated CL is the total amount of load that the subject experienced during a task. Average CL represents the mean intensity of load during the performance of a task. Finally, overall CL is the load based on the whole duration of task performance.

While the classification in Table 19 summarises the main families of CL measurement methods, their applicability and interpretability depend on the context. Engineering design tasks, especially CAD modelling, combine sustained visuospatial reasoning, iterative problem solving, and frequent shifts between information acquisitions, execution, and evaluations,

which makes CL both dynamic and presumably unevenly distributed across the process. Consequently, selecting appropriate CL measures in this domain requires considering not only whether a method is analytical or empirical, but also how well it captures the temporal structure of CAD modelling and the types of cognitive demands it imposes. The remainder of the section therefore reviews how CL has been measured in (engineering) design tasks and what methodological choices have proven most suitable for capturing CL in realistic design settings.
In previous studies within (engineering) design, CL was measured using three common groups of methods: subjective measurement, performance measurement, and psychophysiological measurement methods (see Table 18 for their distribution). Subjective measurement methods, most commonly NASA TLX, alongside questionnaires such as Rating Scale Mental Effort (RSME), and the Workload Profile Assessment, are widely used because they are easy to administer and broadly applicable across tasks. However, their reliance on introspective self-report typically limits temporal granularity, making it difficult to localise periods within a design activity when CL increases without a constant interruption of the execution. Performance-based measurement methods, including dual tasks and protocol analysis [123], provide more direct behavioural evidence that can be linked to design outcomes. Yet, in real-world CAD settings, they can be intrusive; dual tasks may alter the primary task strategy, and protocol analyses can be labour-intensive and still do not provide a continuous physiological trace of changing information-processing demands. More generally, behavioural measures often reflect the consequences of CL (e.g. slower performance, errors, strategy changes, change in distribution of task segments) rather than CL changes, and they may be confounded by expertise, task familiarity, and speed-accuracy trade-offs.
These limitations motivate the growing use of psychophysiological measurement methods, such as eye tracking (ET), EEG, and functional Near-InfRared Spectroscopy (fNIRS), which can capture CL more continuously while engineering designers interact naturally with HCI tools. Among these methods, scholars most frequently selected EEG for studying CL in design (see Table 18), likely because it combines high temporal resolution (the highest among the neuroimaging methods [147]) with increasingly practical, non-invasive, and mobile acquisition setups (including wireless headsets) that allow free movement to participants [147], [148]. Temporal resolution seems to be important for studying CL, among the other aspects of design cognition, since design is a complex cognitive activity related to multiple cognitive processes that change throughout the execution of design tasks [147], [149]. Freedom of the movement is appreciated and often required in engineering design research, especially when dealing with CAD tasks which involve HCI tools, such as a computer mouse and a keyboard.

Recent EEG studies in design aimed to describe a neurocognitive perspective of different aspects of design thinking, primarily creativity. EEG studies explored the neurocognitive differences between design tasks [150] (such as open and constrained tasks) or activities [151] (e.g. decision-making, ideation, sketching). Moreover, EEG proved to help highlight differences in design neurocognition according to the previous experience of engineers [152] (e.g. novice and experts), background [153] (e.g. mechanical engineers, industrial designers, and architects), gender (male/female), etc. At the same time, only a few EEG studies focused on generating visual design representations, CAD modelling, and engineers' interaction with virtual models and environments (e.g. [149]). Some findings may be extracted from the studies that captured engineers' EEG signals while solving design tasks; however, only a few of them segmented design tasks into segments relatable to the visual processing of information (e.g. [154]).

Previous EEG studies from other fields (primarily cognitive psychology) often investigated the visual processing of information through standardized tests (such as the mental rotation test) related to the aspects of visuospatial thinking (such as spatial visualization and mental rotation) [148], [155]–[159]. The results of these previous studies imply that the higher power in the theta and beta frequency bands reflect the cognitive processing of visuospatial information [148], [155]–[157]. For example, Liu et al. reported increased (compared to the baseline) theta and beta band power in the frontal cortical area when solving the mental rotation task [158]. In addition, an increase in theta band power over the frontal cortical area has been related to attention allocation during the task [160]. On the contrary, it has been shown that alpha band power is often reduced during visuospatial information processing and decreases with increasing processing demands [159], [160]. For instance, Riečanský and Katina suggested that the reduced alpha band power in the frontal cortical area during mental rotation task reflects enhanced attention allocation [159]. These implications may be informative for understanding design neurocognition since engineering design seems highly visuospatial in nature [11], [146], [161]. Consequently, visuospatial thinking has been recognized as an essential skill in engineering design, necessary for generating and utilizing visual design representations [11], [161], [162]. It has also been the engineering designers' characteristic whose influence on CAD performance was the most commonly investigated in previous studies (see Table 4).

Accordingly, this thesis proceeds with psychophysiological measurements and focuses on EEG as the main method of this type. This choice aligns with the hypothesis that the prescribed method for measuring and analysing CL, based on monitoring psychophysiological responses during CAD modelling, enables assessment of the engineering designer's level of CAD activity performance. The next sections therefore review EEG-based CL indicators that have been applied in cognitively controlled tasks as well as in (engineering) design tasks and summarises the features that the literature highlights as the most suitable for capturing CL dynamics in CAD.

2.4.3.1 EEG-based CL indicators[2]

The majority of previous EEG studies relied on spectral analysis to discriminate between CL levels [163], [164]. This type of analysis decomposes the EEG signal, consisting of a sum of neural oscillations, into several frequency bands to calculate their individual contributions to the spectrum across different cortical areas [165]. The contribution is quantitatively expressed through frequency band power (POW), calculated as the squared amplitude of the EEG signal within a particular frequency range [165]. The EEG spectrum is typically decomposed into five main frequency bands: delta, theta, alpha, beta, and gamma. However, the boundaries of these frequency bands are not standardised, and slight variations can be found across different publications. In this thesis, they are defined according to the guidelines of Pernet et al. [166] as follows: delta (0.5-4 Hz), theta (4-8 Hz), alpha (8-13 Hz), beta (13-30 Hz), and gamma (30-50 Hz). Additionally, alpha and beta bands are sometimes further subdivided in low and high sub-bands due to the observed functional differences between them (e.g. [163], [167]–[169]). Moreover, the spatial resolution of analysis varies across studies, typically encompassing the following levels: the entire cortex (all available channels), specific cortical areas (usually frontal, parietal, temporal, and occipital, depending on the electrode placement), and individual channels (whose number and positions depend on the EEG device used).

The scope of the overview is limited to EEG features that represent summary measures of EEG signals captured during and collapsed over the entire experimental task used in the reviewed studies. This approach, initially suggested by Gevins and Smith [170], is adopted from their earlier efforts in continuous CL monitoring in more complex, less cognitively controllable computer-based tasks, typical for real-world HCI activities.

2.4.3.1.1 EEG-based CL indicators in cognitively controlled tasks

The extraction of EEG features that may serve as CL indicators originates from cognitive psychology. Studies in this field have employed cognitively controlled experimental tasks to identify EEG-based CL indicators. The high degree of control over these tasks, and the variations in demands they impose, allowed researchers to link changes in EEG features to fluctuations in CL. The overview suggests that the N-back task has been the most frequently

[2] This section is based on:
Lukačević et al. (2025) Identifying the electroencephalography features for measuring cognitive load in computer-aided design. *Journal of Mechanical Design*, 147 (12), 121403.
doi: 10.1115/1.4068746

used cognitively controlled task for studying CL across various domains (e.g. [25], [168], [171]–[175]). Other examples of tasks occasionally used in this category include the Sternberg task (e.g. [167], [176]), the Multitasking SIMKAP task (e.g. [103]), and the Simon task (e.g. [177]). Given the prevalence of cognitively controlled tasks in CL studies, most of the commonly suggested EEG-based CL indicators have been extracted from participants' EEG signals while performing these tasks.

Two of the most prominent EEG features emerging from this set of studies are an increase in theta POW, and a decrease in alpha POW. Theta increase, often associated with higher task difficulty (and thus higher CL), was reported across the entire cortex (e.g. [176]), the frontal cortical area (e.g. [25], [168], [172], [175]), and the Fz channel (e.g. [167]). Decreases in alpha POW were observed across the entire cortex (e.g. [172], [176]), the parietal cortical area (e.g. [25], [168], [175], [176]), the occipital cortical area (e.g. [25], [168], [176]), and the Pz channel (e.g. [172], [173], [176]). In some cases, the ratios of theta to alpha POW appeared to be more indicative of CL differences, particularly when considering the entire cortex (e.g. [103], [171], [174], [175]) or the combination of parietal theta and frontal alpha (e.g. [177]). Additionally, several studies reported a decrease in beta POW across the entire cortex with increasing CL (e.g. [25], [168], [176]).

However, discrepancies and inconsistencies in the EEG features extracted from different cognitively controlled tasks, and even among studies employing the same cognitive task, have been noted. As a result, the literature does not prescribe a single EEG feature that can universally serve as a CL indicator. It remains to be explored whether any of these EEG features are sensitive to varying task complexity within the CAD context. Furthermore, the typical experimental setup for cognitively controlled tasks involves a computer with a monitor for presenting stimuli, along with a keyboard and a mouse as input devices. In such a computerised setup, these tasks can be considered as HCI tasks. Nevertheless, other aspects of cognitively controlled tasks are not easily comparable to the highly complex CAD tasks, which involve a variety of distinct cognitive processes. Given that different cognitive roles have been assigned to both cortical areas and frequency bands, it is possible that their engagement differs when solving complex computer-based tasks, such as CAD tasks, compared to simpler controlled tasks like the N-back task [170].

2.4.3.1.2 EEG-based CL indicators in CAD-related tasks

In recent years, the literature has shown that many EEG features can be descriptive of specific brain behaviours in engineering design tasks [150], [151], [178], showing the potential of this neuroimaging technology for a variety of purposes (e.g. segmenting design activities [179], distinguishing design spaces [150], or predicting design task performance [180]). The remaining part of the overview specifically addresses studies that employed EEG to explore CL in CAD-related tasks within the (engineering) design domain. These studies are summarized in Table 20, alongside the EEG features they used and identified as sensitive to changes in CL. The experimental tasks in this group of studies were more closely associated with CAD activities, making them more complex and less controlled. In addition, experimental tasks within this group of studies varied more when compared to each other. As a result, the EEG features identified as sensitive to CL in these studies (Table 20) are more diverse than those emerging from the overview of studies involving cognitively controlled tasks. Additionally, differences in the behaviour of the same EEG features between the two groups suggest that the type of experimental task indeed influences the results [130]. For instance, beta POW decreased with increasing CL in cognitively controlled tasks, whereas it increased in more complex CAD-related tasks.

Table 20 EEG-based CL indicators in (computer-aided) design or similar HCI tasks

Reference	Experimental task	EEG feature(s) sensitive to the CL changes
[126]	Designing on touchpad	Theta (Frontal) increase; Alpha (Right-Central) increase; Beta (Right-Temporal) increase
[129]	Designing on touchpad	High (20-30 Hz) Beta (Fz) increase
[130]	Computer game with screws	Theta (Frontal) increase
[132]	Solving engineering problems	B-Alert Live software metrics
[133]	Visuospatial task: origami	Theta (Fz) / Alpha (Pz) increase
[134]	HCI task: simulation of a tube rupture	(Beta+Gamma) / (Theta+Alpha) increase (all prefrontal)
[136]	Manual task: engineering (assembly)	Alpha increase
[137]	Drawing a table in AutoCAD	Theta (Frontal; Occipital) increase; Alpha decrease
[138]	Solving listing and concept mapping tasks related to sustainability issues	B-Alert Live software metrics
[141]	Designing on touchpad	Delta increase; Theta (Frontal) increase; Alpha decrease; Beta increase

Sensitivity to variations in CL across different CAD-related tasks was observed and revealed in theta, alpha, beta, and delta POW, as well as in their ratios. The literature frequently emphasizes the sensitivity of both alpha and theta to changes in CL (e.g. [126], [133], [137], [141]). However, some scholars found sensitivity in only one frequency band. For example, Schultze-

Kraft et al. highlighted the role of theta, but not alpha, for classifying CL levels in HCI tasks [130]. On the contrary, Kosch et al. identified only alpha as sensitive to CL variations [136]. Furthermore, the behaviour of EEG features was inconsistent across studies, with some reporting decreases while others observing increases. For example, while most studies noted a decrease in alpha POW, Nguyen and Zeng reported increases in alpha, theta, and beta bands [126].

However, caution is needed when interpreting these reported EEG features associated with CL sensitivity, due to certain methodological limitations. A key concern is that many studies used EEG as the default standard for distinguishing between CL levels, without critically examining or justifying the selection of specific EEG features. In some cases, researchers relied on EEG-based indicators provided by software linked to the EEG device, without clearly specifying which features were analysed (e.g. [132], [138]). Additionally, theoretical explanations for why certain EEG results corresponded to differences in CL often lacked. Consequently, it remains unclear whether the reported EEG differences truly reflected changes in CL or they were influenced by other factors. A few scholars addressed this issue by employing subjective assessment methods alongside EEG measurements. For instance, Nguyen and Zeng [129] and Choi et al. [134] validated EEG results by correlating them with NASA TLX questionnaire, which is widely regarded as a standard in measuring CL [181].

2.4.4 Conclusions of the literature review on CL

The reviewed literature provides valuable foundations for understanding CL in engineering design and its relevance for CAD modelling as an information-processing and HCI activity. Although many authors adopt a resource-based view of CL (task demands relative to limited processing capacity), definitions and operationalisations remain inconsistent, and terms such as CL, mental WL, and cognitive WL are often used interchangeably. This lack of uniformity affects both measurement choices and interpretation of results, making synthesis across studies difficult even when authors appear to address the same underlying construct.

A first key conclusion concerns how CL has been formalised and how the underlying theories can be made informative for CAD. Generic conceptual models consistently distinguish a task-centred component of CL (mental load/effective load/intrinsic load) from an avoidable component that depends on the individual and environment (mental effort/ineffective load/extraneous load). For CAD, this distinction supports separating what is inherently demanding about the CAD modelling task (e.g. geometric and feature complexity, element interactivity, coordinating multiple information streams) from what becomes demanding due to representation formats, interface characteristics, and strategy/workflow choices (e.g. unnecessary visual search, redundant actions, rework). Engineering-design-specific models further extend this structure by introducing concepts such as management load and potential load, which are relevant whenever design work involves coordinating multiple concerns and productive exploration of both design problem and design solution spaces. Cognitive models grounded in cognitive architectures then demonstrate that CL/WL can be quantified as temporally varying indices derived from module/subnetwork utilisation or activation time, and that additional factors (e.g. capacity modifiers, module weights, error penalties, decay of earlier demands) can be incorporated into the estimates. At the same time, these computational approaches have been developed mainly in tasks with clearly specified procedures and timing constraints, implying that their application to CAD is most direct when CAD modelling can be represented as recurring, well-defined CAD modelling segments and less straightforward when CAD modelling is highly exploratory unless those behaviours are explicitly encoded.

A second key conclusion concerns what empirical studies show about factors shaping CL in (engineering) design tasks. Across the reviewed work, CL varies as a function of interacting factors related to 1) the subject, 2) the task, and 3) the working environment and tools. Evidence for subject-related effects indicates that CL is not determined by task demands alone and can vary with expertise and familiarity, sometimes in counterintuitive directions, likely because expertise changes strategies, reasoning depth, and what information is actively maintained. Studies on methods and tools suggest that CL can be influenced by interaction modalities, support methods, and representation media, but findings are often context-specific and occasionally conflicting across formats. Task-focused studies provide comparatively clearer evidence that CL differs across design activity types and increases with internal task complexity, yet they still rarely resolve how CL evolves within longer, realistic processes. Importantly, only a subset of studies explicitly relates CL to performance, and the reported relationships are mixed (positive, null, negative, and potentially curvilinear). For CAD specifically, the evidence remains narrow; only a small number of studies investigated CL during CAD activities, and even fewer support temporally sensitive assessment aligned with CAD process dynamics.

A third key conclusion concerns how CL has been measured and how measurement choices constrain what can be learned about CAD modelling. Across (engineering) design studies, CL has been assessed through subjective ratings (most often NASA TLX), performance-based approaches (e.g., dual tasks, protocol analysis), and psychophysiological methods (e.g., ET, EEG, fNIRS). However, these methods differ in intrusiveness and temporal granularity. Subjective ratings are broadly applicable but typically yield task-level summaries that do not localise when CL increases during modelling. Performance-based approaches can reveal behavioural consequences of CL (e.g. time, errors, rework, strategy shifts), but they may confound CL with expertise and speed–accuracy trade-offs and can be intrusive in realistic CAD execution if temporal granularly is enlarged. Psychophysiological methods, and especially EEG, are therefore increasingly used because they can support continuous monitoring during real-world HCI, which is essential if CL is expected to be unevenly distributed across task segments. At the same time, the EEG literature shows that CL-sensitive features identified in cognitively controlled tasks do not transfer uniformly to CAD-related tasks (including differences in frequency-band behaviour and direction of effects), and CAD-specific CL evidence remains narrow; only a small number of studies measure CL during CAD activity at all, and direct CL–performance relationships in CAD have not yet been established. Together, these findings indicate that progress depends on pairing process-sensitive performance measures with temporally informative CL measures, while treating EEG indicators as candidates that require CAD-specific feasibility checks and validation.

These consolidated conclusions guide the design and purpose of the first empirical study. The literature review identified several factors that may influence CAD performance and CL; however, engineering information format was selected as the focus of the first empirical study because it is a practically relevant factor in CAD modelling, has been shown in prior work to affect design activity and performance, and is theoretically expected to influence cognitive demands. In particular, CAD performance research (Section 2.2) suggests that engineering information format can shape how engineering designers interpret and use input design information, while CL theory (Section 2.3) implies that representation format can modulate the avoidable component of CL (e.g. by affecting visual search and action selection). At the same time, empirical evidence on how different engineering information formats affect CAD modelling performance and CL during realistic CAD activities remains limited. Consequently, the first empirical study varies input engineering information format, specifically projection type and examines whether the change yields detectable differences in CAD modelling outputs and CAD modelling processes. The selected factor therefore serves as an empirically grounded and theoretically motivated case for examining CAD performance and CL, rather than as the sole focus of the thesis. In parallel, the study records EEG during CAD modelling to test feasibility and initial sensitivity of brain activity to the two projection conditions, providing an empirical check of whether psychophysiological monitoring can complement performance analysis in realistic CAD tasks.

3 THE FIRST EMPIRICAL STUDY

The third chapter presents the first experimental study, conducted as a part of the Descriptive study I to provide additional insights and thus corroborate suggestions from the literature review, primarily those informing the theoretical model of CAD performance and the method for measuring CL in CAD. More specifically, the study is used to test the suitability of the candidate approaches for describing, measuring, and relating engineering designers' performance and CL in CAD tasks, and to collect empirical evidence that informs the subsequent prescription of the theoretical models as well as the method.

First, the study informs the theoretical model of CAD performance by evaluating whether candidate output-based and process-based performance measures are sufficiently sensitive and interpretable when CAD modelling is analysed through a process lens. Second, it informs the theoretical model of CL in CAD by testing whether a theoretically motivated influencing factor can induce detectable differences that plausibly reflect changes in information-processing demands. Finally, it informs the method for measuring and analysing CL in two ways; by testing whether human performance modelling is feasible and informative in the CAD context, and by assessing the feasibility and initial sensitivity of EEG-based measurement during realistic CAD modelling, which is prerequisite for later dynamic, segment-level CL analysis.

Accordingly, the first empirical study is framed as an empirical step that bridges the literature review (Chapter 2), and the subsequent prescription of the models (Chapter 4) and the method (Chapter 5). Rather than testing the full thesis hypothesis at this stage, the first empirical study evaluates whether the building blocks required for the later prescribed method (process-oriented performance descriptors, segmentation logic and operator mapping grounded in human performance modelling, and psychophysiological CL measurement) are practically usable and empirically informative in a CAD setting.

3.1 Study motivation, objectives, and research questions

Literature review (see Chapter 2) provides three key premises for designing the first empirical study. First, CAD performance has been investigated with inconsistent definitions and heterogeneous metrics. Many studies emphasise the quality of the final CAD model, whereas fewer analyse CAD modelling as a process that unfolds over time with interpretable behavioural structure. Second, the CL review indicates that CL is widely assumed to affect performance in complex HCI tasks, yet CAD-specific evidence remains limited; CL is rarely assessed dynamically during CAD modelling; and EEG-based indicators established in cognitively controlled tasks do not necessarily transfer to CAD without dedicated validation. Third, taken together, these gaps imply a need for an empirical study that 1) varies a practically relevant factor known to influence CAD performance, and 2) simultaneously evaluates whether process-oriented human-performance modelling and continuous psychophysiological measurement (EEG) are feasible and sensitive to information-processing changes in realistic CAD modelling. Based on these premises and literature review findings, the first empirical study adopts representation format as a focused and theoretically grounded experimental variable. Prior work argued that the way information is presented in visual design representations can influence both the progress and outcomes of design activities [182]. Consequently, improving the ways in which information is utilised in design activities has been proposed as an avenue for enhancing engineering design [183]. Moreover, several studies reported significant relationships between task performance and the format of presenting input engineering information (e.g. [46], [184]). Importantly, much of this evidence compared CAD models to other design representation types (such as technical drawings and physical prototypes), rather than comparing alternative input formats commonly used to drive CAD modelling. From a CL standpoint, conceptual models distinguish between task-centred and avoidable components of CL; representation format is a theoretically plausible driver of the avoidable component because it can change visual search demands, WM load, and action-selection effort even when the underlying modelling is unchanged. Therefore, varying projection type in technical drawings provides a controlled way to test whether differences related to information processing are observable in CAD performance outcomes and processes, and whether such differences are accompanied by measurable changes in brain activity during CAD modelling.

Accordingly, the study focuses on technical drawings as a common form of design representations that visually record the technical systems' characteristics and properties used as information input to CAD modelling [8], [185]. It tests the impact of projection type used to present technical systems in technical drawings – a factor repeatedly identified as potentially influential for CAD performance [162], [184]. Two widely used projection types are examined: isometric and orthographic. Both can communicate the same design content and amount of information, but they represent information differently. Isometric projection provides a single-view depiction of all three principal dimensions from a 120° viewing orientation, with equal angular distortion of each axis (30°) [162], [186]. Orthographic projection provides multiple two-dimensional (2D) principal views (typically front, top, and side) arranged according to the first- or third- angle conventions [186]. Because these projection types differ in how design information is presented, and how it must be perceived and held in the WM during CAD modelling, differences are expected not only in outcome quality, but also in process structure and neurocognitive (EEG) responses when generating 3D CAD models from them.

On the basis of the above premises, the study formulates three research questions that operationalise and narrow the broader thesis aims to guide the data analysis and presentation of results, which are then used to directly inform the later prescription of the theoretical CAD performance and CL models (Chapter 4), and the CL measurement and analysis method (Chapter 5).

The first research question follows from the CAD performance literature (Section 2.2.2.2.1) that has frequently relied on output-based performance measures (e.g. correctness, completeness, accuracy) and from claims that representation format can affect the quality of modelling outcomes [162], [184]:

1.1 Do the CAD modelling outputs differ when using isometric and orthographic projections?

Testing output differences thus provides an initial empirical check that the projection type (representation format) is practically meaningful for CAD performance, and that selected outcome metrics can discriminate between conditions.

The literature review also identified that CAD performance has rarely been interpreted through a process lens using consistent segmentation and terminology. To enable meaningful process comparisons between the two conditions, the CAD modelling activity must first be represented in a standardised way; i.e. decomposed into task segments and organised within a hierarchical structure that is applicable across participants and experimental conditions. Establishing the segmentation framework is therefore a prerequisite for process analysis; it defines what is being compared (the same segments and task/goal levels across conditions) and provides a common basis for quantifying differences in temporal structure and embodied cognition operators demands (P, C, and M). Once this consistent decomposition is in place, the study can examine whether and how CAD modelling processes differ between projection types:

1.2 Do the CAD modelling processes differ when using isometric and orthographic projections?

Testing process differences thereby evaluates the sensitivity and interpretability of the process descriptors, which is an essential prerequisite for prescribing the process-based layer of the CAD performance model (Chapter 4) and for later aligning CL measures with CAD task segments (Chapter 5). Moreover, although the CAD tasks used in the experiment are relatively simple in CAD terms, they are still substantially more complex than the tasks typically modelled in human performance modelling (such as MHP and KLM modelling). This therefore motivated testing whether these modelling approaches remain feasible and informative in a CAD context.

From the CL perspective, the literature review indicates that CL is expected to vary with design representation and HCI characteristics, yet CAD-specific CL evidence remains limited, and dynamic CL assessment during CAD modelling is rare. This gap is especially important because CAD modelling is inherently visual and virtual; engineering designers must interpret rich visual stimuli on-screen, maintain and transform design information in WM, and act through HCI tools while continuously integrating system feedback [146], [187]. Such characteristics place the neurocognitive perspective underlying the utilisation of design information at the centre of efforts to improve design representations and computer support in engineering design [183], [188]. Existing design studies have primarily investigated representation, media, and tool effects using protocol analysis and behavioural methods (e.g. [147], [182], [189], [190]) but comparatively little is known about the underlying cognitive information processing (and neurocognition in particular) during realistic CAD activities, partly due to the lack of reliable methods for continuous monitoring in ecologically valid tasks [188], [191].

Electroencephalography (EEG) has therefore been proposed as a promising method for studying the neurocognitive perspective of design cognition, primarily because it is non-invasive and enables continuous (time-resolved) monitoring of brain activity during ongoing task performance [147]. However, before EEG can be integrated into a CL measurement and analysis method for CAD, its feasibility and sensitivity must be demonstrated in realistic CAD modelling conditions. Accordingly, in this first empirical study EEG is not used to claim a definitive CL indicator; rather, it serves as an initial empirical check that this psychophysiological measurement can be successfully applied during CAD modelling and can detect differences under two controlled conditions related to the projection type. In this context, the study examines whether changing the projection type in the input technical drawings is accompanied by measurable differences in brain activity during CAD modelling:

1.3 Is the brain activity of mechanical engineers different when CAD modelling from isometric and orthographic projections in technical drawings?

In the context of the overall thesis objectives, this step informs the later prescription of the CL measurement and analysis method (Chapter 5) and supports the selection of candidate EEG features for deeper validation in the second empirical study (Chapter 6).

3.2 Methodology

The next sections describe the methodology through which the research was conducted to answer the research questions.

3.2.1 Participants

The study recruited 20 mechanical engineers (young professionals) to participate in the experiment. The inclusion criteria were: being a mechanical engineer, being experienced with CAD modelling in SolidWorks, being a right-handed, and an absence of neurological disorders. Participants were intentionally recruited as professionals experienced in CAD to ensure that task performance reflects skilled CAD system use rather than effects of learning CAD software, and to reduce variability due to novice exploration. This choice also aligns with human performance modelling assumptions (related to MHP and KLM/GOMS), which are most applicable for skilled users that can execute learned methods fluently [86], allowing observed differences on both performance measures and EEG to be attributed more directly to the experimental conditions rather than to differences in basic CAD proficiency. In addition, the participants were instructed to refrain from coffee and caffeine beverages at least two hours before the experiment to mitigate their potential effects on brain activity captured by EEG signals.

The participants' age, professional engineering experience, CAD experience, and technical drawing experience are summarised in Table 21 using their median (Med), Median Absolute Deviation (MAD), and range. In addition, all the participants finished the same engineering graphics course as a part of their studies. Furthermore, Figure 8 presents the participants' frequency of CAD modelling and creating or reviewing technical documentation on the scale from never to always (every day).

Table 21 ES#1: Demographics, education, and prior experience

	Med	MAD	Range
Age [years]	27.50	1.34	25.00 – 30.00
Professional engineering experience [months]	21.50	16.84	0.00 – 72.00
CAD modelling in SolidWorks [months]	8.00	10.97	0.20 – 120.00
CAD modelling [% of working hours]	10.00	14.83	0.00 – 70.00
Creating/reviewing technical documentation with orthographic projections [% of working hours]	5.00	7.34	0.00 – 50.00
Creating/reviewing technical documentation with isometric projection [% of working hours]	1.00	1.48	0.00 – 35.00

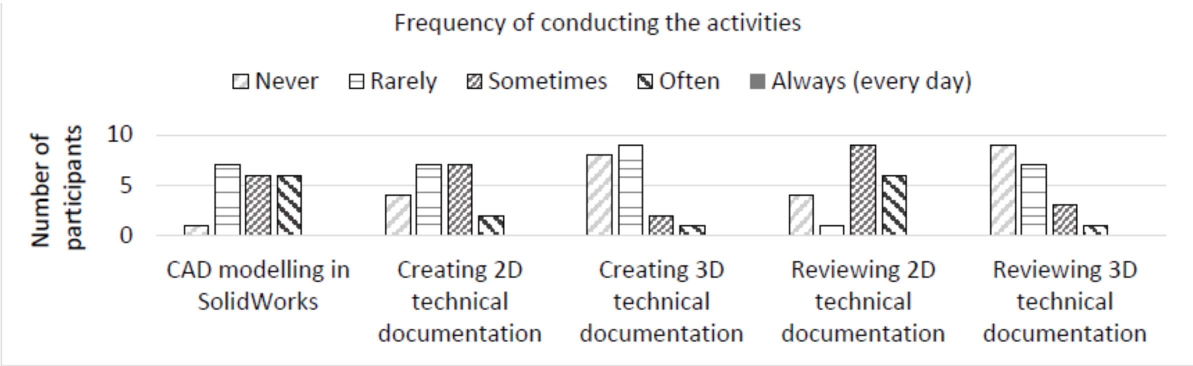

Figure 8 ES#1: Participants' frequency of CAD modelling and creating or reviewing technical documentation

3.2.2 Experimental procedure

The experimental procedure consisted of 15 steps, shown in Figure 9. Firstly, the participants were introduced to the equipment and the experimental procedure. Next, participants were asked to sign a consent. The informed consent was obtained from all the participants. In the third step, the EEG headset was set up. The participants continued to the CAD tasks when the contact and EEG data quality were satisfactory (according to the metrics defined by Emotiv and indicated with green colour within the data gathering application).

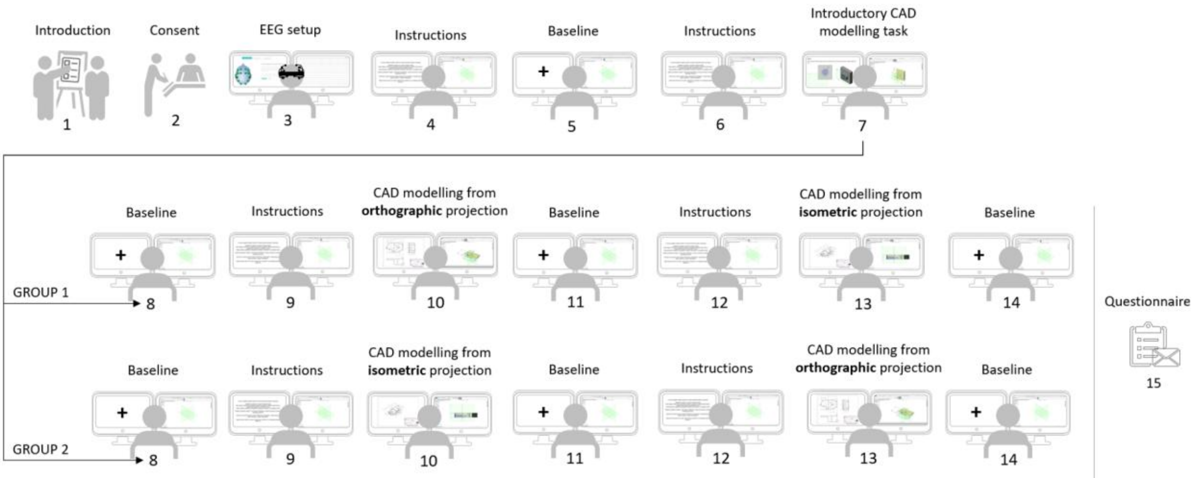

Figure 9 ES#1: Experimental procedure

The parts and the CAD tasks were the same for all the participants. The CAD tasks were not time limited. Each participant was asked to generate 3D CAD models of three parts in total. All the participants started with an introductory CAD task (see Figure 10) that served as a warm-up task for familiarisation with the interaction devices and CAD environment. After the introductory CAD task, participants were instructed to generate 3D CAD models of two parts based on their technical drawings. Each participant had one trial in each condition (orthographic or isometric projection). However, the order of the conditions was controlled; half of the participants (group 1) first generated a 3D CAD model of part 1 (from its isometric projection) and continued to part 2 (from its orthographic projection). The order was reversed for the other half of the participants (group 2). The randomised division was motivated by the goal of bypassing the potential bias of the previous task and cognitive fatigue as its consequence. Each CAD task was preceded and followed by a baseline task; this is a typical procedure in EEG studies, and it is required for normalisation purposes when data analysis is based on comparing POW values between the conditions (see Section 3.5 for details). For the baseline task, participants were asked to stare at the fixation cross on the monitor display until it disappeared (after 20 seconds). The slides with the instructions, the fixation cross, and technical drawings (presented on the left screen) advanced based on participants' keyboard input. Thus, their timing was based on the duration of solving each step. The duration of the active part of the experiment (from step 4 to the end of step 14) ranged from 26.26 to 68.84 minutes, with Med = 45.07 minutes and MAD = 18.48 minutes. There were no breaks during the experiment. Finally, the questionnaire on demographics and prior-experiment experience related to CAD modelling and technical documentation was sent via e-mail to the participants after the experiment. The Ethics Committee of the Faculty of Mechanical Engineering and Naval Architecture, University of Zagreb approved the described experimental protocol.

3.2.3 CAD modelling tasks

During the CAD introductory task, participants were guided through seven CAD modelling steps to create a 3D CAD model of a simple part (presented in Figure 10). This task consisted of seven main steps; creation of three sketches and usage of four CAD features. The resulting model contained a cuboid, a cylinder, a through hole, and chamfers. The introductory task served as a warm-up task to familiarize the participants with the interaction devices and CAD environment.

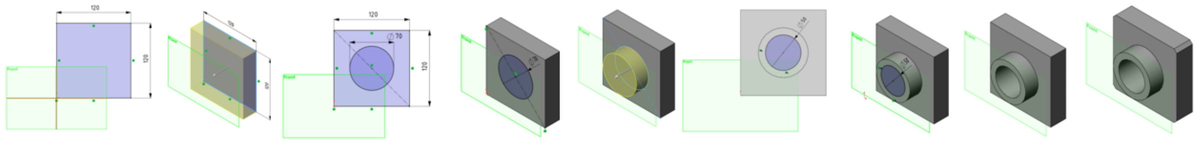

Figure 10 ES#1: Steps of the introductory CAD modelling task

The study incorporated two CAD tasks in which participants were asked to generate 3D CAD models of mechanical parts based on their technical drawings. The study focused on a part modelling as it is considered one of the fundamental aspects of using CAD systems, and thus serves as the basis for most of the more complex CAD activities [58]. This choice also aligns the study with the practical requirements of human performance modelling, where tasks are typically short, well specified, and comparatively simple (e.g. routine HCI procedures with stable method knowledge) [86]. Therefore, the experimental CAD tasks were intentionally kept relatively simple in CAD terms, while still representative of fundamental part-modelling practice. Task complexity was controlled to be comparable across conditions by constraining the resulting models to the same set of features [67]: a cuboid, fillet, chamfer, through hole, slot, and three through slots (Figure 11). Participants were free to choose CAD modelling strategy, feature types, and procedure, without restrictions.

In one CAD task, the technical system (part 1) was presented with the single-view isometric projection in the technical drawing (condition #1). In the other CAD task, the orthographic projection (condition #2) with three main views (front, top, right) in the first angle was used as a 2D visual representation (part 2). Hence, the projection type used to present the parts in technical drawings was an independent variable. In both cases, technical drawings were mediated by the monitor screen as the 2D interface.

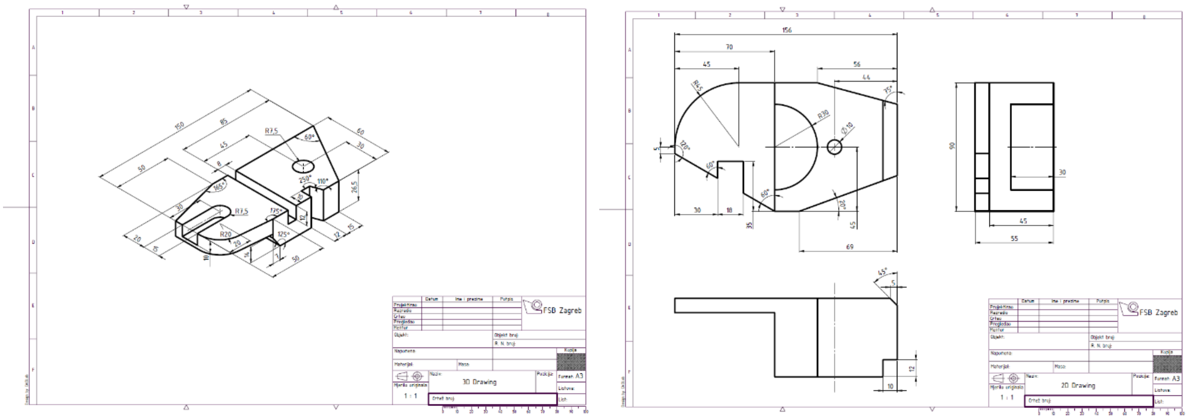

Figure 11 ES#1: Isometric projection of the part 1 (left) and orthographic projection of the part 2 (right)

3.2.4 Experimental setup

The experiment was conducted using one high-performance computer, two 23.8" monitor screens (resolution of 1920 x 1080 pixels, refresh rate of 60 Hz), and a keyboard and a mouse as the interaction devices. The technical drawings and instructions with detailed explanations of what should be done in each task and step were presented through the PsychoPy [192] application on the left monitor. The CAD modelling was conducted in SolidWorks software, presented on the right monitor. As shown in Figure 12, both screens were recorded for the entire experiment duration.

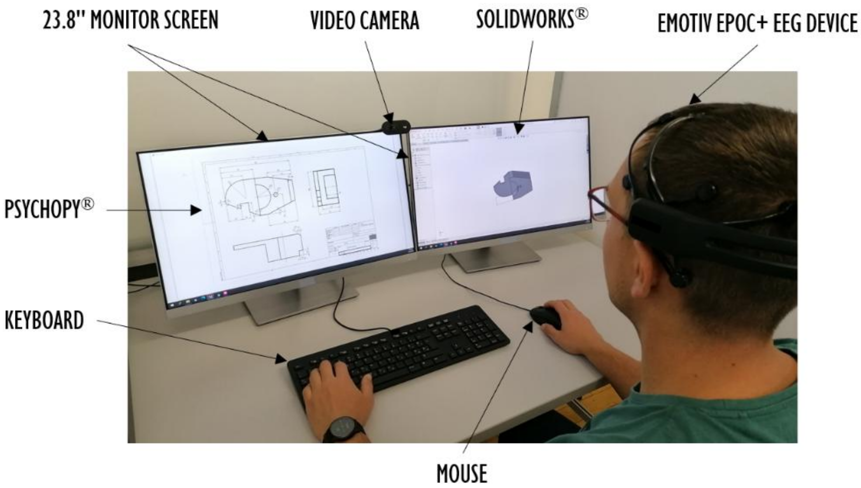

Figure 12 ES#1: Experimental setup

Screen and video recording enabled capturing behavioural data; experiment progress from the left screen and CAD modelling from the right screen. The video camera captured the participant's face, which served the purpose of providing additional explanations of EEG and behavioural data related to CAD if necessary. The PsychoPy enabled the synchronisation of behavioural and EEG data. The device used for EEG data gathering during the CAD tasks was Emotiv EPOC+ with 14 electrodes (AF3, F7, F3, FC5, T7, P7, O1, O2, P8, T8, FC6, F4, F8, AF4) and the integrated amplifier. The sensors' locations followed the international 10-20 system. Two reference sensors were at P3 and P4 locations. These locations were used as the reference in the later data pre-processing steps. Continuously captured EEG data was sent via wireless connection (Bluetooth Low Energy) to a high-performance computer. The sampling frequency used to collect EEG data was 128 Hz. According to the Nyquist-Shannon theorem, such sampling frequency is adequate for analysing frequency bands relevant to the study (see Section 3.5), according to the literature reviewed in Section 2.4.3: theta (4-7 Hz), alpha (8-12 Hz), and beta (13-30 Hz) [193].

3.2.5 Collected experimental data

Four types of data were collected in the experimental study, they are presented in Table 22.

Table 22 ES#1: Description of collected experimental data

Data type	**Data description**	**Main collection purpose**
Questionnaire	Participants' responses	Demographics, experience, expertise
CAD modelling outputs	CAD models, CAD log files	CAD performance assessment Answer RQ #1.1
CAD modelling process	Screen recordings	CAD performance assessment Cognitive model of CAD performance Answer RQ #1.2
EEG data	14 channels, 128 Hz	Test the capability to investigate design cognition in the CAD context. Answer RQ #1.3

3.3 Differences in CAD outcomes and processes when modelling from isometric and orthographic projections[3]

The initial analysis builds on the previous work on CAD performance by adopting the metrics used in the literature to describe and compare CAD outputs and processes (Section 2.2.2). These metrics, initially intended for educational purposes, are here adapted to experimental tasks executed by the practitioners to compare their CAD performance in two conditions that haven't yet been explored – CAD modelling from orthographic and isometric projections. Literature review in Chapter 2 concluded that CAD performance is most consistently interpreted through two complementary aspects, reflected in the quality of the resulting CAD model (outputs), and in how the model is produced (process). It also highlighted a practical challenge; although many outcome- and process-oriented metrics exist, only a subset is both extractable reliably in a realistic CAD setup and comparable across participants when the CAD modelling goal is fixed. Guided by these conclusions, this initial analysis within the first empirical study operationalises CAD performance using a deliberately small set of metrics that align with the output/process distinction in Section 2.2.2.1, and are widely used or directly supported by CAD systems or existing auto-assessment tools. By comparing this initial set of metrics between the conditions, this part of the analysis helps in answering the following research questions:

1.1 Do the CAD modelling outputs differ when using isometric and orthographic projections?

1.2 Do the CAD modelling processes differ when using isometric and orthographic projections?

3.3.1 Data analysis

CAD data for comparing CAD outputs and processes was extracted from two main sources: Graderworks [72] (shape, volume, surface area, duration) and an additional application for CAD performance analysis (feature- and sketch-based data for comparing CAD modelling processes).

CAD modelling outcomes were compared based on completeness [44], which literature review identified as the most common and most directly operationalizable output quality metric (see Section 2.2.2.1). A CAD model is considered complete if it accurately replicates the size and shape of the component [44]. Specifically, completeness is quantified via geometric similarity

[3] This section is based on:
Lukačević et al. (2024) Engineering designer' CAD performance when modelling from isometric and orthographic projections. *Proceedings of the Design Society*, Volume 4: DESIGN 2024, pp. 653-662. doi: 10.1017/pds.2024.68

measures (volume, surface area, and shape) comparing each participant's model to a gold-standard reference (created by the experiment designer), because such reference-based metrics provide transparent interpretation and avoid subjective grading. The volume and surface area of participants' models were extracted from SolidWorks using Graderworks and then compared with the "gold standard" model, according to Equations (28) and (29). In that way, these two metrics reflect percentage (typically 0-100) by which two models differ.

$$Volume = \frac{Volume_{\text{Participant}} - Volume_{\text{Gold Model}}}{Volume_{\text{Gold Model}}} \tag{28}$$

$$Surface\ area = \frac{Surface\ area_{\text{Participant}} - Surface\ area_{\text{Gold Model}}}{Surface\ area_{\text{Gold Model}}} \tag{29}$$

Furthermore, the shape comparison was extracted directly from Graderworks, where it is performed using the D1 geometric similarity algorithm [72]. The algorithm is detailed by Renu and Mocko [194], and Cardone et al. [195]. Although broader rubrics of CAD model quality exist (e.g. conciseness, simplicity, design intent), Section 2.2.2.1 also noted that these dimensions are hard to quantify reliably without extensive manual assessment or task-specific thresholds. Therefore, the first empirical study prioritises completeness-related metrics that can be extracted consistently and used to test whether representation format yields measurable differences in outputs.

CAD modelling processes were described using two complementary groups of metrics. First, time-based metrics operationalise efficiency in a way that is broadly comparable across participants; total time (entire task duration including interpretation and modelling) was extracted from screen recordings), and edit time (time spent creating and manipulating a CAD model in SolidWorks) was extracted using the additional application. Second, feature- and sketch-based metrics captured process structure using data available from the feature tree and the properties of sketches (according to Section 2.2.2.2); they quantified number and type of sketch entities, number and type of sketch relations, and number and type of CAD features the participants used. This first step of the analysis within this empirical study thus uses cumulative process descriptors as an initial check of whether the projection types produce interpretable changes in CAD modelling behaviour, which later motivates the more fine-grained segmentation and dynamic analyses in the subsequent section.

These extracted CAD data were then analysed using the R language. Data analysis included both descriptive and inferential statistics. Descriptive statistics encompassed Med as a measure of central tendency and MAD as a measure of variability. Furthermore, inferential statistics tested differences between the isometric and orthographic conditions in CAD modelling outcomes and processes by comparing several metrics described in the following sections. The paired t-test was used for comparisons of variables when the assumption of normality (tested by Shapiro-Wilk test; $p < 0.05$) and equity of variances (tested by Levene test; $p < 0.05$) were met. If the normality assumption was violated, the Wilcoxon signed-rank test on paired samples was used instead. In addition, the effect size of the tested differences was calculated with Cohen's d.

3.3.2 Results

The following sections describe differences in the CAD modelling outcomes and processes between the two conditions. These differences in the following sections are presented numerically in the tables and graphically in the box plots. The p-values and the related effect sizes are, in the following tables, coupled with the test statistic values: t for the t-test and V for the Wilcoxon signed-rank test.

3.3.2.1 CAD modelling outcome

Comparison of CAD models created by the participants with the "gold standard" model showed smaller differences across all three similarity metrics (volume, surface area, and shape) when CAD modelling from the orthographic projection (i.e. values closer to zero). Differences in the percentage of correctly replicated volume and shape were statistically significant, as presented in Table 23 and Figure 13. These results therefore indicate higher completeness (i.e. better CAD modelling outcomes) when using the orthographic projection. Note that the sign of the difference does not imply a better or worse models; it is determined solely by whether the participant value is in Equations (28) and (29) subtracted from the gold standard or vice versa.

Table 23 CAD modelling outcomes: Completeness

Projection	Metric	Med	MAD	V or t; p	d
Isometric	Volume [%]	-0.87	0.88	38; $1.30 \cdot 10^{-2}$	0.16
Orthographic		$-4.43 \cdot 10^{-2}$	$5.29 \cdot 10^{-2}$		
Isometric	Surface area [%]	-0.15	0.20	67; 0.16	0.12
Orthographic		$4.13 \cdot 10^{-2}$	$7.37 \cdot 10^{-2}$		
Isometric	Shape	-3.63	2.37	32; $4.86 \cdot 10^{-3}$	0.60
Orthographic		-0.29	0.43		

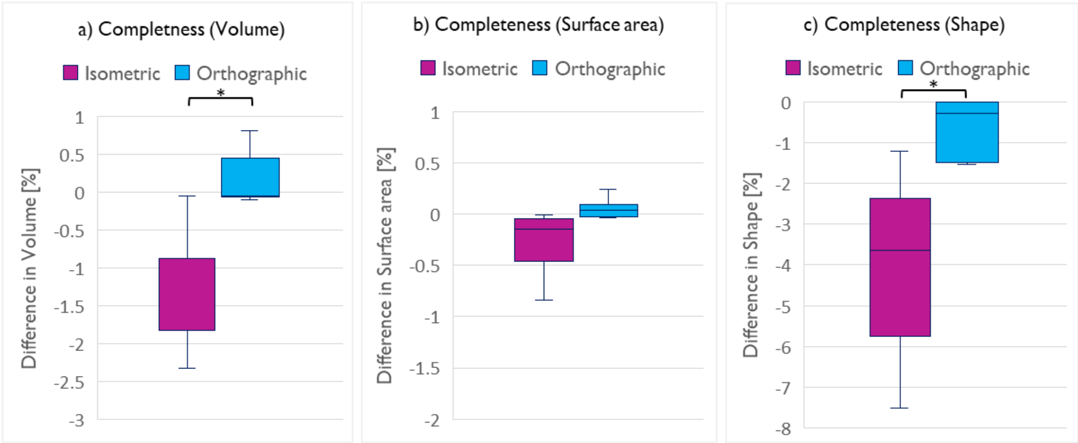

Figure 13 ES#1: Completeness of CAD modelling outcomes

3.3.2.2 CAD modelling process

3.3.2.2.1 Duration

The duration of the entire CAD task performance (total time) was longer when CAD modelling from the isometric projection, as presented in Table 24 and Figure 14 (panel a). If comparing only the duration of CAD modelling, the average edit time was the same for both projections (see panel b in Figure 14). However, these differences were not statistically significant.

Table 24 CAD modelling process: Duration

Projection	Metric	Med	MAD	V or t; p	d
Isometric	Total time [min]	14.77	6.00	-1.79; $9.20 \cdot 10^{-2}$	0.42
Orthographic		16.52	6.75		
Isometric	Edit time [min]	13.50	5.19	92.50; 0.78	$6.00 \cdot 10^{-2}$
Orthographic		13.50	6.67		

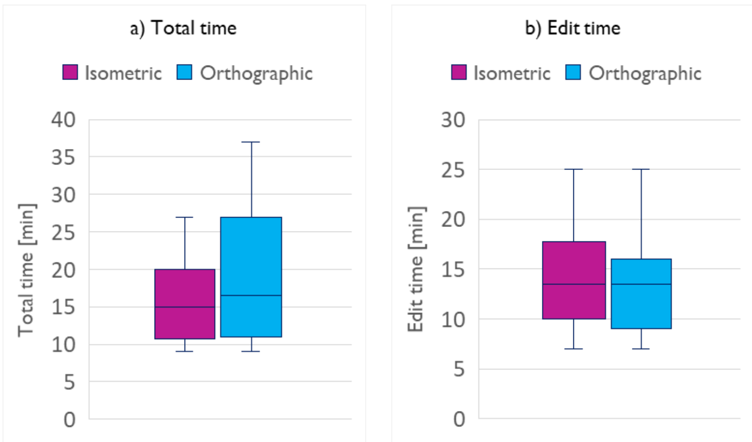

Figure 14 ES#1: Total time (panel a) and edit time (panel b)

3.3.2.2.2 Sketch entities

The average number of entities per sketch and the total number of sketch entities were slightly higher when CAD modelling from the orthographic projection, as presented in Table 25 and Figure 15. The most used sketch entity was *Line* for all participants in both conditions.

Table 25 CAD modelling process: Sketch entities

Projection	Metric	Med	MAD	V or t; p	d
Isometric	Average	4.92	1.36	72; 0.37	$1.84 \cdot 10^{-2}$
Orthographic		5.08	1.36		
Isometric	Total	25.00	2.97	83; 0.64	$8.15 \cdot 10^{-2}$
Orthographic		27.50	5.19		

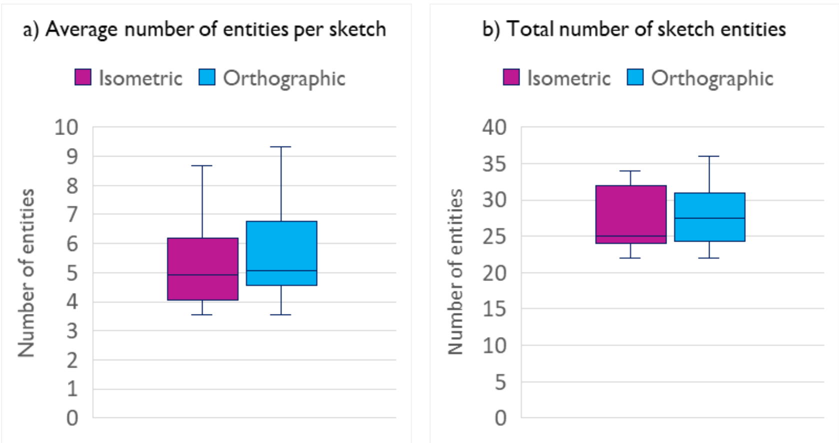

Figure 15 Sketch entities: a) Average number; b) Total number

3.3.2.2.3 Sketch relations

As presented in the panels and Table 26, the average and the total number of relations were slightly lower when CAD modelling the geometry from the orthographic projection. The most used relation was *Coincident* in both conditions (see panel c in Figure 16).

Table 26 CAD modelling process: Sketch relations

Projection	Metric	Med	MAD	V or t; p	d
Isometric	Average	10.50	2.22	59.50; 0.68	0.21
Orthographic		10.00	1.48		
Isometric	Total	57.50	12.60	1.02; 0.32	0.23
Orthographic		56.50	12.60		

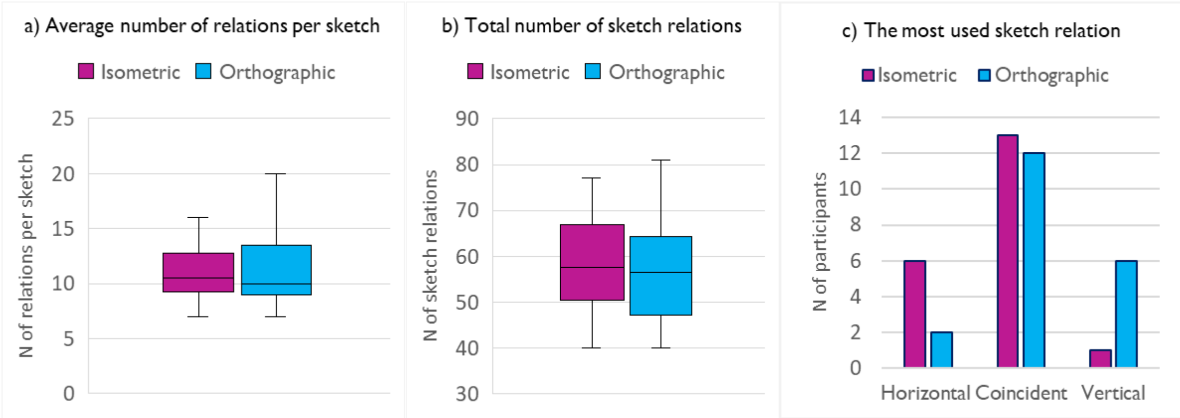

Figure 16 Sketch relations: a) Average number; b) Total number; c) The most used relations

3.3.2.2.4 CAD features

The total number of CAD features was slightly higher when generating the geometry from the orthographic than the isometric projection, as presented in Table 27 and Figure 17 (panel a). On average, the most used CAD feature was *Cut Extrude* in both conditions.

Table 27 CAD modelling process: CAD features

Projection	Metric	Med	MAD	V or t; p	d
Isometric	N of features	5.50	2.22	-0.49; 0.63	0.11
Orthographic		6.00	1.48		

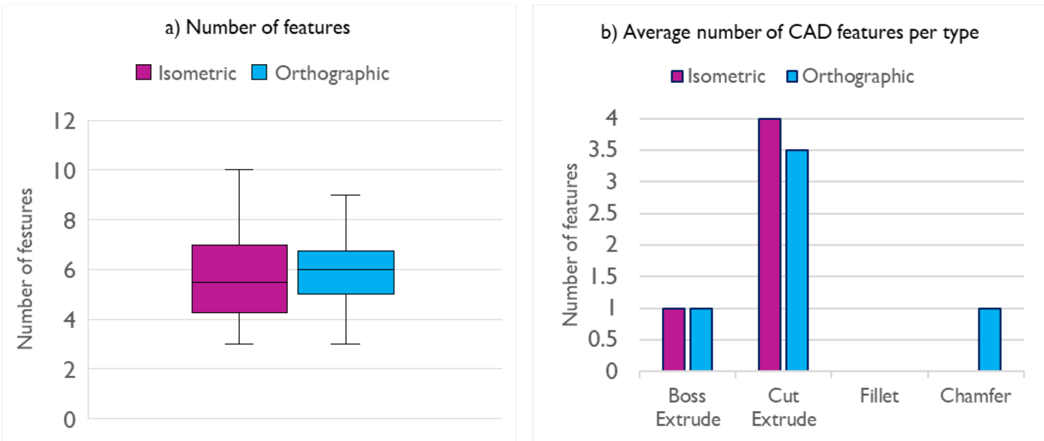

Figure 17 CAD features: a) Total number; b) Distribution per type

3.3.3 Discussion

The analysis showed that participants produced better (completer and more accurate) 3D models when working from orthographic projections, suggesting that they were able to better interpret and reconstruct the geometry under this condition. This result positively answers the first research question and may reflect participants' greater familiarity and experience with orthographic projections, as reported in the questionnaire (see Table 21). Another explanation could be that orthographic projections are inherently more convenient for geometry replication in CAD systems because traditional CAD modelling relies on 2D sketches whose planes typically align with orthographic views.

Interestingly, although the overall CAD modelling time was longer in the orthographic condition, the edit time remained similar across both projection types. This suggests that the additional time was likely spent interpreting the drawings, and potentially facilitating a deeper understanding of the geometry and CAD modelling strategy, which in return led to completer outputs and better CAD outcomes. These temporal differences offer important insights into how engineering designers engage with different types of input representations, but they also highlight the limitations of relying solely on output-based metrics such as completeness or total CAD modelling time.

To address the second research question, the study employed process metrics extracted from SolidWorks, thus following a common practice in prior research (e.g. [12]). However, these metrics did not reveal statistically significant differences in the CAD modelling processes between the two conditions. This negative finding may be partly explained by the components' equivalent complexity levels, which were intentionally matched in terms of geometric features. While this design choice ensured experimental control, it also constrained the emergence of potentially divergent CAD modelling behaviours.

More importantly, this result underscores a critical limitation in current approaches to describing and comparing CAD modelling processes; existing metrics are often too output-focused to capture meaningful differences in user performance, especially related to strategies, decision-making, or cognitive engagement. Despite subtle behavioural variations observed in sketch entities and relations, and CAD features, these differences are not adequately reflected in conventional metrics, thus pointing to the need for more fine-grained, temporally sensitive, and cognitively informed indicators. While CAD model completeness was the primary outcome metric in this analysis, aligned with the only requirement stated in the task instructions, the findings point to a broader need for more process-oriented metrics to describe and compare CAD modelling processes.

The subsequent sections address these challenges through structured task analysis and cognitive modelling, paving the way for a more comprehensive framework to describe and assess CAD performance across users, tasks, and levels of complexity.

3.4 CAD modelling process segmentation and representation

The outcome-based analyses in Section 3.3 established whether orthographic and isometric projections yield detectable differences in CAD modelling effectiveness, measured through completeness of the final CAD models. However, the literature review showed that such outputs are only a partial account of CAD performance; they reveal what was achieved, but not how it was achieved, how information-processing resources were allocated over time, or where difficulties and rework emerged. This limitation is particularly important for RQ 1.2, which concerns process differences in CAD modelling tasks between representation formats (orthographic and isometric projection in technical drawings). Prior CAD performance research remains methodologically fragmented in this respect; segmentation granularity varies widely, terms such as commands, events, actions, and operations are often used interchangeably across studies, and many of them rely on cumulative descriptors available in CAD software (from geometry and feature tree) rather than temporal process representations (see Section 2.2). Consequently, without an explicitly defined and transferable process structure, process-oriented findings are difficult to replicate, compare, and integrate across participants, tasks, and conditions.

These gaps motivate the methodological step taken in this section; before any between-condition comparison, the CAD modelling process must first be expressed in a common, observable, and reproducible representation that supports consistent segmentation and interpretation. The thesis adopts human performance modelling as the main framework for this representation. As concluded in Section 2.3.4, the Model Human Processor (MHP) provides a pragmatic baseline for CAD because it offers explicit constructs for linking observable interaction to resource-limited information processing without requiring cognitive specifications that cannot be validated for complex, variable CAD tasks [7], [92]. Complementarily, KLM/GOMS-derived task analysis provides a practical basis for representing goals, methods, and operator sequences at a controllable level of granularity, while also making transparent the assumptions required when modelling skilled, goal-directed HCI in complex tasks [7], [92]. Together, these foundations justify a process lens in which CAD performance is described across hierarchical task/goal levels and interpreted through perceptual, cognitive, and motor resource engagement, using explicit mapping rules.

Accordingly, the purpose of the present analysis is to first construct a task/process hierarchy grounded in what is directly observable in HCI between an engineering designer and the CAD system, and then use that hierarchy as a stable basis for comparison across participants and task conditions (projection format). This is necessary for two related reasons. First, process differences may not manifest in final geometry; they may instead appear as differences in sequencing, repetition, corrective actions, and redistribution of time and effort across CAD modelling segments at different levels of abstraction. Second, establishing a process representation aligned with MHP supports the broader thesis objective of later relating performance to CL; the same segmentation that makes “what happens when” explicit is required to later align CAD modelling segments with CL indicators (first operator-based and, in later stages, EEG-based indicators).

To achieve a comparable process representation from screen recordings, the study implements a structured procedure that makes segmentation choices explicit and reproducible:

1) Manual extraction of observable CAD events from the screen recordings, as time-stamped interaction logs were not available from the experimental dataset;
2) Bottom-up organisation of CAD events stream into process layers that reflect increasing abstraction and goal structure (CAD actions, CAD operations, unit task operations), thereby addressing the lack of standardised segmentation in prior process-oriented CAD studies;
3) Introduction of recurring unit task phases (acquisition, execution, inspection) as an orthogonal information-processing lens that supports consistent interpretation of *why* and *when* perceptual and cognitive processing is expected to dominate within a CAD modelling process, beyond what can be inferred from observable CAD events alone;
4) Explicit mapping of the time-aligned behavioural stream (CAD actions, CAD operations, unit task operations) to perceptual (P), cognitive (C), motor (M), and CAD system response (R) operators using documented assignment rules, consistent with MHP and KLM practices [7], [92], so that resource engagement can be visualised and quantified in a way that is interpretable and transferable across participants, tasks, and conditions.

Within this section, process-based metrics therefore refer to quantitative descriptors derived from coded, time-aligned scripts, such as time distributions across task segments, frequencies and distributions of CAD actions/operations, and estimated utilisations of P, C, and M operators mapped to the extracted task segments. This choice follows from the literature review conclusions (Section 2.2); time- and segment-based descriptors are among the most common quantitative process measures in CAD studies, yet comparability depends on having a transparent repeatable segmentation scheme, and interpreting those descriptors in information-processing terms requires an explicit operator-based mapping grounded in human performance modelling (e.g. [7], [92]).

This shift from outcome-based evaluation (Section 3.3) to CAD process modelling serves a specific purpose within Descriptive Study I. Output metrics can show whether the condition (projection format) affects what engineering designers ultimately produce, but they cannot reveal how engineering designers arrive there; whether differences arise through alternative sequences, additional navigation and view manipulation, repeated adjustments, or rework within particular segments of the CAD modelling process. The process representation developed here is therefore introduced to enable a structured, hierarchical comparison of CAD modelling behaviour across conditions and to directly address the research question:

1.2 Does the CAD modelling process differ ***(at any hierarchical level)*** *when using isometric and orthographic projections?*

Beyond this empirical study, the same hierarchy and mapping rules also provide the methodological groundwork for the thesis' later stages, where process segments must be aligned with information-processing demands to support analyses of relationships between CL and CAD performance, and prescription of CAD performance model. Accordingly, the next subsection describes how CAD events were extracted from screen recordings and how the task hierarchy and operator timelines were constructed in a consistent, transferable manner.

3.4.1 Coding the scripts

Video screen captures of CAD modelling sessions conducted by three participants in two CAD tasks were manually analysed and coded using ELAN annotation tool [196] to describe CAD modelling behaviour and CAD task structure. The three participants were selected to span the observed range of performance in both process-based terms (CAD modelling duration) and outcome-based terms (CAD model completeness). Specifically, they represent cases from the shortest to the longest CAD modelling sessions, and from the least to the most complete CAD models, as quantified by the outcome-based completeness metrics (see Table 28). In this way, their CAD modelling processes and outcomes represent low, average, and high performance scenarios, similarly to Card et al. [7] and Atman et al. [33]. Moreover, a saturation point in identifying and defining the task structure (described below) was reached through the analysis of these three participants; consequently, no additional participants were included at this data analysis stage. Previous work has employed a similar number of participants for the analysis for comparably detailed coding analysis (e.g. Lang et al. recruited four participants [97]).

Table 28 Data of three participants whose video screen captures were manually coded

Participant	Duration [min]		Completeness [%]					
	Ortho	Iso	Ortho			Iso		
			Volume	SA	Shape	Volume	SA	Shape
High performer	10.88	11.72	99.96	99.91	99.93	99.13	99.85	96.12
Average performer	28.67	19.55	99.94	99.96	100	99.13	99.85	98.58
Low performer	15.42	9.10	85.39	96.80	85.22	90.75	95.46	81.01

After selecting these three representative cases (Table 28), the subsequent analysis was organised to make explicit how an interpretable CAD task hierarchy emerges from what is directly observable in interactions between engineering designers and the CAD system (i.e. CAD events). Importantly, this bottom-up structuring was carried out for each participant separately and for each of the two CAD modelling tasks (orthographic and isometric projection condition). The hierarchy was first derived at the lowest level from each participant's CAD event stream and then progressively built upward; at each step, the resulting categories and definitions were checked for consistency across the three participants and both tasks before proceeding to the next level. This procedure ensured that the final hierarchical decomposition was internally coherent (each higher level was grounded in lower-level observations) and transferable across participants and task conditions.

Consistent with prior work (e.g. [28], [29], [97], [98]), the coding logic followed a bottom-up progression from HCI primitives to higher-level operations:

1) Elementary interactions inside the CAD system were first captured as CAD events, resulting from engineering designers' CAD actions (e.g. cursor movement, clicks, typing; Table 29). This was done separately for each of the three participants in each of the two tasks.
2) Functionally connected sequences of CAD actions were then grouped into CAD operations (e.g. drawing, dimensioning, manipulating view; Table 30), again within each participant and task. The CAD operation definitions were refined iteratively to ensure that the same event patterns were assigned to the same operation types across participants and tasks.
3) Recurrent sequences of CAD operations were subsequently grouped into unit task operations, corresponding to cognitively manageable chunks in feature-based CAD modelling (e.g. sketching a specific part of geometry, applying a CAD feature such as extrude or revolve).

4) Finally, the three process layers (CAD actions, CAD operations, unit task operations) were linked to a coherent hierarchy of goals/subgoals and used as the basis for coding the full scripts in ELAN annotation tool [196].

After the task structure was established, the coded scripts were additionally interpreted through embodied cognition lens following the MHP and KLM practices [7], [92]; by mapping perceptual, cognitive, motor, and CAD system response operators (P, C, M, R) to the time-aligned behavioural stream using the explicit assignment rules (Section 3.4.1.3).

3.4.1.1 Task structure

The task analysis aimed to derive an interpretable structure of CAD modelling behaviour from what is directly observable in interactions between engineering designers and the CAD system. Accordingly, the observable behaviour was first represented in as stream of CAD events; this coded script was then organised bottom-up into three process analysis layers and corresponding goal/task levels that reflect increasing abstraction of engineering designer's behaviour in CAD modelling; they are visualised in Figure 18 and described in the following sections.

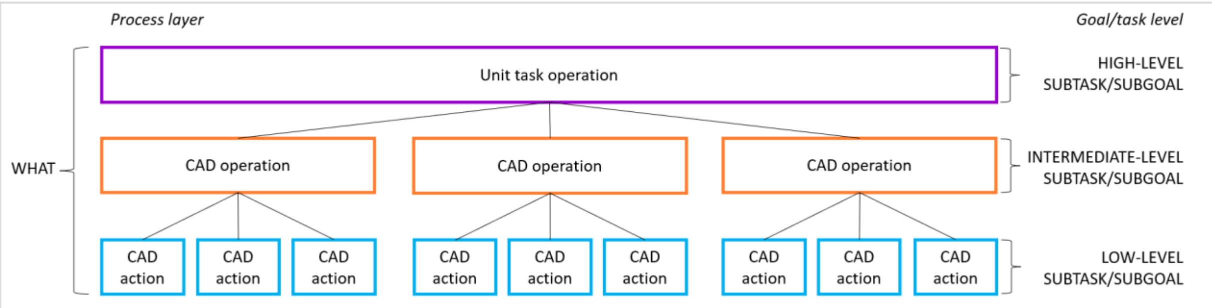

Figure 18 Process layers

3.4.1.1.1 CAD events

The screen recordings were manually analysed to organise engineering designers' behaviour as a sequence of observable CAD events. These events represent repeated low-level interactions between an engineering designer and the CAD system. The coding distinguished: 1) CAD actions executed using HCI tools, 2) periods without any CAD action (pauses), and 3) periods where the CAD system processed input (making an engineering designer to wait).

3.4.1.1.2 CAD action layer

Engineering designers' performance during the CAD modelling tasks was first organised at the lowest level in terms of interaction with the CAD system using HIC tools, i.e. CAD actions. These events represent recurrent types of interactions involving mouse-based (e.g. cursor movement, selection, dragging) and keyboard-based (e.g. typing values or entering commands) inputs. The criteria for marking each CAD action was therefore based on physical input (engineering designer's action) through HCI tools (a mouse or a keyboard).

An initial list of five CAD actions (first five events presented in Table 29) was adopted from previous work on cognitive models, originating from the HCI field [7], [92]. Additional three CAD actions were identified throughout the exploratory coding of High performer's CAD modelling in both tasks. No additional CAD actions were observed from CAD modelling processes of Average and Low performers.

Table 29 List of CAD actions

CAD event/action	Description/Example
Moving a cursor	Moving a cursor on the screen.
Click on an icon/feature/textbox	Clicking a mouse button to select an icon/feature the cursor is positioned over.
Hovering over (pointing to) an icon/feature/point	Positioning a cursor over an icon/feature/point without selecting it.
Typing a value	Striking keyboard key(s) to input a value.
Entering (typing) a command	Striking keyboard command key, such as Escape, Delete, Enter, Undo/Redo, etc.
Rotate view*	Rotating a model view about the specified point by the specified angles in the directions of the screen X and Y axes [56].
Zoom in/out*	Modifying the zoom factor for the model view [56].
Dragging a geometric entity*	Pressing a mouse button to select a geometric entity and moving a mouse to change its position in the virtual global coordinate system.

*CAD events extracted empirically after coding activity

3.4.1.1.3 CAD operation layer

In feature-based CAD modelling, engineering designers rarely perform isolated CAD events due to the structure of CAD systems; instead, they perform functionally connected sequences of events to accomplish unit task goals (drawing a sketch and applying a CAD feature). Due to the assumed memory limits on human cognition [92], it is expected that engineering designers organise CAD modelling into cognitively manageable intermediate chunks when pursuing higher-level CAD modelling goals. In this analysis, such intermediate chunks are represented as CAD operations, defined as groups of CAD actions that are functionally and temporally connected in the CAD modelling sequence to pursue a common subgoal.

Given the large number of possible CAD operations, the identified operations were further grouped into subtask clusters based on the functional role the group of CAD actions plays in the CAD modelling sequence, such as drawing, dimensioning, constraining (the full list is available in Table 30). The definition of CAD operations and subtask clusters was inspired by previous similar work (e.g. [12], [58], [63]) detailed in Section 2.2.2.2.2 and corroborated by the empirical coding. Two additional subtask clusters (bolded in Table 30) were defined throughout the exploratory coding of Average and Low performers' CAD modelling sessions.

Table 30 List of subtasks/subgoals

Subtask cluster	Definition	CAD operations (examples)
Navigating CAD application	Switching between tabs and toolbars, opening and closing toolbar menus.	• Select Sketch tab to reveal sketch toolbar • Select Feature tab to reveal feature toolbar
	Turning sketching or featuring mode on or off.	• Enter/Exit a sketching mode • Enter/Exit featuring mode
	Turning CAD tool on or off	• Turn geometric entity tool on (e.g. select Line, Circle) • Turn geometric entity tool off (e.g. strike Escape or select Confirmation icon) • Turn dimensioning tool on (e.g. select Smart Dimension) • Turn dimensioning tool off (e.g. strike Escape, select Confirmation icon, deselect Smart Dimension icon)
Referencing	Creating or selecting a reference feature/geometry.	• Create a sketching plane • Select a sketching plane • Select an axis of revolution • Select a sketch as a basis for building a volume
Drawing	Combines geometric entities to create a shape.	• Draw a shape by using (and connecting) entities defined by 2 points/parameters • Draw a shape by using (and connecting) entities defined by 3 points/parameters • Drag a geometric entity • Trim geometric entity • Delete geometric entity
Constraining	Setting constraints to individual geometric entities or geometric relations between several geometric entities.	• Add geometric constraint to an entity (e.g. *Horizontal*, *Vertical*) • Add geometric constraint between two entities (e.g. *Tangent*, *Coincident*) • Delete geometric constraint
Dimensioning	Dimensioning in 2D: Defining dimensions of geometric entities and distances between geometric entities in sketches.	• Dimension a length (line, radius/diameter) • Dimension a distance/angle
	Dimensioning in 3D: Defining the value of a third dimension and its direction when applying CAD features.	• Define a dimension value/values • Define a direction of adding a volume (a third dimension)
Manipulating view	Adjusting the representation of a geometry or volume within the graphical area of CAD software. Manipulation does not contain any selection or change of geometry/volume.	• Setting a view • Reorienting/rotating a view • Translating a view • Zooming in/out
Deleting a feature*	Deleting the entire sketch or CAD feature from the feature tree.	• Delete a sketch • Delete a CAD feature
Other*		• Non-specified cursor movement at the end of the Execution part

*Subtasks extracted empirically after coding activity

Unit task operation layer

At a higher level, engineering designers organise their work into unit tasks, adopting the cognitive modelling notion of *small, cognitive manageable, quasi-independent tasks* [92]. While the task and the interactive system (CAD system) influence unit task structure, unit tasks are assumed to owe their existence primarily to memory limits of human embodied cognition [92]. In feature-based CAD systems, 3D models are built as a combination of CAD features while using 2D sketches as a basis [52]. Therefore, the unit tasks are here seen as object-related, which means they are related to a particular part of a geometry (volume) a participant models [97].

Accordingly, the CAD modelling process is represented as cycles of two main unit task types:

1) Sketching a part of a geometry or deleting a sketched part of a geometry;
2) Applying a CAD feature (such as extrude, revolve, sweep, etc.) to generate a new volume from the sketched part of geometry/edit the existing part of a volume (e.g. with fillet or chamfer) or deleting the CAD feature.

While unit task operations describe what object-related CAD modelling step is being pursued at the higher process layer (e.g. sketch-related or feature-related work on a specific portion of geometry), the subsequent operator mapping requires additional context about how the same unit task unfolds over time from an information-processing standpoint. For this reason, each unit task was further segmented into recurring unit task phases. This phase segmentation was introduced to support consistent and reproducible assignment of perceptual and cognitive operator activity to intervals where such processing is expected to dominate, beyond what can be inferred from CAD events alone.

3.4.1.2 Unit task phases

Following cognitive modelling literature, unit tasks are often described as comprising *acquisition* and *execution* cycles [92]. During coding, *inspection* was additionally identified as a recurring third information-processing phase to represent explicit checking behaviour (Table 31). This phase was observed first in the High performer's CAD modelling sessions and then consistently in the Average and Low performers' sessions. A small residual category ("Other" in Table 31) was retained for transitions that did not reflect a new goal/subgoal.

Table 31 Unit task phases

Unit task phases	Definition
Acquisition	Cognitive and perceptual actions required to understand the task and define goals/subgoals. *During acquisition the user builds a mental representation of the task/subtask* [92].
Execution	Cognitive, perceptual, and motor actions required for CAD events. *During execution the user calls on the system facilities to accomplish the task* [92]. These system facilities are C, P, and M actions, which are modelled through C, P, and M operators.
Inspection*	Cognitive, perceptual, and motor actions required for checking the accuracy of the modelled geometry [52], [97]. Example: zoom in/out and rotate at the end of the Execution, after which there was a pause before moving to the next subtask which is different from the previously inspected.
Other*	Cursor movement at the end of Execution or after Inspection, this CAD event does not result in a new goal/subgoal and it's often followed by Acquisition. It seems like a transition period between unit tasks/subtasks.

*Unit task phases extracted empirically after coding activity are bolded

Unit task phases were introduced as a practical coding context for consistent operator assignment. The lowest-level observable data (CAD events/actions) capture what the engineering designer does physically (mouse/keyboard actions and visible CAD system response) and therefore provides a sufficient basis for assigning motor and CAD system response operators. However, they do not by themselves indicate when processing is primarily devoted to perceiving, planning/thinking, executing, or evaluating. Phase segmentation therefore provides a complementary, orthogonal lens that marks recurring information-processing intervals within a unit task and supports reproducible application of operator rules. In practical coding terms, the phase segmentation provides information-processing interpretation needed to consistently assign P and C operators to intervals where perceptual and cognitive processing is expected to dominate.

With unit task phases defined, the coded CAD event/action stream (and its aggregation into CAD operations and unit task operations) can be interpreted in terms of the resources engaged during CAD modelling. The next section therefore introduces the operator set used in this study (P, C, M, and CAD system response R) and specifies how these operators were assigned to the time-aligned scripts.

3.4.1.3 Operators and systems

With the behavioural stream structured into CAD events/actions (and aggregated into CAD operations and unit task operations) and with unit task phases defined as an information-processing context, the next step was to interpret the time-aligned scripts in terms of the processing resources engaged during CAD modelling. Following MHP and KLM practices [7], [92], four operators were used to represent these resources: Perceptual (P), Cognitive (C), Motor (M), and CAD system Response (R). Table 32 lists the associated operation types and systems to clarify what each operator represented in the present coding context; the coding did not include specifying which of these operations occurred during CAD modelling process.

Table 32 List of actions, operators, and associated systems

System	P/C/M/R operation	Description	Operator
CAD	CAD system response	CAD system processes engineering designers' input	R
Motor	Moving a mouse	Move a cursor by moving the mouse with a hand	M
Motor	Clicking a mouse button	Press and instantly release a mouse button	M
Motor	Pressing a mouse button	Press and hold a mouse button	M
Motor	Releasing a mouse button	Release pressed mouse button	M
Motor	Spinning the mouse wheel	Spin the mouse wheel (up or down)	M
Motor	Striking a keyboard key	Press and instantly release a keyboard key	M
Motor	Pressing a keyboard key	Press and hold a keyboard key	M
Motor	Releasing a keyboard key	Release pressed keyboard key	M
Perceptual	Visual scanning	Scanning the screen, often without a one clear target	P
Perceptual	Perceiving	Processing incoming sensory information	P+C
Perceptual	Recognising the target	Identifying an icon/feature as a match to what is stored in WM	P+C
Cognitive	Attention shift	Low-level cognitive action: Moving attention from one location/icon/feature to another	P+C
Cognitive	Visual search	Low-level cognitive action: Intentionally searching for a specific target	P+C
Cognitive	Recalling/Remembering	Retrieve relevant knowledge from the WM or the LTM [197]	C
Cognitive	Understanding	Construct meaning from a technical drawing [197]	C
Cognitive	Applying	Carry out or use a procedure in a given situation [197]	C
Cognitive	Analysing	Break material into its constituent parts and determine how the parts relate to one another and to an overall structure or purpose [197]	P+C
Cognitive	Evaluating	Make judgments based on criteria and standards [197]	P+C
Cognitive	Creating	Put elements together to form a coherent or functional whole; reorganize elements into a new patten or structure [197]	C

Operator assignment was then performed manually by applying explicit rules (Table 33) to the coded scripts. In brief, CAD actions and operations provide the observable basis for assigning M (motor activity during mouse/keyboard actions) and R (CAD system response intervals visible in the interface), whereas the surrounding unit task phase context provides the information-processing interpretation needed to apply rules for assigning P and C consistently. The rules were adopted from the literature (e.g. [92]) and extended where necessary based on recurring patterns observed across the three participants' coded sessions (extensions are bolded in Table 33). Specifically:

- Operator R was assigned to intervals during which the CAD software processed input (e.g. indicated by a processing icon or by temporarily unavailable interface commands until the scene/state updated).
- Operator M was assigned to the full duration of each CAD action involving physical input (mouse/keyboard); this means to every CAD action from Table 29 except *Hover over*, where no motor execution was present.

- Operators P and C were assigned according to cues dependent on unit task phase, CAD action, and CAD operations; *acquisition* and *inspection* were treated as predominantly perceptual and cognitive, while during execution P and C were assigned selectively to CAD actions/operations that plausibly reflect target identification, attention shifts, visual search, or evaluation (as specified in Table 33).

Table 33 Coding rules for assigning operators

Unit task phase	**CAD event/Subtask**	**Operator**
Acquisition	• Entire duration	P+C
Execution	• All CAD events except *Hover over (point to) an icon/feature/point*	M
	• Pause at CAD action level • *Hover over (point to) an icon/feature/*point at CAD action level • CAD system response followed by an input related to an updated processed during the activity of R operator • *Moving a cursor* if it is slow • *Moving a cursor* if it is slow, but intermitted with brief hovering over several icons/features	P+C
	• ***Dragging*** **in** ***Drawing*** **at CAD operation level if it was slow** • ***Rotating*** **at CAD operation level if it was slow**	P+C+M
	• ***Manipulating a view*** **at CAD operation level (short in duration, before or in the middle of a subtask, often to select some tiny part of a geometry)**	P+M
Inspection	• **Entire duration**	P+C
	• **Cursor movement** • ***Manipulating a view*** **(usually longer in duration then** ***Manipulating a view*** **as a part of** ***Execution*****, occurs at the end of Subtask/Unit task)**	P+C+M

*CAD events extracted empirically after coding activity are bolded

The main purpose of this initial operator mapping was to derive a time-aligned operator timeline that supports process visualisation and comparison of resource engagement across conditions and CAD modelling segments. Because the manual coding from screen recordings does not capture fine-grained overlap between perception/cognition and motor execution with high precision, the mapping prioritised consistency and interpretability in identifying when perceptually and cognitively dominant intervals occur around CAD actions and across unit task phases. A more detailed treatment of parallelism is addressed in the second empirical study, where interaction logs of higher resolution are available.

Defining process layers (CAD actions, CAD operations, unit task operations) and the unit task (information-processing) provided a common, time-aligned basis for comparing how engineering designers solved subtasks/subgoals across hierarchical levels (low, intermediate, and high) and how perceptual, cognitive, and motor resources (P, C, and M operators) were engaged while doing so. The next section describes the data analysis procedure used to conduct this comparison.

3.4.2 Data analysis

To enable a qualitative and structured comparison of CAD modelling behaviour across two conditions (isometric and orthographic projection), the study did the following:

1) Extracted and organised participant's CAD modelling process in each condition by manually coding screen recordings and structuring the resulting CAD event stream into process layers (CAD actions → CAD operations/subtask clusters → unit task operations), complemented by unit task phases (acquisition, execution, inspection) and operators (P, C, M, R).
2) Visualised a descriptive model of CAD performance that makes the CAD modelling sequence explicit across the identified layers (timelines of unit task operations and subtask clusters), unit task phases, and operators.
3) Quantified process-based metrics using descriptive statistics and distribution measures, including the number and typical duration of unit tasks, unit task phases, CAD operations (subtask clusters), and CAD actions, as well as time distributions of unit task phases and subtask clusters (within unit tasks and cumulatively at the task level).
4) Computed and visualised operator utilisations (P, C, M; and CAD system response where applicable) across CAD modelling segments, by estimating utilisation values per unit task and per subtask cluster and comparing the resulting patterns between projection conditions.

The coded scripts were exported from ELAN annotation tool [196] and analysed in R. Descriptive statistics were used to summarise the number of occurrences and typical durations of unit tasks, unit task phases, subtask clusters (CAD operations), CAD actions, and operator activity. For duration summaries, both Mean (M) with Standard Deviation (SD) and Median (Med) with Median Absolute Deviation (MAD) were reported, depending on the metric and table/figure context.

For CAD actions coded as instantaneous (*Clicking on an icon/feature/textbox* and *Entering (striking) a command*), a predefined duration of 1 ms was assigned as a fixed coding value to represent instantaneous events, and this value was kept constant across occurrences, tasks, and participants.

Time distributions of unit task phases and subtask clusters were computed as proportions of a reference segment. Within each unit task, the proportion of a given unit task phase/subtask cluster was calculated according to Equation (30):

$$Proportion_{\text{Segment}} = \frac{Duration_{\text{Segment in Unit task}}}{Duration_{\text{Unit task}}} \tag{30}$$

Analogously, cumulative distributions at the task level were computed by dividing the total duration of each unit task phase/subtask cluster by the total task duration.

Operator utilisation was computed as the ration of time during which a given operator was active within a task segment to the duration of that segment (e.g. unit task or subtask cluster) [115], as defined in Equation (31):

$$Utilisation_{\text{Operator}} = \frac{Duration_{\text{Operator in Segment}}}{Duration_{Segment}} \tag{31}$$

Finally, results were compared qualitatively between isometric and orthographic conditions by inspecting differences in 1) sequencing and segmentation patterns (timelines), and 2) the extracted process metrics (counts, durations, distributions, and utilisations). This comparison was conducted at each analysis level (unit task operations, unit task phases, CAD operations/subtask clusters, CAD actions, and operator utilisations) to examine whether projection-related differences are detectable and, if so, at which level they provide the most informative discrimination between conditions.

3.4.3 Results

The results are presented as a working example for one representative participant (Average performer) to demonstrate how the proposed coding scheme and metrics distinguish CAD modelling behaviour between isometric and orthographic projections. Results are reported in the following order: 1) a descriptive timeline model of performance across process layers (Section 3.4.3.1), 2) unit task structure and unit task phases (counts, typical durations, and phase time distributions; Section 3.4.3.2), 3) subtask cluster distributions within unit tasks (Section 3.4.3.2) and cumulatively at the task level (Section 3.4.3.3), 4) CAD action summaries (Section 3.4.3.4), and 5) operator utilisations across unit tasks and subtask clusters (Section 3.4.3.5). Results for High performer and Low performer can be found in Appendix A; they are not presented in this section, but they are used as a part of a discussion (Section 3.4.4) when considering the relationship between process-based performance metrics and output-based performance scores across performance profiles.

3.4.3.1 Descriptive model of CAD performance (timeline)

For the Average performer, the two timelines (Figure 19 for the isometric, Figure 20 for the orthographic condition) show the same basic pattern at the unit task level; a long first *sketch* is used as a basis to which the following unit tasks relate. However, the organisation of unit tasks differs between conditions in terms of the number and the structure of the unit tasks following the first *sketch*. CAD modelling from the isometric projection seems more streamlined at this level; it was sketch-dominant, with a few brief *CAD feature* unit tasks. CAD modelling from the orthographic projection was longer and more iterative, with more *CAD feature* unit tasks, and short *sketching* unit tasks in later stages where visible rework and refinement occurred.

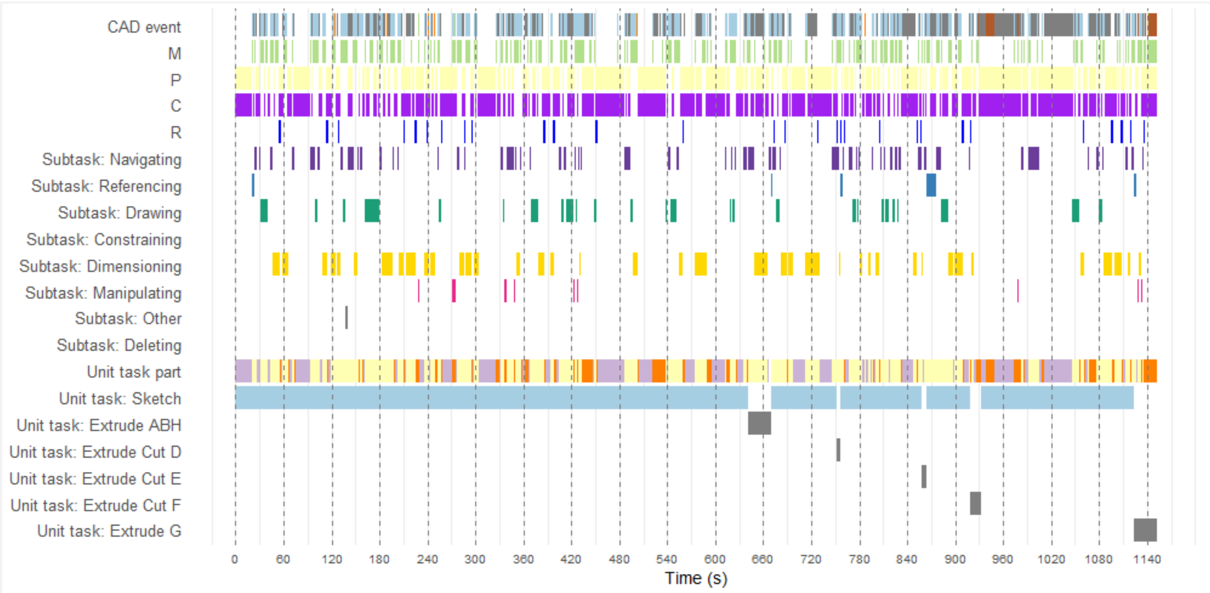

Figure 19 Average performer: Descriptive model of CAD performance - Isometric condition

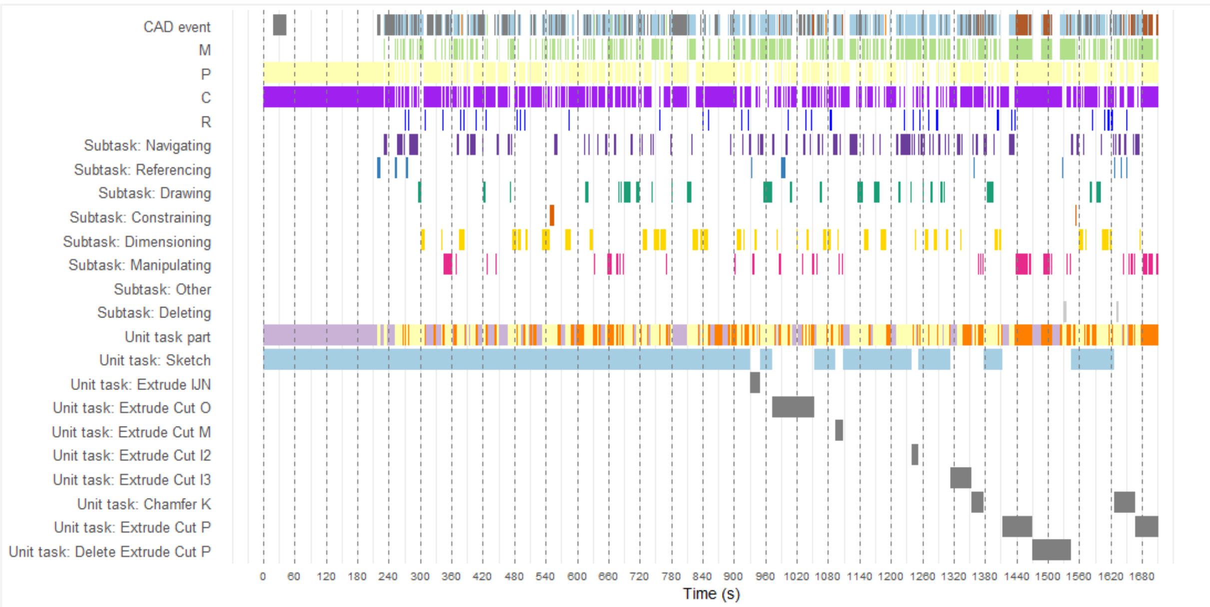

Figure 20 Average performer: Descriptive model of CAD performance - Orthographic condition

3.4.3.2 Unit task operations and unit task phases

Compared to the isometric condition, CAD modelling from the orthographic condition was organised into more unit tasks and included additional unit task types (see Table 34), which is consistent with more iterative workflow described above. *Sketch* occurred more often in the orthographic condition, but with shorter duration, suggesting that sketching was in this case divided into more, shorter unit tasks. *Extrude Cut* was both more frequent and substantially longer when CAD modelling from the orthographic projection, indicating that this type of CAD feature took a much larger share of the CAD modelling process in this condition. Moreover, only the orthographic condition contained rework, including *Delete Extrude Cut*.

Table 34 Average performer: Types, number of occurrences, and average duration of unit tasks

Unit task	N of occurrences		Duration (M) in [s]		Duration (Med) in [s]	
	Iso	Ortho	Iso	Ortho	Iso	Ortho
Sketch	5	7	214	186	101	60
Extrude Cut	3	6	8	41.8	6	42.5
Delete Extrude Cut	-	1	-	74	-	74
Extrude	2	1	29	19	29	74
Chamfer	-	2	-	31.5	-	31.5

3.4.3.2.1 Unit task phases

CAD modelling from the orthographic projection included more overall occurrences of unit task phases (280 vs. 197) and showed a shift in time distribution toward *inspection* (Table 35).

Table 35 Average performer: Number of occurrences, average duration, and distribution of unit task phases

Unit task phase	N of occurrences		Duration (M) in [s]		Duration (Med) in [s]		Distribution [%]	
	Iso	Ortho	Iso	Ortho	Iso	Ortho	Iso	Ortho
Execution	74	107	8.34	7.30	6	5	53.26	44.74
Acquisition	56	71	6.39	7.44	3	2	31.16	30.84
Inspection	67	102	2.67	4.10	1	2	15.58	24.42

Execution was slightly shorter per each occurrence when CAD modelling from the orthographic projection, while *acquisition* had a slightly higher M but lower Med, which is consistent with many short *acquisition* occurrences in addition to few long ones (also visible as an early sustained *acquisition* window in Figure 22). Moreover, *inspection* was clearly longer in *orthographic* condition when compared to the isometric condition, as visualised in Figure 21.

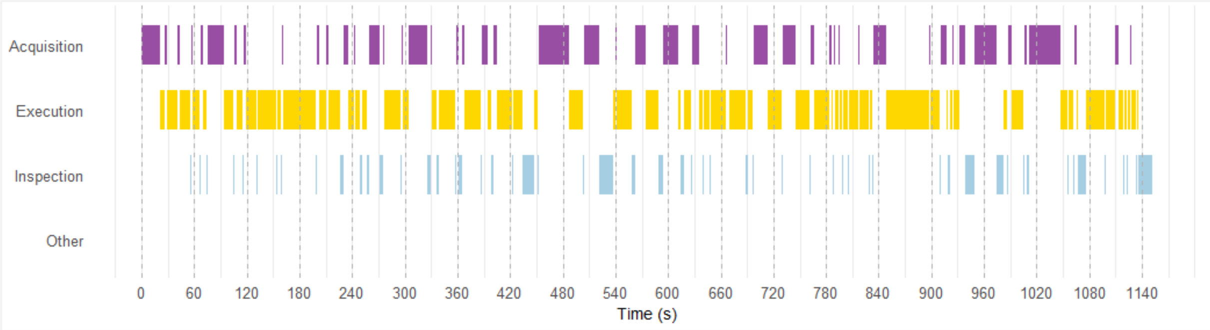

Figure 21 Average performer: Visualisation of unit task phases - Isometric condition

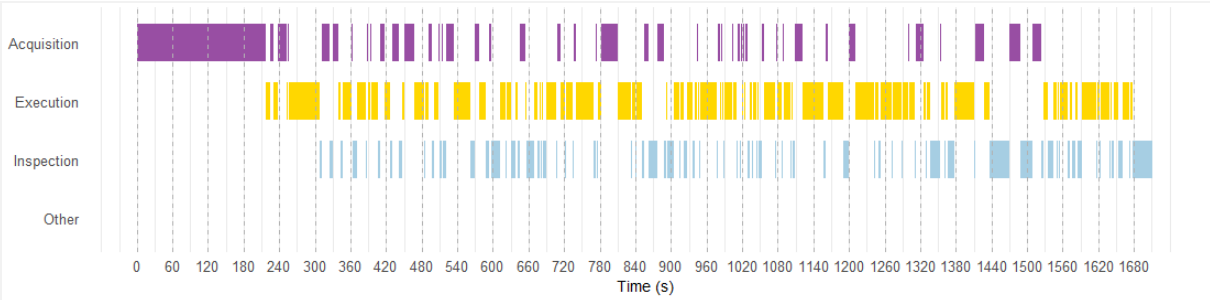

Figure 22 Average performer: Visualisation of unit task phases - Orthographic condition

The number of *acquisition-execution-inspection* cycles was similar in both conditions (38 in isometric and 36 in orthographic). Since there were many more unit task phase occurrences in the orthographic condition, the similar number of cycles implies that, within each cycle, CAD modelling from orthographic projection contained more repeated unit task phases, which aligns with the longer, more iterative unit task phase organisation (Figure 20 and Figure 22).

3.4.3.2.2 Distribution of subtask clusters in unit tasks

In the isometric condition, unit tasks tend to have a clear distribution of subtask clusters (Figure 23). *Sketch* unit tasks were consistently dominated by *acquisition* and *inspection*, typically taking the largest share (45.9% to 62.9%). The remaining time was mainly split between *dimensioning* and *navigating.* A notable exception is *Sketch(4)*, which contained a lengthy *referencing* subtask. *Extrude* and *Extrude Cut* unit tasks were mostly driven by *navigating* and *dimensioning*, with relatively little diversity.

The orthographic condition showed greater diversity within unit tasks when observing the process in CAD operation layer (Figure 24). *Sketches* varied in what subtasks were prevalent in their execution; *Sketch(1)* was dominated by *acquisition and inspection*, *Sketch(2)* by *drawing* and *navigating*, while later sketches were performed as a combination of *drawing* (33.3%) and *dimensioning* (30.6%), and very little *acquisition* and *inspection* (8.3%). *Extrude* and *Extrude*

Cut unit tasks in this condition included more *manipulating* and more mixed distributions when compared to the isometric condition, which is consistent with iterative refinement and rework.

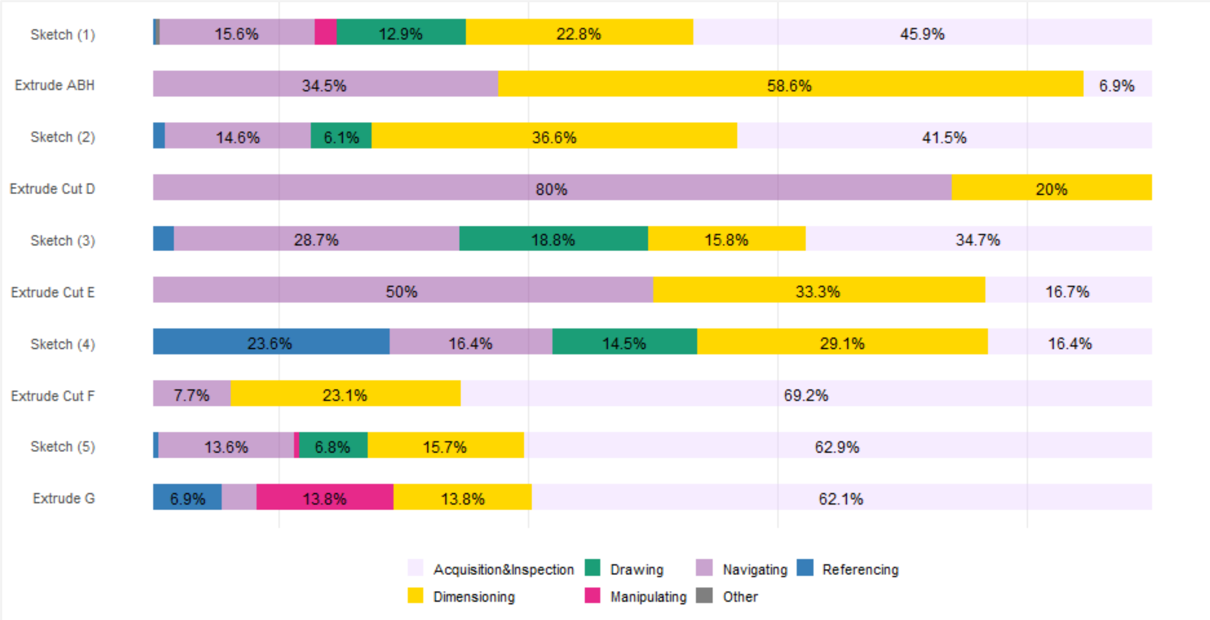

Figure 23 Average performer: Distribution of subtask clusters across unit tasks - Isometric condition

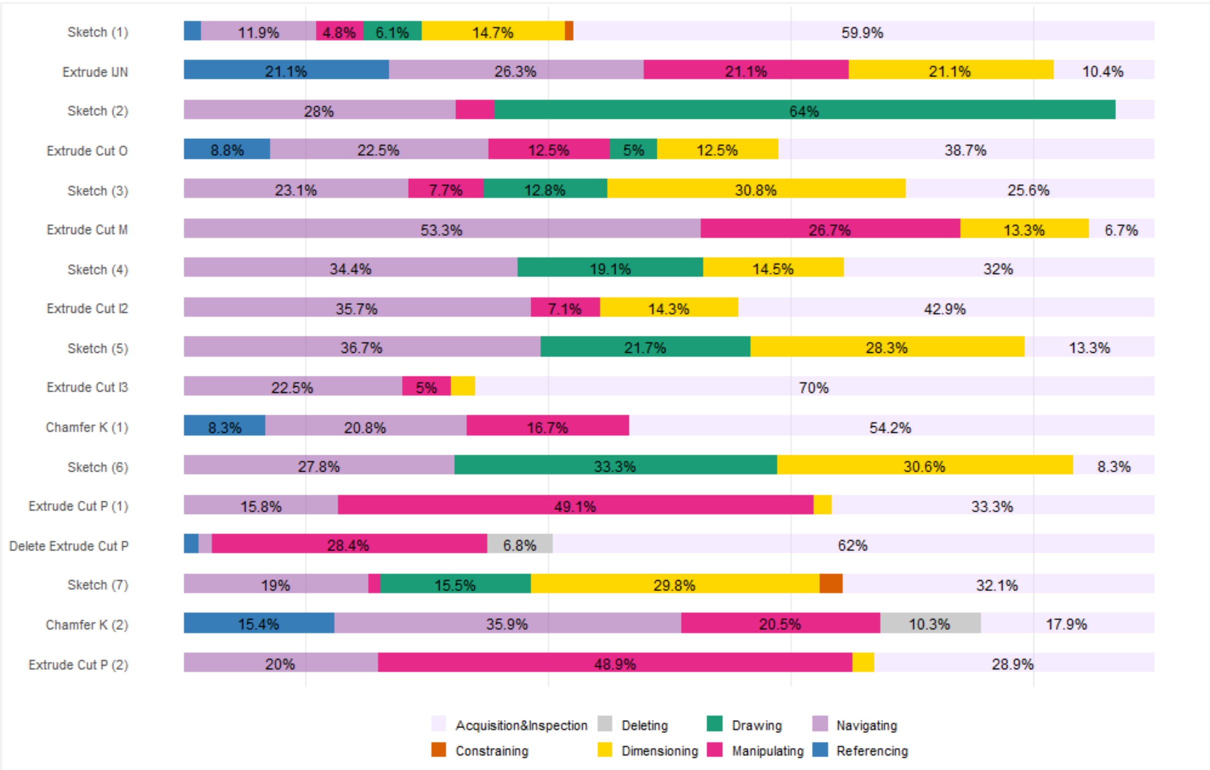

Figure 24 Average performer: Distribution of subtask clusters across unit tasks - Orthographic condition

3.4.3.3 Subtask (CAD operation) clusters

Analysis of the cumulative distribution of subtask clusters (considering the entire CAD modelling task, not individual unit tasks) revealed a difference in what the CAD modelling time was spent on between the projections (see Table 36 and Figure 25).

Compared to isometric, the orthographic condition contained more *navigating* occurrences, and they were slightly longer on average. *Dimensioning* and *drawing* were similarly represented in both conditions, with some differences in durations. The largest increase was noticed in *manipulating*, primarily driven by many more occurrences of this subtask cluster (63 vs. 11), while their typical (Med) length stayed the same. Additionally, more *referencing* was observed in the orthographic condition, while duration was similar to the isometric condition.

Table 36 Average performer: Number of occurrences and average duration of subtasks at the level of the entire task

Subtask	N of occurrences		Duration (M) in [s]		Duration (Med) in [s]	
	Iso	Ortho	Iso	Ortho	Iso	Ortho
Navigating	85	113	2.29	2.68	2	2
Dimensioning	42	44	6.31	5.5	6	5
Other	1	-	2	-	2	-
Drawing	27	25	4.74	5.8	4	4
Manipulating	11	63	1.73	2.44	1	1
Constraining	-	2	-	5.00	-	5
Referencing	7	15	3	2.4	2	2
Deleting	-	2	-	4.5	-	4.5

When looking at how total time was distributed across subtask clusters, Figure 25 suggests that CAD modelling from isometric projection was dominated by *dimensioning*, followed by *navigating* and *drawing*, with *manipulating* contribution very little. In the orthographic condition, *navigating* became the largest component, *dimensioning* dropped noticeably, *drawing* also decreased, and *manipulating* strongly increased. Small additional shares came from *constraining* and *deleting*, which were absent in the isometric condition.

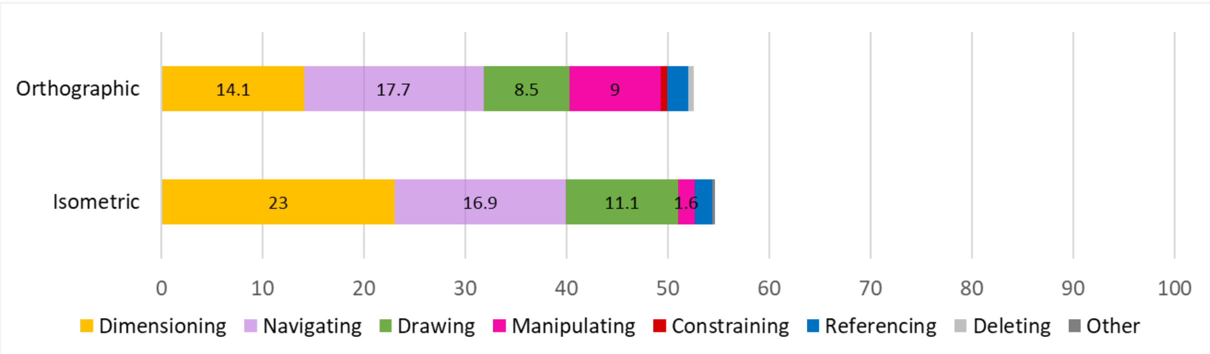

Figure 25 Average performer: Distribution of subtasks

3.4.3.4 CAD actions

The number of occurrences of almost all CAD actions differed between the conditions; the similar number was noticed for *typing a value* and *dragging a geometric entity* (see Table 37). The Average performer executed more CAD actions when CAD modelling from the orthographic projection, while their typical duration stayed similar as in the isometric condition.

Table 37 Average performer: Number of occurrences and average duration of CAD actions at the level of the entire task

CAD action	N of occurrences		Duration (M) in [s]		Duration (Med) in [s]	
	Iso	Ortho	Iso	Ortho	Iso	Ortho
Moving a cursor	402	661	0.92	0.95	1	1
Clicking on an icon/feature/textbox	235	328	0.001	0.001	0	0
Hovering over an icon/feature/point	276	376	1.28	1.03	1	1
Typing a value	33	32	0.27	0.16	0	0
Rotating a view	7	47	3.71	2.11	1	1
Zooming in/out	21	58	0.29	0.19	0	0
Dragging a geometric entity	1	-	0	-	-	-
Entering (striking) a command	-	18	-	0.001	-	0

3.4.3.5 Operator utilisations

Operator utilisations were analysed at two complementary levels across two conditions: 1) averaged within each unit task (Section 3.4.3.5.1), and 2) averaged within each subtask cluster (Section 3.4.3.5.2).

3.4.3.5.1 Operator utilisations across the unit tasks

Figure 26 (isometric) and Figure 27 (orthographic) show time-based operator utilisations for each unit task. Here, utilisation is defined as the proportion of the unit task duration during which an operator was coded as active (Equation (31)). For example. M=0.90 means that motor activity (mouse/keyboard actions) occupied approximately 90% of that unit task's time, whereas P= 0.60 means that perceptual processing (e.g. related to visual scanning) was coded for approximately 60% of that unit task's time. Therefore, these values reflect time allocation, not how often an operation occurred.

In both conditions, operator utilisations depended on the unit task type: S*ketch* unit tasks showed higher P and C utilisation, whereas several feature-application unit tasks (*Extrude*/Extrude *Cut*/*Chamfer*) coincided with higher M utilisation, consistent with more continuous HCI while executing CAD feature commands. In both conditions, the final unit tasks in the sequence showed elevated utilisations across operators, indicating convergent processing demands near task completion; high C and P utilisations were maintained and accompanied by a late rise in M utilisation in the isometric condition, while in the orthographic condition P, C, and M gradually elevated at the similar rate (see Figure 27).

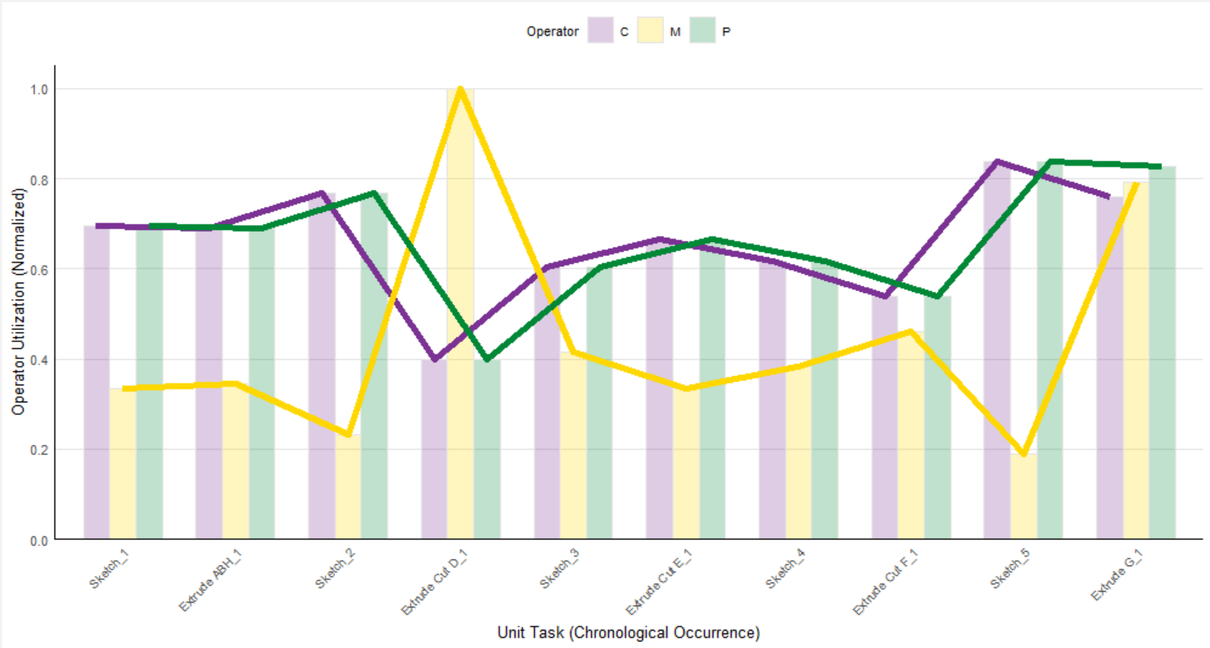

Figure 26 Average performer: Average utilisation of operators in each unit task – Isometric condition

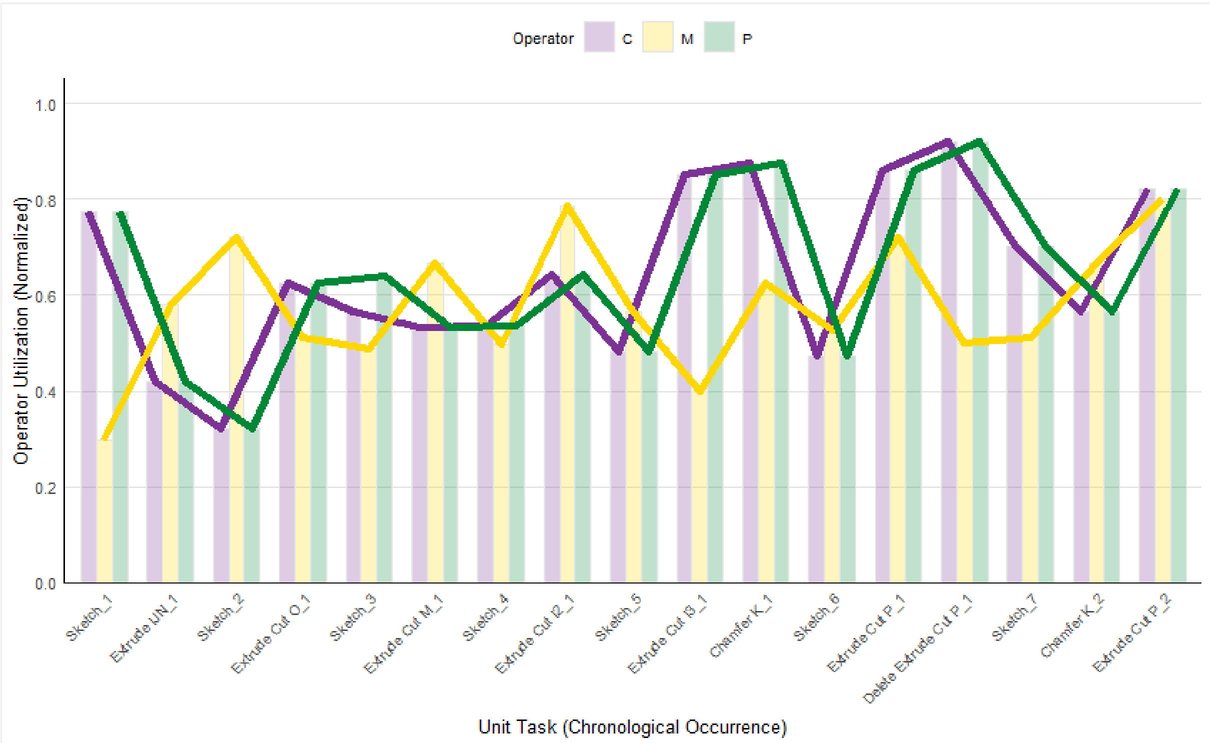

Figure 27 Average performer: Average utilisation of operators in each unit task – Orthographic condition

Despite these shared characteristics, the temporal allocation of processing resources across unit tasks differed between projection formats. In the isometric condition, P and C remained comparatively stable at mid-to-high levels for most unit tasks across the task performance, and the main deviation is a single strongly M-dominant event at *Extrude Cut D*, where M peaked while both P and C dropped to their lowest levels (≈ 0.4). In practical terms, this pattern suggests that this unit task was executed with little time spent in pauses/*hovering over*/*inspection* intervals and comparatively more uninterrupted motor HCI (e.g. selecting the CAD feature, entering parameters, confirming). After this drop, P and C recovered and remained relatively stable through the subsequent *Sketch–Extrude Cut* alternations, until the late increase in M utilisation during *Extrude G* produced convergence at the end of the sequence.

In contrast, the orthographic condition showed more frequent crossovers and larger reallocations between P/C and M from unit task to unit task. This is consistent with the more iterative unit task organisation already observed in the timelines; CAD modelling from the orthographic projection required more frequent transitions between *acquisition*/*inspection* (higher P/C utilisation) and execution (higher M utilisation), rather than maintaining a relatively uniform P/C level. The later refinement segment (around *Extrude Cut I3 → Chamfer K1 → Extrude Cut P*/*Delete Extrude Cut P*) is characterised by plateau of high P and C utilisation with more local variation in M; in practice, this corresponds to spending a larger portion of time in perceiving/deciding/evaluating (e.g. verifying geometry, inspecting dimensions/relations, checking feature results) interleaved with shorter bursts of execution.

Overall, for the Average performer, the orthographic projection is associated with more frequent time reallocation between P/C processing and M execution across unit tasks, while CAD modelling from the isometric projection shows a more uniform P/C allocation disrupted mainly by a single unit task strongly dominated by M utilisation.

3.4.3.5.2 Operator utilisations across the subtask (CAD operator) clusters

The operator utilisation summaries for the Average performer showed that the dominant operator depends on the subtask type, while their values differed depending on the projection format (Table 38).

Table 38 Average performer's average operator utilisation across subtasks

	Navigating		Referencing		Drawing		Constraining		Dimensioning		Manipulating		Deleting	
Operator	**Iso**	**Ortho**	**Iso**	**Ortho**	**Iso**	**Ortho**	**Iso**	**Ortho**	**Iso**	**Ortho**	**Iso**	**Ortho**	**Iso**	**Ortho**
P	0.57	0.43	0.67	0.50	0.52	0.47	-	0.30	0.57	0.51	0.89	0.96	-	0.22
C	0.57	0.43	0.67	0.50	0.52	0.47	-	0.30	0.57	0.51	0.87	0.96	-	0.22
M	0.67	0.71	0.37	0.58	0.54	0.59	-	0.70	0.43	0.53	0.94	0.91	-	0.78

Across both projections, *manipulating a view* showed the highest utilisations across operators (P and C near max; M also high), which in practice reflects that view manipulation operations are dominated by sustained visual processing and control actions (e.g. continuous perception and decision-making while actively rotating/zooming/positioning the view). For this Average performer, CAD modelling from the orthographic projection further increased P and C utilisation during manipulating (P=0.96, C=0.96), suggesting more time spent visually interpreting and verifying while manipulating the view.

In contrast, *navigating* shifted toward a more motor-intensive execution in the orthographic condition; M increased slightly (0.71 vs. 0.67) while P and C decreased (Table 38). In practical terms, this indicates that a larger share of navigation time consisted of continuous clicking/selection/tab switching with comparatively less time spent in perceptual and cognitive processing (e.g. shorter pauses or less visual search within navigation). A similar pattern appears in *referencing*, where CAD modelling from the orthographic projection increased M utilisation (0.58 vs. 0.37) and reduced P/C utilisation (0.50 vs. 0.67), consistent with referencing being implemented through more active interface interaction in that condition (e.g. selecting planes/features, switching views/tools) rather than longer interpretive pauses.

Finally, *dimensioning* showed a clear effect of the projection format for this engineering designer; in the orthographic condition, utilisation of all operators increased, implying that dimensioning operations required more concurrent processing and interaction actions (e.g. more back-and-forth between interpreting the drawing and entering/confirming values). Changes in utilisations when *drawing* were smaller (M: 0.59 vs. 0.54; P/C: 0.47 vs. 0.52), suggesting broadly similar time allocation during this CAD operation across projections.

For this Average performer, moving from isometric to orthographic projections was associated with more motor-intensive *navigating* and *referencing*, higher P and C utilisation during *manipulating a view*, and a more demanding *dimensioning*. These results align with the earlier process evidence of longer and more iterative CAD modelling from the orthographic condition.

3.4.4 Discussion

The initial findings related to CAD performance (explained in Section 3.3) motivated a deeper investigation of how CAD modelling performance unfolds over time. A central contribution of this section is methodological; the study developed a hierarchical structure and associated coding scheme that extracts an interpretable structure of a CAD modelling task (its process) from screen recordings. By manually extracting CAD events and progressively organising them into CAD actions, CAD operations, and unit task operations, complemented by unit task phases (acquisition, execution, inspection) and time-aligned P, C, and M operator mapping, the analysis provides a reusable process representation that can be applied later across engineering designers and CAD modelling tasks, and aligned with CL indicators. This decomposition therefore makes the evolution of CAD modelling visible across multiple process layers (and associated hierarchical levels), rather than reducing performance description and assessment to the final number of sketches and applied CAD features.

Using this hierarchical structure and the coding scheme, the detailed qualitative comparison of conditions across three engineering designers representing the output-based performance profiles (High, Average, Low) provided an answer to the RQ 1.2: the CAD modelling process differs between isometric and orthographic projections, and differences are detectable at multiple hierarchical levels. However, the nature of these differences depends on output-based performance profile; the High performer exhibits comparatively stable segment organisation and CAD system interaction across projections, whereas the Average and Low performers show larger differences, including more increased occurrences of segments and evidence of rework. The following sections discuss these differences at each level, additionally considering three output-based performance profiles to finally highlight the hierarchical level that seems to be the most informative for discriminating process-based performance between conditions.

3.4.4.1 Unit tasks and unit task phases

At the unit-task level, differences between conditions are visible in 1) unit task decomposition (how many unit tasks is formed and which operation types they involve) and 2) phase allocation within unit tasks (how time is distributed across *acquisition*/*execution*/*inspection*). Importantly, the observed effects differ across performance profiles (High, Average, and Low performer). The unit task decomposition (number of occurrences) and rework effects depend on the performance profile. The increase in *inspection* share in the orthographic condition is the most informative and consistent discriminator between projection formats at this level, indicating that CAD modelling from orthographic projections increases verification/validation operations regardless of output-based performance profile.

3.4.4.1.1 Unit tasks

The number of unit tasks was larger when CAD modelling from orthographic than the isometric projection for the Average (17 vs. 10) and Low (17 vs 8) performers, while the High performer showed no increase but a slight reduction instead (12 vs. 14). In addition to this increase in the number of segments, the orthographic condition introduced explicit rework unit tasks for the Average and Low performers only, visible as *delete* unit task operations (e.g. *Delete Extrude Cut* appears for both Average and Low; *Delete Sketch* appears for the Low performer). This pattern indicates that, for the Average and Low performers, CAD modelling in the orthographic condition not only required more frequent decomposition into smaller CAD modelling segments, but also triggered corrective iteration (undoing/removing prior CAD modelling decisions) before the model could be finalised. In contrast, the absence of unit tasks indicating rework for the High performer suggests a comparatively stable CAD modelling plan and fewer corrective resets, even when using multi-view orthographic representations.

A plausible explanation for this differences and their dependence on the performance profiles is that higher information-processing demand requires decomposition into a higher number of unit tasks when the amount of information that must be maintained and integrated in WM exceeds available capacity [7]. In CAD modelling from technical drawings, a key driver of such demand is related to the extent to which the engineering designer must integrate spatial information across representations to infer and maintain a coherent mental 3D model while planning and executing CAD actions and operations. Orthographic projections distribute geometry across multiple 2D views, requiring frequent cross-referencing and mental integration of design characteristics across views (i.e. higher element interactivity), whereas an isometric view provides a single pictorial representation that can support faster formation and maintenance of the mental 3D model. From this perspective, the marked increase in unit task segmentation for the Average and Low performers when CAD modelling from the orthographic projection suggests higher integration demands in this condition and, consequently, a greater need to externalise or re-instantiate intermediate representations through additional *acquisition-execution* cycles to manage WM limitations. In contrast, the High performer executed a similar number of unit tasks across projections, suggesting that the orthographic projection did not increase the required task decomposition for this engineering designer. This can be interpreted as greater proficiency in rapidly constructing and maintaining an accurate internal mental 3D model from multi-view projections, decomposing the object into constituent elements, and planning the CAD modelling sequence with fewer resets or re-checks of the geometry. Such results align with Rynne and Gaughran's account of expert performance in constructing mental models from given views and organising CAD modelling sequences accordingly [35].

3.4.4.1.2 Unit task phases

While unit task decomposition varied with performance profile, the time allocation across unit task phases showed a consistent effect of the projection format; across all High, Average, and Low performer, the orthographic condition was associated with an increase in *inspection* share (8.8% to 9.7%), indicating that CAD modelling from orthographic projection required more verification/validation activity, not merely longer execution (Table 39). Specifically, *inspection* rose from 10.1% to 18.9% for the High performer, 15.6% to 24.4% for the Average performer, and 10.2% to 19.9% for the Low performer.

Table 39 Number of occurrences, and relative duration of unit task phases across conditions and performance profiles

Performance profile	N of occurrences		N of AEI cycles		Acquisition [%]		Execution [%]		Inspection [%]	
	Iso	Ortho	Iso	Ortho	Iso	Ortho	Iso	Ortho	Iso	Ortho
High	149	159	29	28	35.70	30.90	54.20	50.20	10.10	18.90
Average	197	280	38	36	31.16	30.84	53.26	44.74	15.58	24.42
Low	128	215	17	29	31.40	19.40	56.10	60.50	10.20	19.90

The way this inspection increase is achieved, however, differs by performance profile (Table 39). For the High performer, the orthographic projection reallocates time away from *acquisition* and *execution* (execution: 54.2% to 50.2%; acquisition: 35.7% to 30.9%) toward *inspection*, while the number of Acquisition-Execution-Inspection (AEI) cycles remains essentially unchanged (29 to 28). This pattern suggests that orthographic projections primarily increased the need for verification without triggering additional decomposition into more AEI loops. For the Average performer, the orthographic condition produces the strongest shift away from *execution* (53.26% to 44.74%) toward *inspection* (15.58% to 24.42%), with *acquisition* nearly unchanged (31.16% to 30.84%) and AEI cycles slightly reduced (38 to 36). This again points to orthographic projections increasing CL mainly via *inspection* demands, rather than by forcing more frequent re-initiation of unit tasks. The Low performer showed a qualitatively different trial-and-error performance; orthographic projection increased unit task phase occurrences (128 to 215) and increased AEI cycles (17 to 29), alongside higher *inspection* (10.2% to 19.9%) and higher *execution* (56.1% to 60.5%) shares, while *acquisition* dropped (31.4% to 19.4%). Taken together with the emergence of rework unit tasks (discussed in Section 3.4.4.1.1), this pattern is consistent with repeated re-initiation and corrective iteration; spending proportionally less time preparing (acquisition) and more time alternating between doing (execution) and checking (inspection) as the model is incrementally built.

Across the three performance profiles and both projection formats, the per-occurrence durations of unit task phases are generally shorter than the typical values reported by Card et al. for acquiring an entire unit task in routine design (approximately 5–30 s) [92]. Acquisition was typically brief (M: 2.5–7.4 s; Med: 1–3 s), which is consistent with acquisition being operationalised here as a phase embedded within partially pre-defined CAD goals supported by external representations (technical drawings), rather than as acquisition of a completely new unit task (e.g. defining design characteristics and properties of a technical system) in the user's mind. Execution durations align with Card et al.'s expectation that execution rarely exceeds 20 s; M execution durations remained in the 5–8 s range (Med: 4–6 s), supporting the interpretation that longer tasks were decomposed into smaller unit tasks [92]. Inspection episodes were consistently the shortest component (Med: 1–2 s; M: 1.3–4.1 s), but tended to be longer in the orthographic condition, especially for the Average and Low performers, which is consistent with increased verification demands when CAD modelling from orthographic projections.

3.4.4.2 Subtask (CAD operation) clusters

At the level of subtask clusters (CAD operations), effects of projection format are visible both in how often CAD operations occur (Table 41) and in how CAD modelling time is allocated across CAD operations (Table 40). Across the three performance profiles, *dimensioning*, *navigating*, and *drawing* consistently take the larger share of CAD modelling time, but orthographic projections reshape this allocation toward more *navigating* and *manipulating a view*. The direction of this effect of the orthographic projection is consistent, while its magnitude depends on performance profiles.

Table 40 shows that the orthographic projection tends to shift CAD modelling time away from *dimensioning* and toward *navigating/manipulating a view*. This shift is most pronounced for the Average and Low performers. For the Average performer, orthographic projection reduces the *dimensioning* share from 23.0% to 14.1% while increasing *manipulating a view* from 1.6% to 9.0%; *navigating* changes only slightly (16.9% to 17.7%), and *drawing* decreases (11.1% to 8.5%). For the Low performer, the redistribution is even clearer; *navigating* increases from 15.8% to 23.8% and *manipulating a view* from 6.4% to 14.1%, while dimensioning decreases (22.7% to 16.2%). In contrast, the High performer shows a comparatively stable allocation across projections, with only modest shifts (e.g. *manipulating a view* increases from 4.1% to 9.2%; *dimensioning* decreases from 20.1% to 18.7%; *navigating* decreases from 14.9% to 13.4%).

Table 40 Relative duration of subtask clusters (CAD operations) across task conditions and performance profiles

Performance profile	Navigating [%]		Referencing [%]		Drawing [%]		Constraining [%]		Dimensioning [%]		Manipulating [%]		Deleting [%]	
	Iso	Ortho	Iso	Ortho	Iso	Ortho	Iso	Ortho	Iso	Ortho	Iso	Ortho	Iso	Ortho
High	14.9	13.4	1.7	2.0	13.7	12.1	3.4	5.4	20.1	18.7	4.1	9.2	0.0	0.0
Average	16.9	17.1	1.8	2.1	11.1	8.5	0.0	0.6	23.0	14.1	1.6	9.0	0.0	0.5
Low	15.8	23.8	1.3	2.3	11.4	13.1	0.8	0.0	22.7	16.2	6.4	14.1	0.0	0.8

The duration shifts are mirrored and clarified by occurrence counts (Table 41). Orthographic projection increased *navigating* and *manipulating a view* occurrences for all three performance profiles, but with different magnitudes; *manipulating a view* increases from 11 to 63 for the Average performer and 21 to 63 for the Low performer, compared to 8 to 16 for the High performer; *navigating* increases from 85 to 113 (Average) and 60 to 130 (Low), while remaining essentially unchanged for the High performer (54 to 51). Moreover, orthographic projection introduces corrective operation (*deleting*) for the Average (0 to 2) and Low (0 to 3) performers. Together with the time redistribution, this observation indicates that orthographic projection does not simply add more work, but shifts the CAD modelling process toward searching, re-orienting, and iteratively modifying the model while translating multi-view 2D information into a 3D model.

Table 41 Occurrences of subtask clusters (CAD operations) across task conditions and performance profiles

Performance profile	Navigating		Referencing		Drawing		Constraining		Dimensioning		Manipulating		Deleting	
	Iso	**Ortho**	**Iso**	**Ortho**	**Iso**	**Ortho**	**Iso**	**Ortho**	**Iso**	**Ortho**	**Iso**	**Ortho**	**Iso**	**Ortho**
High	54	51	6	6	14	24	6	5	33	26	8	16	0	0
Average	85	113	7	15	27	25	0	2	42	44	11	63	0	2
Low	60	130	5	13	22	39	1	0	31	40	21	63	0	3

These results can be interpreted through the two-dimensional view of CL described in the literature review (Section 2.4.1). Specifically, one component is task-centred (mental load/effective load/intrinsic load) and stems from the inherent demands of the CAD modelling problem, while the other is avoidable (mental effort/ineffective load/extraneous load) and depends on representation, interface, and strategy choices [13], [102], [107]. In CAD terms, the task-centred component is largely fixed here by the required geometry and CAD-feature structure [109], whereas the projection format primarily modulates the avoidable component by shaping how much additional information integration, visual search, and context re-establishment is required [13]. Orthographic projections distribute design information across multiple 2D views, increasing cross-referencing and mental integration demands needed to maintain a coherent mental 3D model in the WM. When these demands approach WM limits, engineering designers may compensate by externalising portion of the integration through interaction with the CAD system; this is operationalised through frequent navigation and view manipulations to re-orient, verify, and reduce what must be held internally, rather than progressing primarily through CAD operations that directly instantiate the 3D model geometry [7]. Occurrences and distribution of subtask clusters align with this interpretation; CAD modelling from the orthographic projection disproportionately increases *navigating* and *manipulating a view* activity and introduces corrective CAD operations for the Average and Low performers, indicating a larger rise in avoidable (extraneous/ineffective) CL for these performance profiles. The High performer shows only modest rebalancing and mainly structural micro-adjustments, consistent with greater proficiency in constructing and maintaining in the

WM an internal representation from multi-view projections (due to schematic organisation of chunks) and planning the CAD modelling sequence with fewer view manipulations and corrective operations [35].

3.4.4.3 CAD actions

At the CAD action layer, the two projection formats are compared in terms of interaction density (number of low-level CAD actions) and the typical time per each CAD action. This layer captures how intensively the engineering designer interacts with the interface (mouse/keyboard inputs) while executing the same CAD modelling goal. Across the three performance profiles, orthographic projections increase CAD action counts for the Average and Low performers, whereas the High performer remains largely stable in basic CAD action counts (Table 42).

Table 42 Number of occurrences (N) and mean (M) duration of CAD actions across task conditions and performance profiles

Performance profile	High				Average				Low			
	Iso		Ortho		Iso		Ortho		Iso		Ortho	
CAD action	**N**	**M [s]**	**N**	**M [s]**	**N**	**M [s]**	**N**	**M [s]**	**N**	**M [s]**	**N**	**M [s]**
Moving a cursor	402	0.952	274	0.909	402	0.92	661	0.95	275	0.88	543	0.86
Clicking on an icon/feature/textbox	235	0.001	173	0.001	235	0.001	328	0.001	180	0.001	312	0.001
Hovering over an icon/feature/point	276	1.01	185	0.714	276	1.28	376	1.03	157	0.40	299	0.46
Typing a value	33	0.185	21	0.143	33	0.27	32	0.16	25	0.58	26	0.19
Rotating a view	7	1	15	3.87	7	3.71	47	2.11	6	2	27	2.30
Zooming in/out	21	0.0769	10	0.1	21	0.29	58	0.19	30	0.2	91	0.12
Dragging a geometric entity	1	2	5	2	1	0	-	-	-	-	1	1
Entering (striking) a command	-	0.001	1	0.001	-	-	18	0.001	12	0.001	26	0.001

CAD modelling from orthographic projections increased the overall number of CAD actions by 55.9% for the Average performer, 93.4% for the Low performer, and only 3.8% for the High performer (Table 42). For the Average and Low performers, the largest increases were in cursor movements, clicking, hovering, rotating, and zooming, i.e., many additional short CAD actions consistent with repeated re-orientation and verification while CAD modelling from the orthographic projection. In contrast, the High performer shows almost constant counts of cursor movements and clicks, but executes fewer, more deliberate viewpoint adjustments; rotations are only slightly more frequent (13 vs. 15), yet their mean duration increases markedly (1.00 s vs. 3.87 s), indicating longer, purposeful rotations rather than many brief ones (Table 42).

Overall, the direction of the effect of projection format is consistent in the sense that orthographic projections demand more view management and behaviour related to visual scanning, but the magnitude depends on the performance profile; it is negligible for the High performer, large for the Average performer, and largest for the Low performer (Table 42).

3.4.4.4 Operator utilisations

Segmentation of CAD modelling processes to the level of CAD actions and their mapping to perceptual (P), cognitive (C), and motor (M) operators enabled a time-based characterisation of how processing resources were allocated across CAD modelling segments. Utilisation provides a critical interpretive lens; it indicates whether behavioural changes related to projection format correspond mainly to increased perceptual and cognitive processing (P/C) or increased motor execution (M), and it reveals whether the same projection format elicits different strategies across output-based performance profiles.

Across the three performance profiles, the overall utilisation patterns are largely preserved when moving from isometric to orthographic projections; the same subtask clusters remain relatively more perceptually and cognitively intensive (P/C) or more execution intensive (M), and the rank ordering of "higher-demand" versus "lower-demand" subtasks is broadly stable. Most notably, *manipulating a view* remains the most demanding subtask cluster in both conditions and for all three performance profiles. This stability implies that, regardless of projection format, *manipulating a view* is inherently characterised by sustained P/C processing interleaved with continuous M control.

A second subtask cluster of relative invariances is *dimensioning*, particularly for the High performer, where utilisation is essentially unchanged across projections (P=C=0.50 in both conditions; M: 0.53 to 0.54). The Low performer also shows only modest shifts (P/C 0.44→0.47; M 0.62→0.58), suggesting that the information-processing balance of dimensioning remains broadly comparable even when CAD modelling from orthographic projections. The Average performer exhibits a clearer redistribution in dimensioning (P/C: 0.57 to 0.51; M 0.43 to 0.53), but the direction of dominance does not invert; rather, dimensioning remains a mixed activity that engages all systems, with orthographic projections increasing the relative share of motor execution for this performer. Overall, these findings indicate that projection format does not uniformly reorganise utilisation profiles across all subtasks; instead, for several core CAD operations (especially *manipulating* and, for some performance profiles, *dimensioning*), the utilisation patterns are robust to the representation format and appear to reflect stable task demands inherent to CAD.

Importantly, these stable patterns are largely consistent across performance profiles (i.e. they do not depend on whether the participant is a High, Average, or Low performer). Where small differences occur (e.g. orthographic projection slightly increases P/C during *manipulating* more strongly for Average and Low), they mainly reflect magnitude changes rather than qualitative shifts. This suggests that some utilisations represent task-inherent characteristics of CAD operations, relatively independent of both projection condition and performance profile.

In contrast to the relatively stable utilisation patterns in *manipulating* (and partly *dimensioning*), the clearest projection-related differences in operator utilisations were notices in *navigating* and *referencing*, and these differences are not equally uniform across performance profiles.

The most consistent effect of the projection format across the performance profiles appears in *navigating*, where CAD modelling from orthographic projections is associated with a reduction in P and C utilisation for all three performers. P and C decrease from 0.56 to 0.52 for the High performer and from 0.57 to 0.43 for both the Average and Low performers. This pattern implies that, under orthographic projections, *navigating* segments contain proportionally less time in P/C processing (e.g. shorter visual search/interpretation pauses) and proportionally more time allocated to interaction-driven execution. However, the accompanying motor changes depend on profile: M increases for High (0.67 to 0.76) and Average (0.67 to 71), but decreases for Low (0.90 to 0.81) performer. Thus, the direction of the P/C change is consistent, but the way navigation is enacted differs by performance profile.

The strongest effect of the projection profiles was noticed in *referencing*. For the High performer, *referencing* became more P/C intensive and less motor intensive (P/C: 0.56 to 0.64; M: 0.67 to 0.50), consistent with spending more time interpreting/selecting reference geometry with fewer continuous interaction (M) actions. The Low performer showed an even stronger shift in the same direction (P/C: 0.33 to 0.86; M 0.86 to 0.69), indicating that referencing in orthographic projection condition requires more P/C involvement for this participant, likely reflecting increased interpretive effort when extracting and integrating information from multiple views. In contrast, the Average performer shows the opposite pattern: *referencing* becomes more motor intensive and less P/C intensive (P/C: 0.67 to 0.50; M: 0.37 to 0.58), indicating that it is carried out through more continuous interface interaction (e.g., selecting planes/features, switching tools/views) rather than sustained interpretive pauses.

3.4.4.5 Which process analysis layer is the most informative?

Considering the three process analysis layers (unit task operations, CAD operations, and CAD actions), the CAD operation layer (subtask clusters) provides the most informative discrimination between projection conditions. At this level, CAD modelling from orthographic projections yields the clearest and most systematic redistribution of behaviour (most visibly for the Average and Low performers) toward *navigating* and especially *manipulating a view*, alongside the emergence of corrective CAD operations (e.g. *deleting*) that are absent or negligible when CAD modelling from isometric projections. In contrast, the High performer shows comparatively smaller rebalancing across subtask clusters. The CAD action layer also discriminates conditions strongly for the Average and Low performers (increases in *rotate* and *zoom in/out* and interaction density), but it is less robust for the High performer and more sensitive to differences in HCI strategies at the lowest level; therefore it is best interpreted as supportive evidence rather than the primary discriminating process analysis layer.

Beyond the process layers, the most consistent effects of projection formats across performance profiles are visible in unit task phases; the share or *inspection* increases in the orthographic projection condition for all three performers (approximately 8.8% to 9.7 %), indicating that CAD modelling from the orthographic projection reallocated time toward verification/checking regardless of the output-based performance profile.

Operator utilisations explain the observed behavioural shifts in information-processing terms; namely, whether additional time in CAD modelling from orthographic projections is expressed primarily as increased motor execution (M; more continuous interface interaction) or increased perceptual and cognitive processing (P/C; more visual interpretation, evaluation, and checking). Importantly, utilisation also revealed strategies that are not always visible in counts alone, most notably in *referencing*, where the direction and magnitude of P/C versus M reallocation differs across performance profiles. In this sense, utilisations reinforce the subtask clusters as the most informative process analysis level, because utilisation effects become most interpretable and comparable when attached to functional CAD operations. Utilisation at the level of unit tasks remain valuable primarily for describing within-participant temporal dynamics (e.g. more frequent switching between P/C-dominant and M-dominant unit tasks in the orthographic projection condition).

Finally, the thesis aims to use utilisations as indicators of CL changes throughout the CAD modelling. However, they reflect when P, C, and M resources were employed, while not implying the magnitude of processing demands which is assumed to differ across CAD task segments. Based on the literature (Section 2.4.3, Table 18), such differences in processing demands might be tested by using EEG to compare engineering designers' brain activity CAD task segments. To confirm the suitability of EEG for capturing and detecting differences in engineering designers' brain activity when CAD modelling, the next section presents the results of an initial EEG data analysis.

3.5 Differences in engineering designers' brain activity when CAD modelling from isometric and orthographic projections[4]

EEG data analysis concerns differences in engineers' brain activity while CAD modelling from orthographic and isometric projections by dividing the CAD task into two segments. In particular, CAD task segment #1 refers to interpreting a technical system from a visual representation (isometric or orthographic projection). Therefore, CAD task segment #1 represents *acquisition* part of the CAD modelling task. In addition, CAD task segment #2 refers to generating a 3D CAD model based on the interpreted projections, thus reflecting primarily the *execution* part of the CAD modelling task.

Considering the division in the CAD task segments, the third research question this experimental study aims to answer is divided into the following two questions:

RQ #1.3.1: Is the brain activity of mechanical engineers different when interpreting isometric (condition #1) and orthographic (condition #2) projections in technical drawings?

RQ #1.3.2: Is the brain activity of mechanical engineers different when generating CAD models from isometric (condition #1) and orthographic (condition #2) projections in technical drawings?

It is assumed that the engineering designers' brain activity will be different when interpreting isometric and orthographic projections in technical drawings as well as when generating CAD models from them. The study analyses neural oscillations over the entire cortex (14 electrodes), individual electrodes, cortical hemispheres, and cortical areas to answer research questions and test the assumption. The analysis focuses on theta (4-7 Hz), alpha (8-12 Hz), and beta (13-30 Hz) since previous EEG studies often reported changes in the power of these frequency bands when visually processing sets of visuospatial information, which is common when solving visuospatial tasks such as CAD modelling (see Section 2.4.3).

With reference to the results of the previous studies, an increase in theta and beta frequency bands and a decrease in alpha frequency band power (compared to the baseline) may be expected during both CAD task segments and for both conditions (#1 and #2). Additionally, larger theta and beta power increase and smaller alpha decrease may be expected when interpreting the orthographic projection and generating a 3D CAD model from it (compared to

[4] This section is based on:

Lukačević et al. (2023) Differences in engineers' brain activity when CAD modelling from isometric and orthographic projections. *Scientific Reports*, 13, 9726. doi: 10.1038/s41598-023-36823-9 the condition with the isometric projection). Furthermore, it is assumed that the right hemisphere (RH) will be more activated than the left one as the RH in right-handed human beings seems specialised for processing visuospatial information [198]. For instance, Roberts and Ann Bell reported a larger alpha power decrease in the right than the left parietal area during visuospatial information processing, thus implying the importance of the RH and the frontal cortical area [199]. Regarding the cortical areas, higher activation in frontal theta (increase), frontal beta (increase), and rear alpha (decrease) may be expected during both CAD task segments when using orthographic projections [158], [200]. Such activation is often related to higher mental effort and CL, which is expected when using orthographic projection. Namely, it is assumed that one must allocate additional cognitive resources to mentally manipulate 2D information presented in the three 2D views of the orthographic projection to combine into a mental 3D model of the represented object [201].

3.5.1 EEG data pre-processing

The EEG data processing is presented in Figure 28, and it was conducted in MATLAB using the EEGLAB toolbox [202]. An original script for data processing was developed according to the pipelines described by Li et al. [203], Vieira et al. [153], and Jia et al. [141].

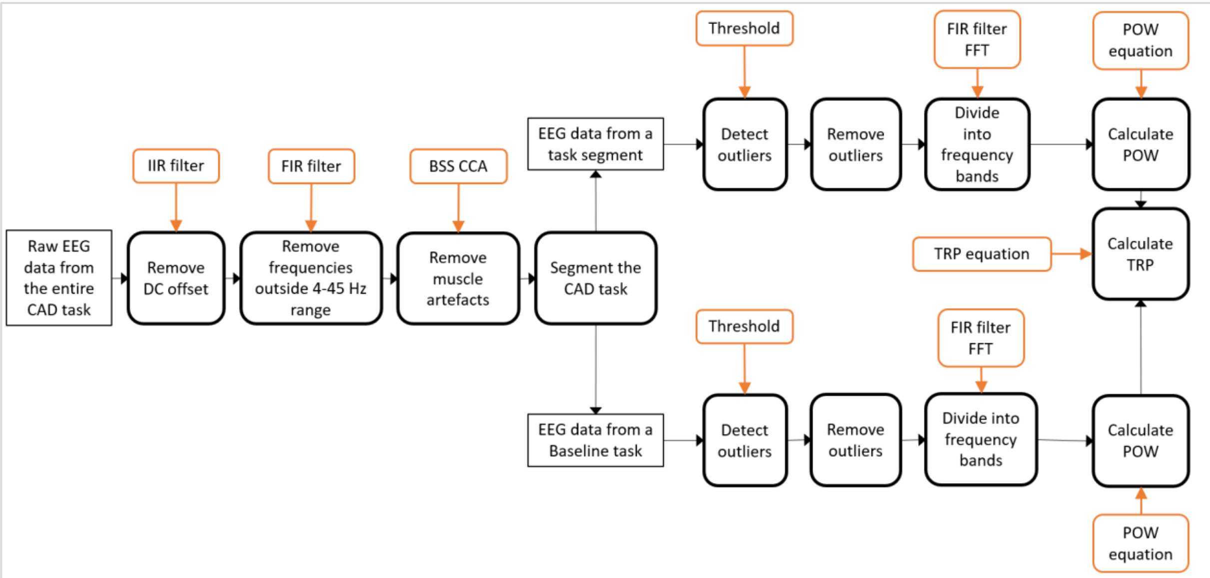

Figure 28 ES#1: EEG data pre-processing pipeline

In the first step, DC offset specific for Emotiv EPOC+ devices was removed with the infinite impulse response (IIR) filter (0.16 Hz first order high-pass filter). Secondly, frequencies outside the 4-45 Hz range were removed with the Finite Impulse Response (FIR) filter. After that, muscle artefacts were removed with the Blind Source Separation (BSS) technique based on Canonical Correlation Analysis (CCA) [204]. The filtering parameters were set as follows:

window length of 2.5 s, window shift of 1.2 s, and four as the number of the least correlated components to be removed. In the next step, the EEG recording was segmented into epochs representing the baselines and the tasks. Furthermore, the CAD tasks were segmented into CAD task segment #1 - interpretation of a technical system from the 2D visual representation (isometric or orthographic) and CAD task segment #2 - generation of a 3D CAD model. The start of the first CAD task segment was defined by the timing of the transition to the slide with 2D representation, and it was derived from the PsychoPy log files. The end of the first CAD task segment corresponds to the start of the second CAD task segment. Namely, the start of the CAD task segment #2 was defined as a moment when the participant started generating the first sketch element (e.g. by drawing the first line) in SolidWorks. An example of segmentation is presented in Figure 29.

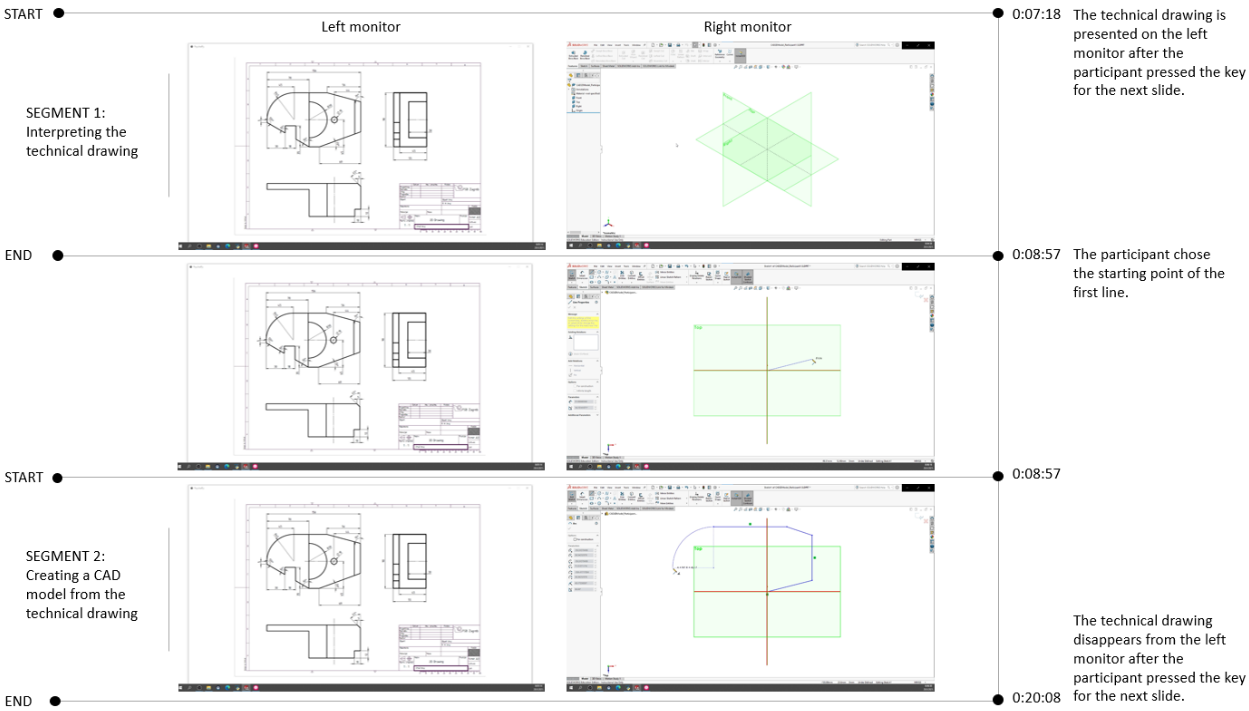

Figure 29 CAD task segmentation example

The segmentation was followed by removing the windows (length of 1 s, shift of 1/128 s) with an amplitude exceeding ±100μV and/or the threshold value calculated for each participant individually. The threshold was calculated as a value three standard deviations greater than the mean (M) of the entire epoch across the electrodes. In this way, any 1 s long epoch of the EEG data with the M amplitude higher than the calculated threshold (or 100μV if the absolute threshold value was above it) was discarded. The percentage of the bad EEG data ranged from 0.19 to 4.02% for the condition #1 (Med = 0.83, MAD = 0.76) and from 0 to 9.92% for the condition #2 (Med = 0.76, MAD = 0.94). In the next step, EEG data was divided into theta (4–7 Hz), alpha (8–12 Hz), and beta (13–30 Hz) sub-frequency bands using the FIR filter. After the threshold was applied, the power of EEG signals (POW) was calculated as the M of the squared values, resulting from the band-pass filtering of the EEG signal and using Fast-Fourier Transformation. The EEGLAB function pop_eegfiltnew, hardcoded to Hamming window, was used for the filtering. In the final pre-processing step, task-related power (TRP) was calculated by subtracting the transformed POW average of a subject j at an electrode i during a baseline task recorded before each CAD task from the transformed POW average of a subject j at an electrode i during a CAD task. Hence, TRP values were calculated according to Equation (32):

$$\mathrm{TRP}_{ij} = \log\left(\mathrm{POW}_i(\mathrm{task})_j\right) - \log\left(\mathrm{POW}_i(\mathrm{baseline})_j\right). \tag{32}$$

Positive TRP values reflect an increase of power during the CAD task (compared to the baseline task), whereas negative TRP values reflect a power decrease [205].

3.5.2 Data analysis

Data analysis was conducted using the R language. Descriptive statistics encompassed the calculation of the Med as a measure of central tendency and MAD as a measure of variability. These parameters were used for data distribution since they are robust to the effects of eventual outliers that potentially persisted after the data pre-processing. Besides, they are more suitable for describing the non-normal distributions (as tested by the Shapiro-Wilk test; $p < 0.05$). In addition, inferential tests enabled the calculation of differences in duration and TRP values between two CAD modelling task segments (interpreting a technical drawing and generating a 3D CAD model). The analysis encompassed a comparison of the tasks (their segments) based on TRP in three frequency bands (theta, alpha, and beta). Data of the two participants from the original sample ($n = 20$) were discarded from the EEG data analysis; one participant reported diagnosed neurological issues, and the other left-handedness.

A nonparametric repeated measures ANOVA approach was adopted [206], [207] to study the differences in TRP values between the projections in each CAD task segment and in each frequency band. Such an approach is based on the Aligned Rank Transform (ART) procedure devised to handle data that violates ANOVA assumptions without inflating the Type I error rates [207]. For each setting (i.e. segment and frequency band), the factors of interest included the projection, electrode, hemisphere (LH and RH), and cortical area (frontal area -FA- and rear area -RA-). The odd-numbered electrodes were grouped under the LH, while the even-numbered ones were under the RH. To compare the TRP values between the cortical areas, the electrodes were distributed as follows: FA: AF3, F7, F3, F4, F8, AF4, FC5, FC6, and RA: O1, O2, P7, P8, T7, T8. The EEG device used in the experiment has good coverage of the FA, but a low spatial resolution in other areas (central, occipital, parietal, and temporal) since only two electrodes are in each of them. Consequently, the division into smaller cortical areas would offer results with a lower statistical rigour.

Significant interaction and main effects detected using the nonparametric repeated measures ANOVA were further decomposed into simple interactions, simple main effects and pairwise comparisons to enable further insights. Herein, the pairwise Wilcoxon signed-rank test with Bonferroni correction was used for the posthoc comparisons. In addition to the (adjusted) p-values, the effect size (reported as r-value) of the Wilcoxon singed-rank test was calculated by dividing the test statistic by the square root of the number of observations. The p-values and the related effect size are, in the following section, coupled with the test statistic values: partial eta squared for the nonparametric factorial ANOVA test and V for the Wilcoxon signed-rank test. Significant differences are presented graphically in the box plots and numerically in the tables.

3.5.3 Results

The first section presents the differences in duration when considering the entire CAD task and its two segments. The second section compares the theta, alpha, and beta TRP values between conditions #1 and #2 for each CAD task segment. First, the significant differences in TRP values and/or large effect sizes are presented for the CAD task segment #1 in the following order. The significant main effects of projection (i.e. differences among projections considering the entire cortex) are discussed first. Then, the significant main effects and interactions concerning the cortical hemispheres are reported. Finally, the main effects and interactions concerning the cortical areas are listed. The same reporting structure is then followed for the CAD task segment #2. Note that, of the explored interactions, only several two-way interactions were found significant and are discussed accordingly (whereas non-significant results are omitted).

3.5.3.1 Duration

The average duration of the CAD task performance (expressed in seconds) was higher when using the orthographic (Med = 991.04, MAD = 405.29) than the isometric projection (Med = 886.44, MAD = 360.26), as shown in Figure 30 (panel a). However, the difference between the two conditions was not statistically significant.

The completion time of the first CAD task segment (interpreting the 2D visual representations) was significantly longer when interpreting the orthographic (Med = 71.5, MAD = 37.81) than the isometric (Med = 32, MAD = 12.6) projection, with V = 170, $p = 1.53 \cdot 10^{-5}$, and r = 0.87.

Participants spent similar time generating CAD models from the orthographic (Med = 904, MAD = 473.69) and the isometric (Med = 861, MAD = 360.27) projections.

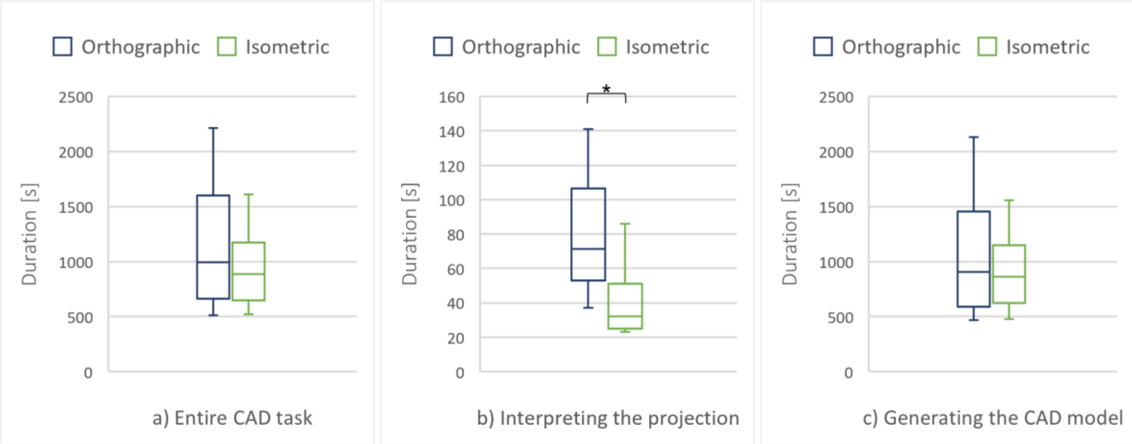

Figure 30 Duration of the a) entire CAD task, b) first CAD task segment, and c) second CAD task segment

3.5.3.2 Theta, alpha, and beta TRP

3.5.3.2.1 CAD task segment 1: Interpreting the projections

The nonparametric repeated measures ANOVA revealed a significant main effect of the projection on the TRPs in all three frequency bands. For completeness, the differences in theta, alpha, and beta TRPs over the entire cortex when interpreting the orthographic and isometric projection were also assessed using the Wilcoxon signed rank test. Figure 31 and Table 43 give further details on these differences.

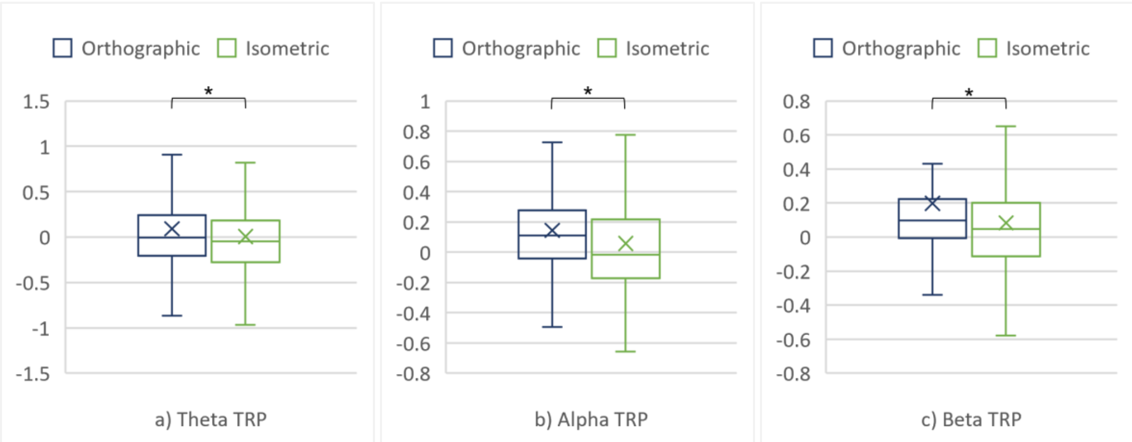

Figure 31 a) Theta TRP, b) Alpha TRP, and c) Beta TRP over the cortex when interpreting the projection

Table 43 Comparison of the TRPs over the cortex when interpreting the projections

Frequency band	Projection	Med	MAD	ART ANOVA		Wilcoxon	
				F, p	ηp2	V, p	r
Theta TRP	Orthographic	$-8.32 \cdot 10^{-3}$	0.33	5.61, $1.82 \cdot 10^{-2}$	$1.21 \cdot 10^{-2}$	13426, $3.00 \cdot 10^{-2}$	0.14
	Isometric	$-4.61 \cdot 10^{-2}$	0.32				
Alpha TRP	Orthographic	0.11	0.24	22.09, $3.44 \cdot 10^{-6}$	$4.59 \cdot 10^{-2}$	12188, $1.20 \cdot 10^{-3}$	0.2
	Isometric	$-1.81 \cdot 10^{-2}$	0.27				
Beta TRP	Orthographic	$9.65 \cdot 10^{-2}$	0.16	11.81, $6.44 \cdot 10^{-4}$	$2.51 \cdot 10^{-2}$	12864, $7.94 \cdot 10^{-3}$	0.17
	Isometric	$4.88 \cdot 10^{-2}$	0.23				

The main effect of the cortical hemisphere was found on the theta and alpha TRPs (Table 44).

Table 44 The main effect of the cortical hemisphere

Frequency band	Cortical hemisphere	Med	MAD	ART ANOVA		Wilcoxon	
				F, p	**ηp²**	**V, p**	**r**
Theta TRP	LH	$5.98 \cdot 10^{-2}$	0.31	$47.68,\ 1.68 \cdot 10^{-11}$	$9.41 \cdot 10^{-2}$	$23412,\ 1.68 \cdot 10^{-11}$	0.41
	RH	-0.13	0.31				
Alpha TRP	LH	$6.99 \cdot 10^{-2}$	0.25	$6.96,\ 8.61 \cdot 10^{-3}$	$1.49 \cdot 10^{-2}$	$19673,\ 1.27 \cdot 10^{-3}$	0.2
	RH	$3.91 \cdot 10^{-2}$	0.28				

No significant interaction effects between hemisphere and the remaining factors were found in the three frequency bands. Nevertheless, pairwise Wilcoxon signed rank tests with Bonferroni correction revealed a significant difference in alpha TRP between the projections over the RH. In addition, alpha TRP significantly differed between the hemispheres, but only for the condition with the isometric projection, whereas significant differences in theta TRP were found between hemispheres when interpreting both projections (see Figure 32 and Table 45).

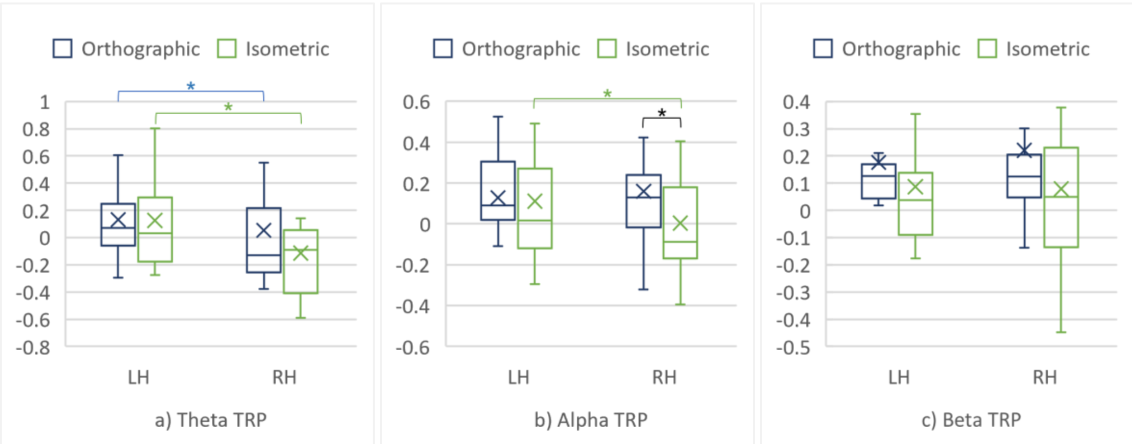

Figure 32 a) Theta TRP, b) Alpha TRP, and c) Beta TRP over the cortical hemispheres when interpreting the projections

Table 45 Significant differences in TRP regarding the cortical hemispheres and projections

Frequency band	Projection	Cortical hemisphere	Med	MAD	Wilcoxon	
					V; p	**r**
Theta TRP	Orthographic	LH	$6.49 \cdot 10^{-2}$	0.27	$6373;\ 7.68 \cdot 10^{-9}$	0.3
		RH	-0.12	0.28		
	Isometric	LH	$4.37 \cdot 10^{-2}$	0.34	$5404;\ 6.36 \cdot 10^{-4}$	0.52
		RH	-0.17	0.35		
Alpha TRP	Isometric	LH	$2.22 \cdot 10^{-2}$	0.25	$5713;\ 3.07 \cdot 10^{-5}$	0.37
		RH	$-7.02 \cdot 10^{-2}$	0.27		
Alpha TRP	Orthographic	RH	0.12	0.24	$2964;\ 1.20 \cdot 10^{-2}$	0.26
	Isometric		$-7.02 \cdot 10^{-2}$	0.27		

No main effects of the cortical area were found in the TRP values of theta, alpha, and beta when interpreting the projections. However, there was a statistically significant interaction between the projection and the cortical area in the theta frequency band ($F = 7.56$, $p = 6.22 \cdot 10^{-3}$, and ηp2 $= 1.62 \cdot 10^{-2}$). As shown in Figure 33 (panel a), further analysis of the simple main effects of projection showed significant differences between the theta TRP over the FA, with $V = 3603$, $p = 1\cdot10^{-3}$, and $r = 0.13$.

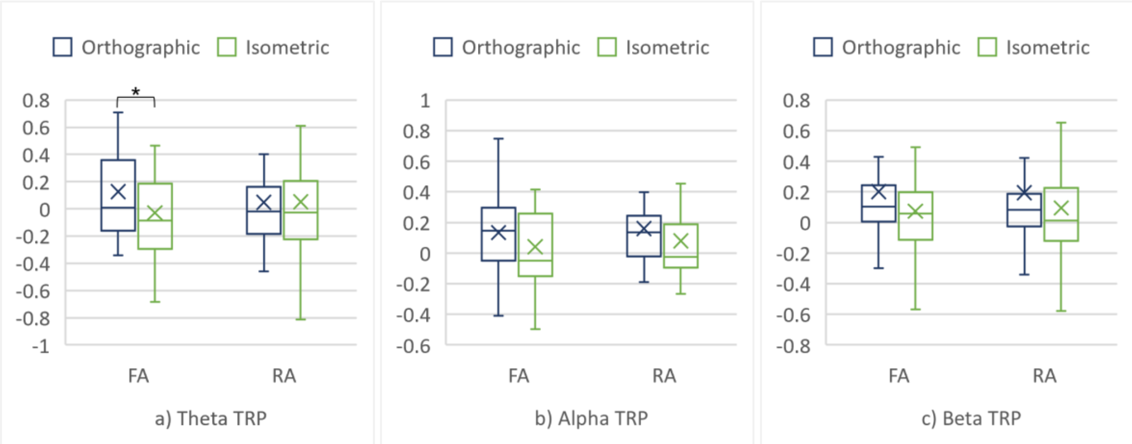

Figure 33 a) Theta TRP, b) Alpha TRP, and c) Beta TRP over the cortical areas when interpreting the projections

3.5.3.2.2 CAD task segment 2: Generating the CAD models

The main effect of the projection on the TRPs over the entire cortex was found in all three frequency bands when generating the CAD models. As presented in Figure 34 and Table 46, theta, alpha, and beta TRPs over the entire cortex significantly differed when generating the CAD models from the orthographic and isometric projections.

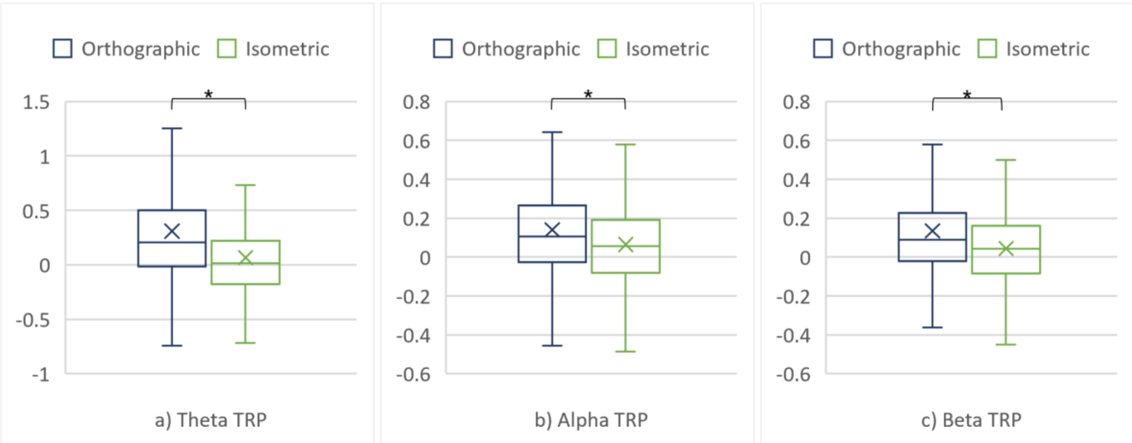

Figure 34 a) Theta TRP, b) Alpha TRP, and c) Beta TRP over the cortex when generating the 3D CAD models

Table 46 Comparison of the TRPs over the cortex when generating the 3D CAD models

Frequency band	Projection	Med	MAD	ART ANOVA		Wilcoxon	
				F; p	ηp^2	V; p	r
Theta TRP	Orthographic	0.2	0.37	61.40; $3.27\cdot10^{-14}$	0.12	13426; $3.00\cdot10^{-2}$	0.14
	Isometric	$1.20\cdot10^{-2}$	0.29				
Alpha TRP	Orthographic	0.11	0.22	9.74; $1.92\cdot10^{-3}$	$2.08\cdot10^{-2}$	12188; $1.20\cdot10^{-2}$	0.20
	Isometric	$5.62\cdot10^{-2}$	0.20				
Beta TRP	Orthographic	$9.01\cdot10^{-2}$	0.18	16.73; $5.10\cdot10^{-5}$	$3.52\cdot10^{-2}$	12864; $7.94\cdot10^{-3}$	0.17
	Isometric	$4.16\cdot10^{-2}$	0.18				

The main effect of the cortical hemisphere was found on the TRPs in all three frequency bands; details are available in Table 47.

Table 47 The main effect of the cortical hemisphere

Frequency band	Hemisphere	Med	MAD	ART ANOVA		Wilcoxon	
				F; p	ηp²	V; p	r
Theta TRP	LH	0.16	0.36	29.45; $9.31\cdot10^{-8}$	$6.03\cdot10^{-2}$	22141; $8.59\cdot10^{-8}$	0.34
	RH	$3.58\cdot10^{-2}$	0.31				
Alpha TRP	LH	0.12	0.21	10.92; $1.03\cdot10^{-3}$	$2.32\cdot10^{-2}$	20656; $4.66\cdot10^{-4}$	0.26
	RH	$3.85\cdot10^{-2}$	0.22				
Beta TRP	LH	$8.19\cdot10^{-2}$	0.20	6.49; $1.12\cdot10^{-2}$	$3.89\cdot10^{-3}$	20420; $1.10\cdot10^{-4}$	0.24
	RH	$4.46\cdot10^{-2}$	0.17				

Furthermore, the interaction effect of the projection and cortical hemisphere was found in theta ($F = 5.14$, $p = 2.38\cdot10^{-2}$, and $\eta p2 = 1.11\cdot10^{-2}$) and alpha ($F=3.73$, $p=0.05$, and $\eta p2 = 8.05\cdot10^{-3}$) frequency bands. Decomposing the two-way interaction into simple main effects revealed the significant differences between the projections in the theta TRPs over both hemispheres. In other words, the simple main effect of the projection was significant both in the LH ($V = 2492$, $p=2.41\cdot10^{-4}$, $r = 0.33$) and the RH ($V = 1414$, $p = 3.05\cdot10^{-10}$, $r = 0.56$), as presented in Figure 35 (panel a). In addition, simple main effect of the cortical hemisphere was significant in the condition with the isometric projection ($V = 6511$, $p = 9.88\cdot10^{-2}$, and $r = 0.55$).

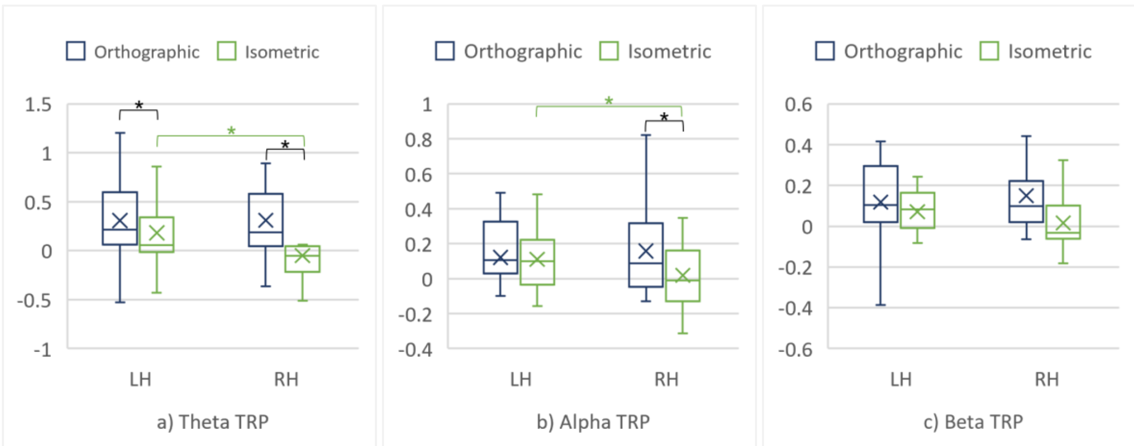

Figure 35 a) Theta TRP, b) Alpha TRP, and c) Beta TRP over the cortical hemispheres when generating the 3D CAD models

At the electrode level, difference between the LH and the RH was found in all electrodes for the isometric projection condition. These electrodes are presented in Table 48.

Table 48 Difference between the LH and the RH in theta TRP over the individual electrodes

Projection	Electrode	LH		RH		LH vs. RH	
		Med	MAD	Med	MAD	V; p	r
Isometric	AF3 | AF4	0.17	0.24	$-7.16\cdot10^{-2}$	0.29	165; $1.07\cdot10^{-4}$	0.43
	F3 | F4	$4.71\cdot10^{-2}$	0.29	$-7.80\cdot10^{-2}$	0.13	169; $2.29\cdot10^{-5}$	0.21
	F7 | F8	$3.84\cdot10^{-2}$	0.24	-0.13	0.22	154; $2.00\cdot10^{-3}$	0.38
	FC5 | FC6	0.14	0.40	-0.11	0.29	168; $3.81\cdot10^{-5}$	0.42
	T7 | T8	0.21	0.38	$-9.98\cdot10^{-2}$	0.33	152; $2.00\cdot10^{-3}$	0.37
	P7 | P8	$5.65\cdot10^{-2}$	0.43	$-4.04\cdot10^{-2}$	0.32	153; $2.00\cdot10^{-3}$	0.22
	O1 | O2	0.15	0.34	$1.44\cdot10^{-2}$	0.11	132; $4.30\cdot10^{-2}$	$9.49\cdot10^{-2}$

Additionally, the interaction effect of the projection and cortical hemisphere area was found for alpha frequency band. Decomposing this interaction into simple main effect of the projection revealed the differences in alpha TRP over the RH, as presented in Figure 32 (panel b), with $V=2964$, $p = 1.2\cdot10^{-2}$, $r = 0.23$. Furthermore, the simple main effect of the cortical hemisphere was found on alpha TRP for the condition with the isometric projection, with $V = 5713$, $p = 3.07\cdot10^{-5}$, $r = 0.37$. Finally, significant differences in alpha TRP between the LH and the RH were found at several electrodes presented in Table 49.

Table 49 Difference between the LH and the RH in alpha TRP over the individual electrodes

Projection	Electrodes	LH		RH		LH vs. RH	
		Med	MAD	Med	MAD	V; p	r
Isometric	AF3 | AF4	0.13	0.23	$3.78\cdot10^{-2}$	0.22	139; $1.80\cdot10^{-2}$	0.20
	F3 | F4	0.12	0.14	$4.84\cdot10^{-2}$	0.17	132; $4.30\cdot10^{-2}$	$8.44\cdot10^{-2}$
	F7 | F8	0.12	0.24	$-2.65\cdot10^{-2}$	$8.97\cdot10^{-2}$	135; $3.00\cdot10^{-2}$	0.25
	T7 | T8	0.14	0.20	$1.80\cdot10^{-2}$	0.27	133; $3.80\cdot10^{-2}$	0.23
Orthographic	T7 | T8	$9.16\cdot10^{-2}$	0.14	$5.47\cdot10^{-2}$	0.16	139; $1.80\cdot10^{-2}$	0.18

The main effect of the cortical area was found on theta TRP, with $F = 13.88$, $p = 2.19\cdot10^{-4}$, and $\eta p2 = 2.94\cdot10^{-2}$. Furthermore, the interaction effect of the projection and cortical area was found in the same (theta) frequency band ($F = 9.99$, $p = 1.67\cdot10^{-3}$, and $\eta p2 = 2.13\cdot10^{-2}$). Further decomposing the interaction into the simple main effects revealed significant differences in theta TRP over both the FA and the RA when compared between the projections (see Figure 36 and Table 50). In addition, the simple main effect of the cortical area was found on theta TRP when generating the CAD model from the orthographic projection.

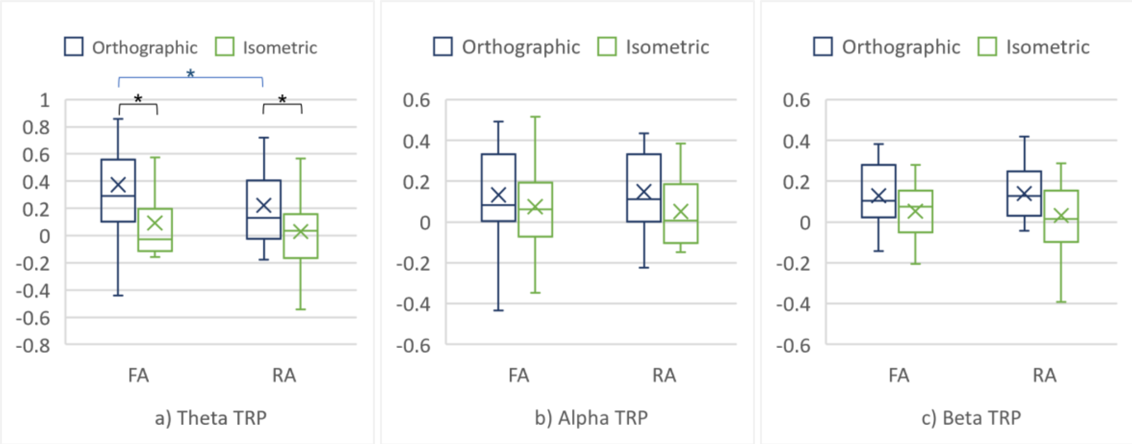

Figure 36 a) Theta TRP, b) Alpha TRP, and c) Beta TRP over the cortical areas when generating the 3D CAD models

Table 50 Simple main effects on theta TRP

Cortical area	Projection	Med	MAD	Wilcoxon	
				V; p	r
FA	Orthographic	0.31	0.42	1921; $4.76 \cdot 10^{-11}$	0.33
	Isometric	$7.25 \cdot 10^{-3}$	0.29		
RA	Orthographic	$9.88 \cdot 10^{-2}$	0.30	1937; $2.00 \cdot 10^{-3}$	0.17
	Isometric	$3.47 \cdot 10^{-2}$	0.31		
FA	Orthographic	0.31	0.42	9608; $1.00 \cdot 10^{-3}$	0.20
RA		$9.88 \cdot 10^{-3}$	0.30		

For completeness, the interaction among projection and individual electrodes (encoded as: AF3, AF4, F3, F4, F7, F8, FC5, FC6, P7, P8, T7, T8, O1 and O2) was studied. No significant interaction between projection and individual electrodes factors was found in any of the bands. Nevertheless, pairwise Wilcoxon signed rank tests with Bonferroni correction identified several electrodes at which theta TRP differs significantly between the projections (see Figure 37 and Table 51).

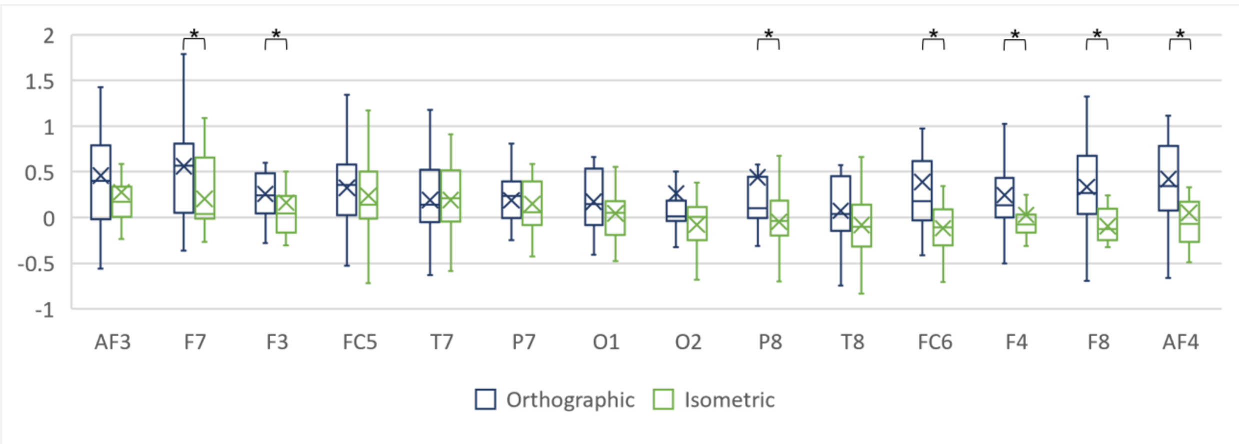

Figure 37 Theta TRP at the individual electrodes when generating the CAD models

Table 51 Theta TRP at and among the individual electrodes

Electrode	Orthographic		Isometric		Orthographic vs. Isometric	
	Med	**MAD**	**Med**	**MAD**	**V; p**	**r**
F7	0.56	0.56	$3.84 \cdot 10^{-2}$	0.24	37; $3.40 \cdot 10^{-2}$	0.24
F3	0.24	0.30	$4.71 \cdot 10^{-2}$	0.29	19; $2.00 \cdot 10^{-8}$	0.26
P8	0.10	0.23	$-4.04 \cdot 10^{-2}$	0.32	35; $2.70 \cdot 10^{-2}$	0.31
FC6	0.18	0.42	-0.11	0.29	21; $3.00 \cdot 10^{-3}$	0.66
F4	0.13	0.26	$-7.80 \cdot 10^{-2}$	0.13	28; $1.00 \cdot 10^{-2}$	0.42
F8	0.27	0.34	-0.13	0.22	16; $1.00 \cdot 10^{-3}$	0.52
AF4	0.34	0.43	$-7.16 \cdot 10^{-2}$	0.29	21; $3.00 \cdot 10^{-8}$	0.46

3.5.4 Discussion

Engineers' brain activity was captured and analysed during the CAD modelling tasks in two conditions; when technical systems were presented with isometric (condition #1) and orthographic (condition #2) projections in technical drawings. An increase in theta and beta TRP (compared to the baseline) was expected in both CAD task segments (interpreting the projections in technical drawings as segment #1 and generating the 3D CAD models from them as segment #2) since it has often been related to cognitive processing of visuospatial information [148], [155]–[157]. In line with that assumption, beta TRP increased in both CAD task segments. However, theta TRP increased when generating the 3D CAD models while surprisingly decreased when interpreting the technical drawings (for both conditions). In addition, alpha TRP decreased only when interpreting the isometric projection. A decrease of alpha TRP was expected in both CAD task segments and for both conditions since such a response has often been observed when processing visuospatial information [159], [160]. Moreover, it is generally considered that the RH is specialised for processing visuospatial information in right-handed human beings [198]. Since the tasks used in the experiment are visuospatially-intensive and all the participants included in the analysis were right-handed, higher activation over the RH was expected when performing the CAD tasks. However, the TRP values over the RH were either similar to those of the LH (theta TRP during the first CAD task segment) or lower (in all the other cases).

A possible explanation for unexpected brain activity may be related to cognitive characteristics of the CAD tasks and visuospatial information processing in the CAD context. In addition to that argument, Willis et al. suggested that brain activity in visuospatial information processing highly depends on the task requirements posed to the participants (what they should do with presented information) and not only the type of information that should be processed [208]. For example, several studies have reported increased alpha TRP while solving standardized tests related to aspects of visuospatial thinking [155], [157], [198]. Hence, it may be that increase in alpha TRP revealed when performing CAD tasks is due to using similar cognitive mechanisms. Similarly, revealed brain activity in individual hemispheres aligns with the results reported by Ornstein et al. [198] and Roberts and Ann Bell [199], who argued that higher activation of the LH may be caused by the analytic strategies specific for the mental rotation tasks. Both mental rotation and interpretation of the visual representation ask for a visual transformation as an aspect of visuospatial thinking (according to Shah et al. [11]). Hence, it may be that interpretation of visual representations causes higher activation of the LH because of the underlying visual transformation required to conduct it. Furthermore, the dimensionality of visual representation used in the mental rotation task seems to affect alpha TRP in the cortical hemispheres. Namely, the 2D mental rotation task was previously associated with the higher parietal activation in the LH than the RH. On the contrary, the activation was higher in the right than the left parietal area during the 3D mental rotation task [199]. It is yet to be explored why the results diverge across the studies that used the same standardised tests and what it means for the CAD context.

The study assumed differences in engineering designers' brain activity when interpreting isometric and orthographic projections in technical drawings as well as when generating CAD models from them. EEG results revealed statistical significance and considerable effect sizes in both CAD task segments when analysed across the cortex, cortical hemispheres, cortical areas, and individual electrodes. Thus, both research questions are answered positively, and the assumption is confirmed; the brain activity of engineering designers was different when interpreting isometric and orthographic projections in technical drawings, as well as when generating 3D CAD models from them. The discussion of these differences is detailed for each segment in the following sections.

3.5.4.1 Differences in engineering designers' brain activity when interpreting the projections

The analysis of the TRP values in all three frequency bands suggested statistically significant differences in brain activity when interpreting the orthographic and isometric projections. In particular, a decrease of theta TRP was smaller when interpreting the orthographic projection (as compared to the baseline). Furthermore, alpha TRP increased when interpreting the orthographic projection while decreased when using the isometric one. Finally, an increase in beta TRP was larger for the orthographic than the isometric projection. The smaller decrease or larger increase in theta and beta TRP confirm the assumption of using more cognitive resources to interpret the technical drawing in which the technical system is presented with the orthographic projection (due to the need to combine three 2D views to a mental 3D model). Such results are in line with previous studies (e.g. the work of Fajen and Philips [201]). Alpha, on the other hand, behaved differently from what was expected; there was more alpha TRP in condition #2 although it was expected to see alpha TRP decrease with increasing processing demands (imposed by the orthographic projection) [159], [160].

Moreover, engineering designers used the cortical hemispheres similarly when interpreting both projections since no significant differences in the TRP values were found between the hemispheres when considering the projections. However, alpha TRP suggests different activation of the RH when interpreting the isometric than the orthographic projection. In addition, a significant difference in alpha TRP was found between the LH and the RH when interpreting the projections. Similarly, significant differences were found in theta frequency band when TRPs were compared between the hemispheres. These results imply the asymmetric hemisphere activation when interpreting the projections.

3.5.4.2 Differences in engineering designers' brain activity when generating the 3D CAD models from the projections

Similarly to the first CAD task segment, the TRP values were significantly different when generating the CAD models from the orthographic and isometric projections in all three frequency bands. However, the difference in TRP values for the CAD task segment #2 between the conditions was reflected in a larger increase when using the orthographic than the isometric projection in all three frequency bands.

Furthermore, significant differences in theta TRP were found over the individual electrodes when comparing the conditions (see Figure 37). These electrodes were mainly located in the FA, thus implying the important role of theta over the FA in distinguishing the effects of the projections on engineers' brain activity.

The effect of the projections when generating the CAD models was reflected in the theta and alpha TRPs, which both significantly differed for conditions #1 and #2 over the RH. The difference among the projections in theta TRP was also significant over the LH. When comparing theta TRP values over the hemispheres, it is noticeable that a decrease (with respect to the baseline task) is present over the RH and only for the isometric condition. In addition, the differences in TRP values over the LH and the RH were significant in the theta and alpha frequency band when generating the CAD models from the isometric projection. For this condition (#1), the revealed difference in theta TRP between the hemispheres was corroborated by the significant differences among all the seven pairs of the electrodes (as shown in Table 48). At the level of the electrodes, the effect of the hemispheres on the alpha TRP was noticed mainly over those located in the FA (see Table 49).

Considering the cortical areas, significant differences between the conditions were found in theta TRP over both the RA and the FA. Theta TRP over the FA increases in value for the orthographic and a decreased for the isometric projection. Such results are aligned with previous studies on the visual processing of information. For example, Liu et al. reported high theta TRP over the FA during a mental rotations task [158]. An increase in theta TRP over the FA for condition #2 might be related to higher requirements on attention and the level of the cognitive load imposed by interpreting and generating CAD models from the orthographic projection [160]. Higher alpha TRP in the RA for condition #2 in both CAD task segments suggests similar explanations (although differences were not statistically significant). For instance, Gerlič and Jaušovec related higher FA alpha power with more efficient processing of presented information, and higher alpha activation in the RA (temporal area in particular) with higher CL [200].

Therefore, the analysis revealed significant differences in engineering designer’s brain activity in both examined CAD task segments. At the same time, the behavioural results showed an asymmetry in time-on-task across segments; segment #1 (interpreting the technical drawing) was significantly longer in the orthographic than the isometric projection condition, whereas the time spent generating the 3D geometry in the CAD system (segment #2) was similar across conditions. These duration results imply that the orthographic projection primarily prolongs the initial visual transformation and synthesis required to construct a coherent 3D mental representation from multiple 2D views [11], i.e. the dominant cognitive operation in segment #1.

However, the EEG results discussed above suggest a different emphasis; the projection-related differences in TRP appear more evident during CAD model generation (segment #2) than during initial interpretation (segment #1). This apparent mismatch might be explained by the fact that the two segments differ not only in duration, but also in their information-processing composition and in how frequently WM is likely accessed. Interpreted through the operator-based (P, C, M) lens adopted in this thesis (Section 3.4), segment #1 largely corresponds to an extended acquisition interval dominated by perceptual and cognitive processing (P and C), where the key requirement is to construct, transform, and maintain (in the WM) an internal 3D representation from the technical drawing. Segment #2, in contrast, is characterised by repeated acquisition–execution–inspection cycles, in which cognitive activity, including frequent WM access to retrieve dimensions/feature parameters and to update the evolving mental model, is tightly interleaved with motor execution and perceptual checking as the CAD model is incrementally built and verified.

This distinction provides one possible explanation for why effects of projection formats can be more clearly expressed in segment #2 even when segment #1 is significantly longer in the orthographic projection condition. In practice, EEG-based TRP indicators are typically interpreted as reflecting the momentary intensity of oscillatory engagement, and thus more sensitive to changes in information-processing intensity (i.e. the level of neural engagement per unit time) rather that being a direct function of total time spent in a segment [148], [163], [209], [210]. Consequently, if the orthographic projection prolongs segment #1 mainly by extending acquisition time through additional visual search, re-checking, or slower information synthesis (i.e. adding time without a proportional increase in processing intensity per unit time), then the mean TRP contrast between conditions can be attenuated when averaged over the entire segment. By contrast, in segment #2 the orthographic projection may increase the frequency and tight interleaving of the WM updating and control operations within interactive acquisition, execution and inspection cycles, which can yield more pronounced TRP differences even when overall duration between the conditions is similar.

In the second empirical study, this interpretation can be examined more directly by aligning EEG-based CL indicators to the process representation developed in Section 3.4; specifically to unit task phases (acquisition, execution, inspection), unit task operations, and CAD operation segments, so that effects of conditions can be evaluated at the level where differences in operator composition are explicitly represented.

3.6 Experimental study limitations

As stated in the chapter introduction, the first empirical study primarily served as an initial feasibility test of candidate methods for describing CAD performance from the information-processing lens and for capturing neurocognitive differences during realistic CAD modelling. The goal was to evaluate whether the proposed building blocks (output metrics, process segmentation and operator mapping, and EEG measurement) are practically usable and empirically informative in the CAD context, rather than to provide final, generalisable claims about all effects of the projection format on engineering designers' performance in CAD modelling tasks. Still, several limitations should be noted as a precaution against over-interpreting the findings beyond these aims.

A first limitation concerns the EEG sample size and group balance. The statistical analysis of EEG was conducted on data from 18 participants, which provides an empirical basis for detecting differences related to projection format but remains modest for drawing strong population-level inferences. In addition, both participants whose data were discarded from EEG analysis belonged to Group 2 (as defined in Figure 9), which started CAD modelling from the isometric and then moved to the orthographic projection condition. This created an imbalance between counterbalancing groups, which could have reduced statistical power for detecting some effects and may have contributed to non-significant results in certain comparisons.

A second limitation concerns the spatial resolution and interpretability of EEG results. The EEG device used in the study comprised 14 electrodes, and therefore the analysis did not attempt fine-grained functional localisation of brain activity to specific cortical sources. With relatively low spatial resolution, conclusions were intentionally restricted to TRP differences at the level of the entire cortex (all 14 electrodes), hemispheres, cortical areas, and individual 14 electrodes, without strong anatomical claims, because such localisation can be unreliable with this configuration [211]. Accordingly, the EEG results should be interpreted as evidence of condition-related differences in brain activity (oscillatory responses), rather than as precise statements about activation of specific brain regions.

A third limitation concerns the scope and purpose of the process-based CAD performance analysis. The segmentation and process representation were developed through detailed manual coding of screen recordings for three representative participants (High, Average, and Low performer). This choice reflects the thesis objective at this stage; to derive and demonstrate a transferable hierarchical task structure, coding scheme, and operator-mapping logic that can support consistent comparison across conditions and later alignment with CL indicators. In other words, the process analysis in this chapter is not intended to estimate population-level effect sizes between the conditions; rather, it provides a rigorous working example that makes the CAD modelling process observable and measurable across multiple levels. While previous work has also used small samples for similarly detailed, multi-layer process coding, extending the process-based analysis to more participants is advisable in future work to quantify variability, strengthen generalisability, and test whether the observed process-based differences hold across the broader sample.

Moreover, because the process coding was performed by a single trained researcher, interrater reliability could not be computed for this dataset at this stage. This is a practical constraint common to high-resolution behavioural coding, but it is important for interpretation. To mitigate this limitation, the coding was constrained by 1) an explicit, rule-based segmentation logic, 2) operational definitions for each category, and 3) iterative refinement and consistency checks across three participants and two conditions before finalising each level of the hierarchy. Nevertheless, future work should include a second coder for a subset of the data (or a structured recoding protocol) to establish interrater agreement for the segmentation and operator mapping. In addition, the participant characteristics and task context constrain external validity. The conclusions about performance and neurocognitive responses are specific to the sampled population of young engineering designers (young professionals) and to CAD modelling tasks in which design characteristics are predefined and communicated via technical drawings, rather than being generated through open-ended design synthesis. Likewise, the observed interactions between engineering designers and the CAD system reflect CAD modelling performed with common HCI tools (conventional monitor, keyboard, and mouse) and within a feature-based CAD environment; the proposed segmentation into task/goal levels is therefore most directly applicable to feature-based CAD modelling workflows and may require adaptation for other CAD paradigms or interaction modalities.

Taken together, these limitations do not undermine the purpose of the first empirical study within the thesis. Rather, they clarify the appropriate level of inference; this chapter provides an initial empirical validation of feasibility and interpretability of the candidate methods and demonstrates that projection type can induce detectable differences in both behavioural process structure and EEG responses under realistic CAD modelling conditions. The second empirical study builds on this foundation by scaling up the process-based analysis and strengthening validity checks (including reliability and segment-level alignment between behavioural structure and psychophysiological indicators).

3.7 Conclusions

As framed in the chapter introduction, the first empirical study was designed as an initial empirical step, conducted within Descriptive Study I, to test whether the candidate building blocks required by the thesis (output-based performance metrics, a transparent process representation grounded in human performance modelling, and EEG-based monitoring) are feasible to implement in realistic CAD modelling and sufficiently sensitive to detect differences induced by a theoretically motivated influencing factor (a projection type). Accordingly, the central contribution of this chapter is twofold: 1) it demonstrates that CAD modelling can be decomposed into a transferable hierarchical structure (unit task operations, CAD operations/subtask clusters, and CAD actions), complemented by unit task phases and an explicit P, C, M operator mapping that enables interpretation in information-processing terms;

and 2) it provides initial evidence that EEG recording during CAD modelling is practically feasible and can detect condition-related differences in engineering designers' brain activity. These contributions were established by answering the three research questions stated in the study introduction (Section 3.1). Answering RQ 1.1 showed that CAD modelling outputs differ between isometric and orthographic projections, confirming that projection type is a practically meaningful influencing factor and that the selected output metrics are sensitive enough to register its effects. At the thesis level, this justifies keeping outcome-based measures as one layer of the CAD performance model and provides a baseline for interpreting process differences. Answering RQ 1.2 showed that CAD modelling processes also differ between the projection types and that these differences are most apparent when process is analysed at the CAD operation (subtask cluster) level, thus marking it as the most informative layer for discriminating the conditions in this study. Importantly, the P, C, and M operator utilisation results provide the interpretive link that makes these behavioural differences meaningful in information-processing terms by showing how time is distributed between perceptual, cognitive, and motor demands within the same CAD task segments. At the thesis level, this supports the use of human performance modelling as a complementary approach for interpreting process differences as changes in information-processing demands, thereby justifying the adoption of an operator-based lens as part of the process-oriented CAD performance model and as a bridge to later CL analyses. This in turn motivates the need for mapping rules between operators representing the embodied cognition and CAD task segments that are explicit, transparent, and reproducible. Finally, answering RQ 1.3 showed that engineering designers' brain activity during CAD modelling differs between projection types, demonstrating that EEG can be applied in a task context that closely reflects typical CAD modelling work and can detect changes related to representation format (projection type in particular). At the thesis level, this provides the empirical justification for including EEG as a modality in the CL measurement and analysis method, while also motivating the later work of selecting and validating EEG-based indicators applicable in the CAD context and aligning them with process segments in the second empirical study. These insights are carried forward into the next chapters, where the theoretical models of CAD performance and CL are prescribed (Chapter 4), followed by prescription of the method for measuring and analysing CL (Chapter 5).

4 THEORETICAL MODELS

Chapter 2 showed that both CAD performance and CL are widely discussed in engineering design research, yet the comparison across studies remains difficult because definitions, segmentation schemes, and metrics are inconsistent and often constrained by data availability. In CAD performance research, outcome measures of the final model are essential for assessing effectiveness, but they provide limited insight into how performance unfolds over time; process measures are frequently cumulative summaries, while temporally structured descriptions of CAD modelling behaviour are still relatively rare and use inconsistent terminology (commands, events, actions, operations). In the CL literature, a resource-based view is common, but operationalisations differ, CAD-specific evidence remains narrow, and temporally informative measurement aligned with CAD process dynamics is still scarce. Together, these conclusions motivate a synthesis that provides a transferable structure for describing CAD modelling as a process and makes the information-processing perspective explicit so that CL can later be represented and related to performance at the segment level.

Accordingly, this chapter synthesises the theoretical foundation developed in Chapter 2 and empirical insights gathered in Chapter 3 into two complementary models: 1) a theoretical model of engineering designers' performance in CAD activities and 2) a theoretical model of CL in CAD activities. The models formalise the key elements and relationships required by the thesis objectives; to represent CAD performance considering both outcomes and processes, to structure CAD modelling activity in a way that supports consistent segmentation and interpretation, and to enable later prescription of a method for measuring and analysing CL and relating it to CAD performance.

4.1 Key elements of CAD performance and CL in CAD activities

Both engineering designer's performance and CL in CAD activities are conceptualised as a function of the characteristics and interactions of four elements shown in Figure 38: an engineering designer, a CAD task, a CAD system, and an environment. This framing follows the foundations discussed in Chapter 2. Cognitive architectures and human performance models (in particular the Model Human Processor; MHP [7]) describe performance as emerging from the coupling between a human information processor (with perceptual, cognitive, and motor systems) and an interactive (HCI) system, while cognitive load theory [13] emphasises how task demands and context determine required resources and thereby shape both performance and experienced CL. In this thesis, these four elements are therefore not treated independently; rather, they are used to make explicit how information flows between them, how feedback loops are formed, and how these relationships shape behaviour (e.g. what must be processed, what can be perceived at a given moment, what actions are possible given the interface, and what external conditions affect perception and action execution).

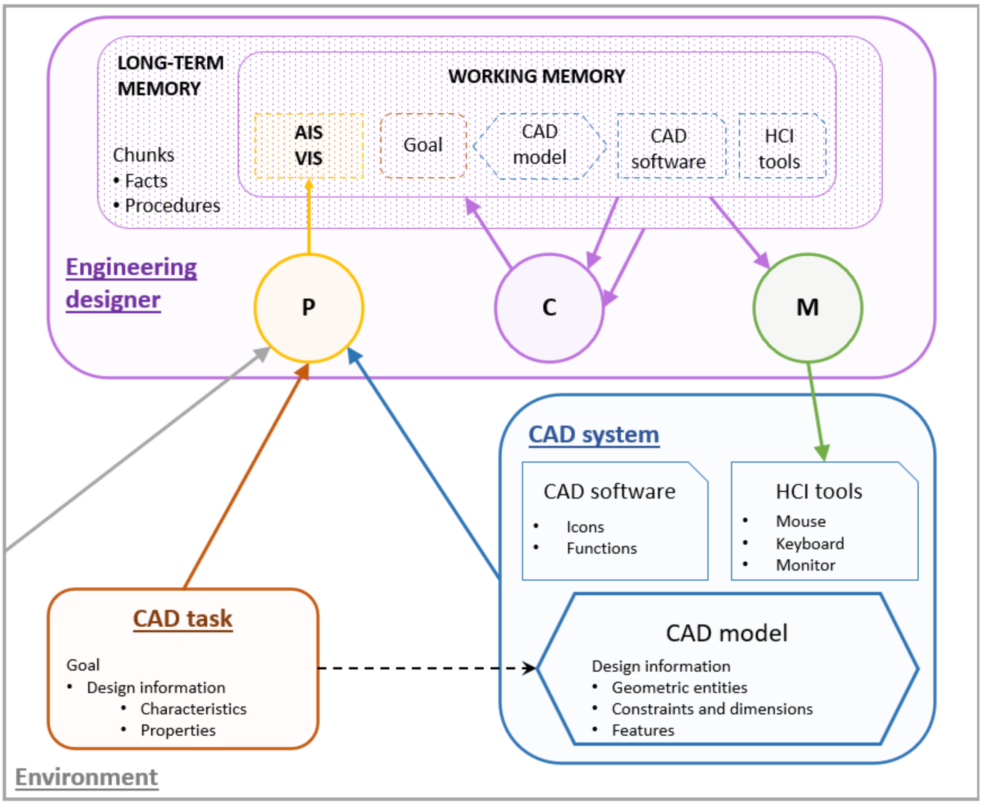

Figure 38 Key elements of CAD performance and CL in CAD

The engineering designer is represented through embodied cognition, as proposed by the MHP [7]; perceptual, cognitive, and motor processing supported by memory structures (WM and LTM). Perception encodes stimuli into internal representations (yellow arrow in Figure 38); cognition operates on these representations in the WM (purple arrows in Figure 38) using knowledge (both facts and procedures) stored in the LTM; motor processing translates selected actions into physical interaction with HCI tools (green arrow in Figure 38). This representation of the engineering designer is intentionally simplified (as in MHP [7]) to provide a usable and empirically compatible structure for modelling HCI performance in complex tasks without requiring fine-grained assumptions that cannot be validated in CAD settings.

The CAD task defines the goal state and the design information that must be transformed into a CAD model. In this thesis, design information refers to the information content provided to the engineering designer that specifies characteristics and properties of the technical system and requirements it must satisfy [68]. Depending on the task setup, this typically includes geometry and dimensions, geometric relations and constraints, feature requirements or CAD modelling intent cues, and other specification criteria embedded in the task representations (e.g. drawing views, annotations, tolerances, reference entities, or instructions that restrict CAD modelling strategy). Thus, the CAD task determines what must be interpreted and remembered, how much must be held or manipulated in the WM at once, and how strongly LTM schemas can be leveraged (e.g. by chunking familiar feature patterns).

The CAD system comprises the CAD software (icons/commands, functions, interaction states, and system feedback) and HCI tools (mouse, keyboard, monitor) through which CAD actions are executed. The CAD system therefore defines how design information can be accessed and manipulated (available commands, constraints imposed by interfaces states, latency/system responses), as well as what externalised information is available to perception at any time (visual feedback of geometry, dimensions, constraint symbols, feature tree structure, prompts, etc.). The system's responses also shape time allocation between active processing and waiting, which is relevant to process-based performance metrics and later CL modelling.

The environment provides boundary conditions that influence perception and actions (e.g. physical workstation properties such as display size/resolution and input device characteristics, lighting and noise, posture/ergonomics, interruptions, and situational constraints). These factors can affect perceptual clarity, motor precision, and the likelihood of disruptions, and thus contribute indirectly to performance variability and CL. In Figure 38, this influence is represented by the grey arrow into the perceptual operator (P), indicating that the environment can act as an additional information stream that modulates what is perceived and how actions are executed. In this thesis, environmental effects are acknowledged at this conceptual level but are not parameterised or analysed in detail; they are treated as out of scope and are controlled as far as feasible by the experimental setup.

Figure 38 explains relationships between these key elements at the conceptual level by showing how the engineering designer's P, C, and M systems interact with two main information sources during CAD modelling: design information from the CAD task (orange arrow), CAD system feedback and CAD model state from the interface (blue arrow), and environmental stimuli (grey arrow). The P system captures perception of both the task representation and the evolving CAD environment (e.g. reading dimension/relations, monitoring geometry and UI feedback). Perceived information is encoded and supplied to the C system (yellow arrow), where it is integrated with the WM contents (current goal state, intermediate CAD modelling state) and the LTM knowledge (CAD procedures, feature schemas, domain factors) to support decision-making (purple arrows). The C system then selects CAD actions that are executed by the M system through HCI tools (green arrow). These CAD actions update the CAD software state and the external CAD model representation, creating a feedback loop in which new system/model information becomes available to perception (blue arrow). In this way, the model explicitly represents not only what the elements are but also the interaction cycle that connects them (perception → cognition → action → system/model update → perception), which is central for later process-based segmentation, operator mapping, and CL estimation.

Finally, Figure 38 is built as a CAD-specific synthesis of MHP construction reviewed in Chapter 2 (P, C, M processors and WM/LTM roles, recognise-act cycles, and bounded capacity assumptions) with an HCI framing of CAD modelling as transformation of design information into externalised model information via interaction with CAD software and HCI tools. It is included here to summarise underlying theoretical assumption and make them explicit before the next subsection operationalises them by defining the operator set and mapping logic. The conceptual relationships in Figure 38 are then turned into a process-based model and later used as the foundation for the method (prescribed in Chapter 5) that can be applied to real CAD modelling sessions.

4.2 Theoretical model of engineering designers' performance in CAD

CAD performance is defined as engineering designer's behaviour when interacting with a CAD system to solve a CAD modelling task. Consistent with the process view adopted in Section 2.2, CAD performance can be interpreted as the effectiveness and efficiency with which engineering designers carry out the transformation of input information (design problem descriptions, requirements, design information such as characteristics and properties of technical systems) into output information (CAD representations of the technical system) across CAD modelling activities [27]. The literature review showed that prior CAD performance studies operationalise this transformation using metrics at different abstraction levels and from different lenses. Some work proposes layered process representations that connect what engineering designers do to what they produce, such as the event–operation–product model by Ishino and Jin [28], and the multi-layer model that distinguishes product and process information proposed by Sadeghi et al. [29]. Empirical studies then often quantify CAD processes either at a higher level (e.g. segmenting the activity into problem definition, information gathering, CAD modelling, evaluation [33]) or at a lower level, where commands/events/actions are counted, grouped, and linked through distributions and transitions (e.g. [12], [45], [58], [59]). These approaches are highly valuable for describing behaviour and comparing CAD modelling strategies at a coarse level; however, as concluded in Section 2.2.3, they also remain difficult to integrate across studies because the underlying metrics and segmentation schemes vary, and because they do not make explicit which information-processing mechanisms drive observed performance aspects.

To address this gap, the thesis adopts a human performance modelling perspective and conceptualises CAD performance through engineering designers' embodied cognition, i.e. perceptual, cognitive, and motor aspects of behaviour during CAD modelling [23], [83]. This lens is grounded in the Model Human Processor (MHP) [7] reviewed in Section 2.3.1.1 and tested in Section 3.4, which provides a simplified but usable structure for explaining interactive performance under limited processing capacity. Importantly, the purpose here is not to replace established CAD performance metrics, but to add an interpretable embodied cognition layer that supports consistent segmentation and enables process explanations that connect to CL modelling later in the thesis.

Viewing CAD performance through P, C, and M processing offers three explicit benefits that complement the dominant metrics reviewed in Section 2.2. First, it provides a principled way to interpret process-related outcomes (e.g. total duration, rates of progress, time allocation) by decomposing where time is spent; increased time can be expressed as a different mix of perceptual activity (e.g. visual scanning, inspection, monitoring constraint/relations), cognitive activity (e.g. interpreting geometry, maintaining and switching subgoals, planning, evaluating, recovering from errors), and motor activity (e.g. interaction execution and interface overhead). Seconds, it creates a bridge between coarse-grained process analytics and embodied cognition by mapping P, C, and M operators to observed CAD events. Third, because CL is defined as resource demand relative to limited capacity, representing CAD performance in terms of P, C, and M resource engagement provides a direct and theory-consistent foundation for the CL model prescribed in Section 4.2.

This thesis therefore focuses on CAD performance not only as an end-state outcome (effectiveness in terms of model quality), but also as a temporally unfolding process (efficiency and process quality) that can be compared across engineering designers and conditions using a transferable structure. The model operationalises this by representing CAD modelling as sequence of P, C, and M operators mapped to observed CAD actions and aligned to a hierarchical task structure. While classic GOMS models are often developed with a time-prediction objective, the predictive logic is used here in a restricted, pragmatic way; the model primarily serves as a formalised representation of observed behaviour in a CAD modelling session (from screen capture and behavioural traces) that supports reproducible segmentation and interpretable process metrics. This structure representation can later support predictive reasoning, but its immediate role is measurement, comparison, and integration with CL indicators rather than predicting performance.

In that sense, the proposed model complements existing CAD process metrics (e.g. [59], [78]) by keeping their descriptive strengths (process-level comparison and temporal structuring) while adding the embodied cognition interpretation that distinguished perceptual, cognitive, and motor actions. The following sections detail this lens by defining the operator set, hierarchical structure, and mapping logic used to construct the process-based cognitive model of CAD performance.

4.2.1 Perceptual, cognitive, motor, and CAD system operators

Perceptual, cognitive, and motor operators represent the activity of three main systems that define engineering designer's embodied cognition. The list of operators and associated operations in Table 52 was generated through a two-stage synthesis that combined 1) established human performance modelling models and task analysis frameworks (MHP [7] and KLM [92]) and 2) CAD-specific evidence from the exploratory coding activity conducted in the first empirical study (Section 3.4).

The central theoretical framework in this thesis is the Model Human Processor (MHP) [7], reviewed in Chapter 2 and tested in Chapter 3, which decomposes human behaviour in interactive tasks into perceptual (P), cognitive (C), and motor (M) processing supported by memory structures. At the same time, the practical operator-based modelling logic is aligned with the GOMS family and, in particular, the Keystroke-Level Model (KLM) [92], which represents skilled, error-free interaction as a sequence of primitive operators and includes an explicit operator R for system response time that must be waited for. This is directly relevant to CAD modelling, where noticeable processing delays can occur during feature computation and graphical updates. Therefore, Table 52 adopts R to represent CAD system response periods, consistent with the KLM convention, and retains a clear separation between engineering designer's processing (P, C, and M) and CAD system processing (R).

Table 52 List of operators, operations, and associated systems

System	P/C/M/R operation	Description	Operator
CAD	CAD system response	CAD system processes engineering designers' input	R
Motor	Moving a mouse	Move a cursor by moving the mouse with a hand	M
Motor	Clicking a mouse button	Press and instantly release a mouse button	M
Motor	Pressing a mouse button	Press and hold a mouse button	M
Motor	Releasing a mouse button	Release pressed mouse button	M
Motor	Spinning the mouse wheel	Spin the mouse wheel (up or down)	M
Motor	Striking a keyboard key	Press and instantly release a keyboard key	M
Motor	Pressing a keyboard key	Press and hold a keyboard key	M
Motor	Releasing a keyboard key	Release pressed keyboard key	M
Perceptual	Visual scanning	Scanning the screen, often without a one clear target	P

Perceptual	Perceiving	Processing incoming sensory information	P+C*
Perceptual	Recognising the target	Identifying an icon/feature as a match to what is stored in WM	P+C*
Cognitive	Attention shift	Low-level cognitive action: Moving attention from one location/icon/feature to another	P+C*
Cognitive	Visual search	Low-level cognitive action: Intentionally searching for a specific target	P+C*
Cognitive	Recalling/Remembering	Retrieve relevant knowledge from the WM or the LTM [197]	C
Cognitive	Understanding	Construct meaning from a technical drawing [197]	C
Cognitive	Applying	Carry out or use a procedure in a given situation [197]	C
Cognitive	Analysing	Break material into constituent parts and determine how the parts relate to one another and to an overall structure or purpose [197]	P+C*
Cognitive	Evaluating	Make judgments based on criteria and standards [197]	P+C*
Cognitive	Creating	Put elements together to form a coherent or functional whole; reorganize elements into a new patten or structure [197]	C

*P+C indicates concurrent involvement of the perceptual and cognitive processors

The motor operations in Table 52 were derived by translating the low-level interaction primitives used in the coding within the first empirical study (Section 3.4) into operations. These include mouse and keyboard actions (move, click, press/hold/release, wheel, key strike/press/release), which correspond to the kinds of primitives KLM represents (e.g. keystrokes and cursor pointing actions) [92], but are expressed here in a form tailored to the CAD context and the available data streams. The R operation was defined as periods where the CAD system is visibly processing input and the engineering designer is required to wait (e.g., processing icon shown, temporarily unavailable interface response), again consistent with KLM's definition of response time [92].

The perceptual operations (e.g. visual scanning, perceiving/encoding, recognising targets) reflect the MHP view of how task and system stimuli are acquired and converted into internal representations that can support action selection [7]. The cognitive entries are defined at two levels. Low-level cognitive operations are closely tied to interactive behaviour (e.g. attention shifting and visual search) to capture common HCI demands in CAD (e.g. locating commands/icons/features, switching focus between 3D model regions and UI elements). Higher-level cognitive operations (recalling/remembering, understanding, applying, analysing, evaluating, creating) are included to describe reasoning that occurs during CAD modelling (e.g. interpreting technical drawings/specifications, selecting CAD modelling procedures, evaluating intermediate geometry). These categories are adopted from the revised Bloom's taxonomy by Krathwohl [197] and were not directly observed nor coded from the empirical data within the first empirical study. In Table 52, the notation P+C indicates operations in which perceptual encoding is tightly coupled with concurrent cognitive processing and therefore are expected to recruit both the perceptual and cognitive processors. This includes operations in which the stimulus must be interpreted or matched to WM/LTM and/or because perception is guided by top-down attentional control (e.g. recognising an icon/feature as the intended target, shifting attention between UI/model regions, or conducting goal-directed visual search).

Finally, the operator list was checked against the exploratory coding in the first empirical study (Section 3.4), where recurrent CAD events were identified and then used to derive mapping heuristics for assigning operator activity. Table 52 therefore functions as the reference list that makes the mapping rules reproducible; it combines 1) the MHP-based decomposition into P/C/M resources, 2) the KLM/GOMS operator logic including explicit system response time (R), and 3) CAD-specific interaction evidence and cognitive-process terminology required to interpret operator activity within higher-level processing.

4.2.2 Hierarchical structure of CAD tasks

The model of CAD performance is also seen as a stage model of engineering designer's information processing when interacting with a CAD system to solve a CAD modelling task. Consistent with CPM-GOMS [94], it adopts a hierarchical structure of CAD tasks to which the operators are associated. The proposed hierarchical structure consists of three levels at which engineering designers' performance can be observed and described: high-level subtasks/subgoals, intermediate-level subtasks/ subgoals, and low-level subtasks/subgoals (see Figure 39).

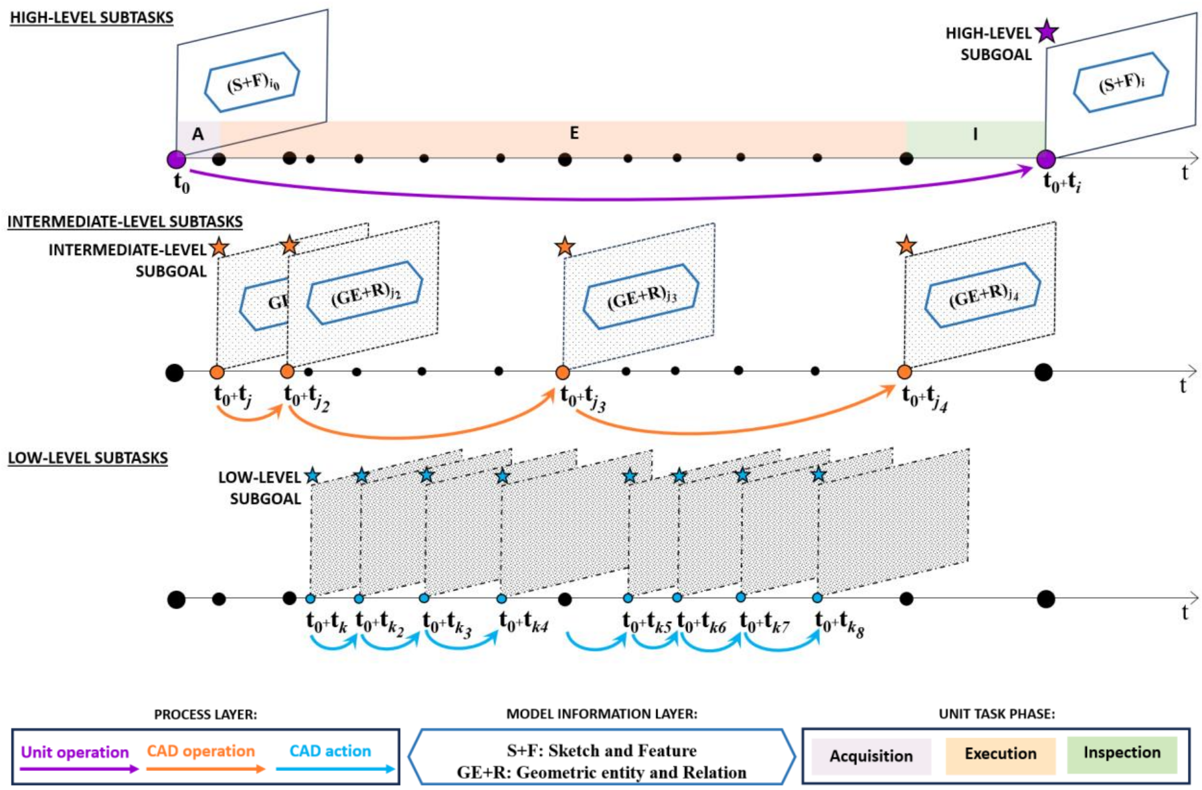

Figure 39 Visualisation of the theoretical model of CAD performance, depicting hierarchical levels of CAD tasks/goals, process layers, and model information layers

Division into these three levels primarily exists due to engineering designers' WM limits on the one side [92] and high cognitive complexity of CAD modelling tasks on the other side [5], [9]–[11]. Thus, dividing the overall CAD modelling task/goal into smaller, quasi-independent subtasks makes their execution cognitively manageable to individual engineering designers [92], [212]. The exact definition of subtasks is influenced by the design of CAD systems; here prescribed definitions focused on solid modelling using feature-based CAD software and a traditional monitor screen, a keyboard, and a mouse.

Each CAD task level is related to a corresponding process layer, which provides information on the process at the level of CAD actions/CAD operations/unit task operations through which the associated low-level/intermediate-level/high-level subgoal is reached. Therefore, these subgoals drive the performance of the subtasks. The outputs of performing unit task operations and CAD operations are visible as transformations of design information represented by model information and observable in two layers: 1) sketches and features as outputs of unit task operations (S+F), and 2) geometric entities, and geometric and dimensional relations as outputs of CAD operations (GE+R). In that sense, operation is seen (in both layers) as a *step in a chain of the activities necessary for the performing of the transformation* [213]. Outputs of CAD actions typically do not change the model information content but can affect how it is presented in the virtual environment (e.g. by zooming in or out).

CAD actions, CAD operations, and unit task operations are defined generic enough to be reusable across engineering designers, CAD tasks, and CAD modelling strategies, but specific enough to preserve the functional meaning of what is being done; they thus represent abstractions at which CAD performance can be observed and compared [214]. Actions and operations are formulated to capture recurrent interaction and CAD modelling functions that appear in most feature-based CAD sessions, rather than being tied to software-specific command names, toolbar locations, or task-specific geometry details. At the same time, the definitions are not made so broad that they collapse distinct behaviours with different cognitive implications (e.g. treating all interface activity as *editing*), because that would reduce interpretability and prevent meaningful comparison across task segments. In practice, the feasible level of generality is bounded by two constraints: 1) the need for reliable coding/mapping from observable interaction traces (so categories must remain operationally distinguishable), and 2) the need for context sensitivity to feature-based CAD (so the categories reflect CAD modelling logic such as sketch–feature cycles and common subgoals as well as actions). Therefore, the proposed action/operation levels should be understood as a reusable baseline for feature-based solid modelling with mouse–keyboard interaction; when the CAD system, input modality, or task type change substantially (e.g. surface modelling, VR/AR CAD, or assembly tasks), the same abstraction principle applies, but the category set may require extension or adaptation while keeping the overall hierarchical structure intact.

4.2.2.1 Unit task operations

The unit tasks in CAD modelling are seen as object-related, which means they are related to a particular part of a geometry (volume) an engineering designer models, as described by Lang et al. [97]. Such a definition is guided by the functionality of feature-based CAD systems, where the model is built as a combination of CAD features while using 2D sketches as a basis [52]. Accordingly, the CAD modelling process can be represented by two main types of unit task operations through which unit tasks are executed:

1) Creating/Editing/Deleting a *sketch* of a part of geometry;
2) Applying/Editing/Deleting a *CAD feature* (such as extrude, revolve, sweep, etc.) to generate a new volume from the sketched part of geometry or to change the existing part of a volume (e.g. with fillet or chamfer).

The outputs of unit task operations are visible as transformations of design information (characteristics and properties of a technical system provided by the task goal) into sketches and/or volumes by applying CAD features; they therefore belong to Sketch and Features (S+F) model information layer depicted in Figure 40.

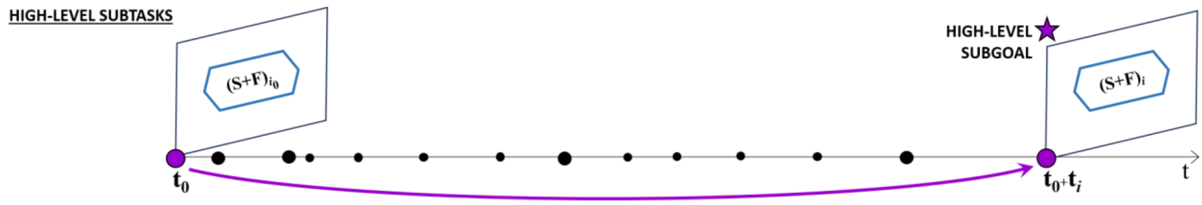

Figure 40 High-level subtasks/subgoals with corresponding unit task operations and (S+F) model information layer

Importantly, it is not necessary that both sketch- and feature-related operations occur within every unit task, nor that they always occur as a fixed pair. While creating new geometry commonly follows a sketch → CAD feature sequence (because many CAD features require a sketch profile), CAD modelling also frequently involves working directly at the feature level without modifying the underlying sketch (e.g. editing feature parameters, or adding feature-based edits such as fillets or chamfers to an existing volume). Thus, after an initial volume is generated, subsequent unit tasks may involve sketching new portions of geometry followed by applying new CAD features, or may involve applying or editing additional CAD features that operate on existing geometry without sketching.

4.2.2.2 CAD operations

CAD systems are structured in a way that a set of CAD actions must be performed to reach the unit task goal (drawing a sketch and applying a CAD feature). The first experimental study showed that engineering designers divide unit tasks into intermediate-level subtasks to achieve the high-level subgoals related to unit tasks, as expected based on the assumptions on memory limits on human cognition [92]. These intermediate-level subtasks are solved through related groups of CAD operations, defined as CAD actions that are functionally and temporally connected to pursue a common intermediate-level subgoal (Figure 41).

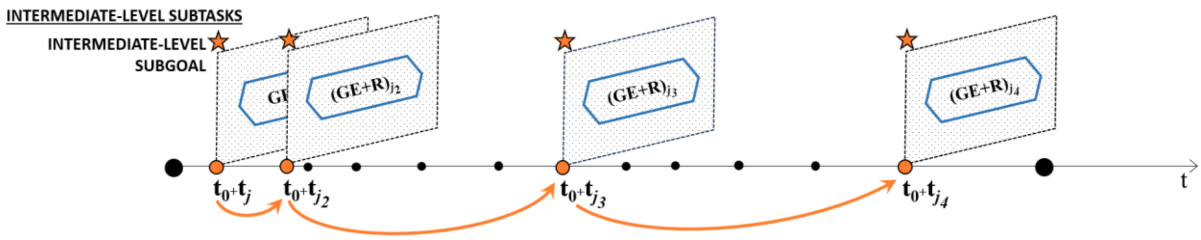

Figure 41 Intermediate-level subtasks/subgoals with corresponding CAD operations and (GE+R) model information layer

Due to the vast number of possible CAD operations an engineering designer might execute when CAD modelling, they are further grouped into seven *subtask clusters* considering the functional role the group of CAD actions played in the CAD modelling sequence. Their definition (see Table 53) was inspired by previous similar work detailed in Section 2.2.2.2.2. (e.g. [12], [58], [63]) and corroborated by empirical data from the first experimental study (see Section 3.4.1). In that way, the subtask clusters are abstracted with an intention to be generalisable across different CAD modelling tasks and engineering designers, which is required for modelling and comparing purposes [214].

Table 53 List of subtask clusters and CAD operation examples

*T	Subtask cluster	Definition	Example of CAD operation	Time [s]
	Navigating CAD application	Switching between tabs and toolbars, opening and closing toolbar menus	• Select Sketch tab to reveal sketch toolbar • Select Feature tab to reveal feature toolbar	1.4-2.7
		Turning sketching or featuring mode on or off	• Enter/Exit a sketching mode • Enter/Exit featuring mode	
		Turning CAD tool on or off	• Turn geometric entity tool on (e.g. select Line, Circle) • Turn geometric entity tool off (e.g. strike Escape or select Confirmation icon) • Turn dimensioning tool on (e.g. select Smart Dimension) • Turn dimensioning tool off (e.g. strike Escape, select Confirmation icon, deselect Smart Dimension icon)	
	Manipulating a view	Adjusting the representation of a geometry or volume within the graphical area of CAD software. Manipulation does not contain any selection or change of geometry/volume.	• Setting a view • Reorienting/rotating a view • Translating a view • Zooming in/out • Showing/hiding entities/features	1.6-4.9
⬣	Referencing	Creating or selecting a reference feature/geometry.	• Create or select a sketching plane • Select an axis of revolution • Select a sketch as a base for a volume	1.4-3.0
⬣	Drawing	Combining geometric entities to create a shape.	• Draw a shape by using (and connecting) entities defined by 2 points/parameters • Draw a shape by using (and connecting) entities defined by 3 points/parameters • Drag a geometric entity • Trim geometric entity • Delete geometric entity	2.5-6-9
⬣	Constraining	Setting constraints to individual geometric entities or geometric relations between several geometric entities.	• Add geometric constraint to an entity (e.g. *Horizontal*, *Vertical*) • Add geometric constraint between two entities (e.g. *Tangent*, *Coincident*) • Delete geometric constraint	4-7
⬣	Dimensioning	Dimensioning in 2D: Defining dimensions of geometric entities and distances between geometric entities in sketches.	• Dimension a length (line, radius/diameter) • Dimension a distance/angle • Check the dimension	3.7-6.3
		Dimensioning in 3D: Defining the value of a third dimension and its direction when applying CAD features.	• Define dimension value/values • Define direction of adding a volume (a third dimension) • Check the dimension (e.g. Measure tool)	
⬣	Deleting a feature	Deleting the entire sketch or CAD feature from the feature tree.	• Delete a sketch • Delete a CAD feature	

*T stands for Transformation: ⬣ indicates that the CAD operation transforms design information

The outputs of performing CAD operations are visible as transformations of design information (characteristics and properties of a technical system provided by the task goal) into geometric entities that are defined by geometric and dimensional relations; they therefore belong to Geometric Entities and Relation (GE+R) model information layer depicted in Figure 41. Exceptions are two CAD operations belonging to the following two subtask clusters: *navigating CAD application* and *manipulating a view*. The first subtask cluster concerns interaction with icons and functions of CAD application, but it does not directly transform design information. These CAD operations primarily exist as a necessity of using CAD software and HCI tools to interact with a virtual environment. The latter subtask cluster (*manipulating a view*) affects how design information is graphically presented by a CAD model in the virtual environment, but it does not change the content or amount of information transformed in a CAD model.

4.2.2.3 CAD actions

CAD actions represent repeated low-level interactions between an engineering designer and the CAD system, based on physical input through HCI tools (typically a mouse or a keyboard). These actions represent recurrent types of interactions involving mouse-based (e.g. cursor movement, selection, dragging) and keyboard-based (e.g. typing values or entering commands) inputs. As such, their outputs typically do not change the model information content; therefore, no additional model information layer is related to low-level subtasks/subgoals (Figure 42).

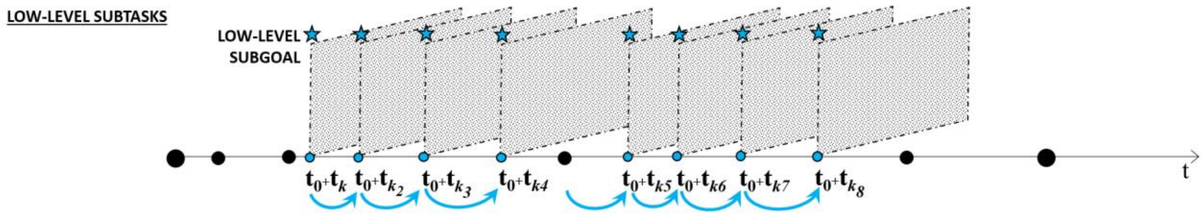

Figure 42 Low-level subtasks/subgoals with corresponding CAD actions

The final list, consisting of CAD actions originating from studies within the HCI field [7], [92] and those added based on the empirically gathered data are listed in Table 54. The first five CAD actions are generic and may occur in many HCI tasks; however, the last three (*rotating a view, zooming in/out, dragging a geometric entity*) are CAD-specific.

Table 54 List of CAD actions

CAD action	Description/Example	Time [s]
Moving a cursor	Moving a cursor on the screen.	0.86-0.95
Clicking on an icon/feature/textbox	Clicking a mouse button to select an icon/feature the cursor is positioned over.	0.001 (predefined)
Hovering over (pointing to) an icon/feature/point	Positioning a cursor over an icon/feature/point without selecting it.	0.40-1.28
Typing a value	Striking keyboard key(s) to input a value.	0.16-0.58
Entering (striking) a command	Striking keyboard command key, such as Escape, Delete, Enter, Undo/Redo, etc.	0.001 (predefined)
Rotating a view	Rotating a model view about the specified point by the specified angles in the directions of the screen X and Y axes [56].	1.00-3.87
Zooming in/out	Modifying the zoom factor for the model view [56].	$7.69 \cdot 10^{-2}$-0.29
Dragging a geometric entity	Pressing a mouse button to select a geometric entity and moving a mouse to change its position in the virtual global coordinate system.	1.00-2.00

The duration of each CAD action was estimated from the first experimental study, and offered here as a range considering all three participants whose data was analysed in Section 3.3. A predefined duration of 1 ms was assigned to the CAD actions *Clicking on an* *icon/feature/textbox* and *Entering (striking) a command* as a fixed coding value to represent instantaneous actions, and this value was kept constant across occurrences, tasks, and participants.

4.2.3 Unit task phases: acquisition, execution, inspection

To describe CAD modelling as a process, this thesis uses a hierarchical task structure, which is a goal–content decomposition; higher-level goals are achieved through sequences of lower-level operations/actions, and the levels primarily capture what CAD modelling operation/action is being performed to progress the CAD model (e.g. referencing, drawing, constraining, dimensioning). In addition to this hierarchy, each unit task is described through a second, complementary lens that captures how information processing unfolds while a unit task is being carried out. These components are referred to in this thesis as *unit task phases*. A unit task is decomposed into three recurring phases: *acquisition*, *execution*, and *inspection* (represented in Figure 43 and described in Table 55).

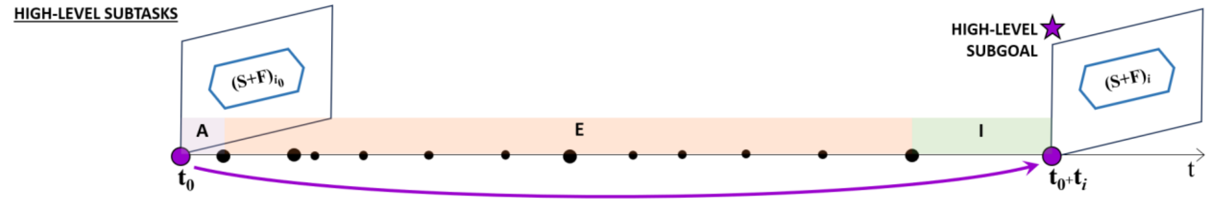

Figure 43 Visualisation of unit task phases within the high-level subtasks/subgoals

Here, phases are not another hierarchical level and should not be interpreted as a further decomposition of unit task operations into smaller subgoals. Rather, phases represent information-processing modes that tend to recur within a unit task regardless of the specific CAD operations used. Unit task operations specify what object-related CAD modelling goal is pursued (e.g. creating a sketch for a particular geometric portion, applying a feature, editing or deleting a feature), whereas unit task phases specify how that goal is approached in time (preparing/doing/checking). Because they address different questions, the two decompositions are intentionally orthogonal and used together; the hierarchy localises behaviour to CAD modelling goals and functional operations, while phases enable consistent comparison of processing demands across engineering designers, strategies, and tasks through mapped P, C, and M operators (see Table 55). Specifically, unit task phases provide the coding context needed to assign operators consistently, because the same observable CAD action/operation (e.g. moving a cursor or manipulating a view) can reflect different information-processing demands depending on whether it supports acquisition, execution, or inspection.

The proposed phases are grounded in human performance modelling assumptions and complemented with the insights from the first empirical study. Classic task analysis and performance modelling in cognitive psychology commonly describe a unit task as a cycle of acquisition (understanding what must be done and preparing to act) and execution (carrying out the required actions), with the purpose of representing and, in some contexts, predicting performance for skilled, largely error-free behaviour [92]. In CAD modelling, however, verifying intermediate results is a frequent and functionally important activity. Results from the first empirical study (Chapter 3) showed that engineering designers often engage in explicit checking of geometry and constraints, typically by visually inspecting the model and, when needed, manipulating the view before moving on. For this reason, *inspection* is introduced as a third phase that complements *acquisition* and *execution* in the CAD context. Although the phases often appear in an acquisition–execution–inspection sequence, the empirical evidence also indicates that phases can be interleaved or repeated (e.g. brief inspection episodes within execution, or returning to acquisition after detecting a mismatch).

Importantly, the same three-phase logic is also consistent with descriptions used in the engineering design literature, where design activity has been characterised using higher-level design operations and information-processing states that enable comparison of processes across individuals and tasks [69], [215]–[217]. In engineering design, these phases are typically described as cycles of analysis–synthesis–evaluation or closely related state-based descriptions [69], [215]–[217]. The unit task phases adopted here align with this view at the functional level: acquisition corresponds to analysis-oriented processing (interpreting requirements, selecting relevant information, structuring subgoals), execution corresponds to synthesis-oriented processing (constructing or modifying the artefact representation through CAD modelling actions), and inspection corresponds to evaluation-oriented processing (assessing the produced state against requirements and deciding whether correction is needed) [69], [215]–[217]. This mapping is not intended as a one-to-one claim about internal cognitive states; rather, it provides a theoretically grounded functional analogy that supports interpretation and comparison across task segments and tasks.

Table 55 summarises the three unit task phases, their definitions, typical operator involvement, and indicative absolute and relative duration ranges (based on the results of the first empirical study, see Section 3.4). In brief, acquisition is primarily perceptual–cognitive (P+C) as engineering designers form or refresh a mental representation of the subgoal and plan the CAD modelling approach. Execution combines cognitive, perceptual, and motor activity (P/C/M) as engineering designers interact with the CAD system to carry out the intended model transformation, with the exact operator mixture depending on the CAD operations performed. Inspection is primarily perceptual–cognitive (P+C), often accompanied by motor activity (M) when engineering designers manipulate the view or cursor to support checking, and it provides an explicit mechanism for capturing verification and potential rework that are characteristic of realistic CAD modelling.

Table 55 Parts of unit tasks/goals

Unit task phase	Definition	Operator	*AD [s]	RD [%]
Acquisition	Cognitive and perceptual actions required to understand the task and define goals/subgoals. *During acquisition the user builds a mental representation of the task/subtask* [92].	Always P + C M may co-occur	2-7 s	20-30%
Execution	Cognitive, perceptual, and motor actions required for executing CAD actions. *During execution the user calls on the system facilities to accomplish the task* [92]. These system facilities are P, C, and M actions, modelled through the operators.	Depends on a CAD action	5-8 s	45-60%
Inspection	Cognitive, perceptual, and motor actions required for checking the accuracy of the modelled geometry [52], [97].	Always P + C M may co-occur	2-4s	10-24%

*AD stands for Absolute Duration [s]; *RD stands for Relative Duration [%]

4.2.3.1 Acquisition

In the acquisition part of a unit task an engineering designer gathers or recalls information for understanding the goal/subgoal, required to execute it in the following step. Therefore, acquisition is primarily perceptually and cognitively demanding, which is modelled with the parallel activity of P and C operators during the entire phase (both are active during the entire phase), while allowing brief concurrent motor actions (M) when needed, without changing the phase's primary information-processing role.

The initial acquisition that occurs early in the CAD modelling session (typically as a part of the first unit task) is expected to be longer because it includes familiarisation with the presented information and its broader comprehension [183]. When CAD modelling from technical drawings, engineering designers first need to perceive a visual representation, and visualise the represented technical system in 3D in their mind (generate their mental 3D model of a technical system). Visualising a technical system requires understanding its characteristics and properties from a visual representation, and often asks for visuospatial transformations of presented information as well as mental integration of multiple views and annotations [35], [52].

Once formed (visualised), the model of technical system is then typically deconstructed into its constituent elements (parts of geometry), in accordance with CAD features through which it can be built [52]. The acquisition of information is here directed toward finding the match between the parts of geometry and CAD features. Sketches and CAD features are cognitively assembled in a sequence that serves as the CAD modelling procedure. This procedure is

commonly maintained internally (in the WM) rather than written down, and it is shaped by both CAD knowledge and CAD modelling intent considerations (defined by task goal) [52].
After the initial acquisition, this phase is expected to recur at the start of each subsequent unit task, but usually with shorter duration and a narrower focus. Rather than reconstructing the full presented technical system, engineering designers here search for the specific subsets of information relevant to the next CAD modelling step (e.g. a particular dimension or relation), refresh the subgoal, or prepare the next action/operation sequence [183].

4.2.3.2 Execution

During the *execution* phase, an engineering designer carries out the planned sequence of CAD actions through interaction with the CAD system using a mouse and a keyboard as input devices. Execution therefore includes motor activity (M) by definition, while perceptual (P) and cognitive (C) activity co-occur depending on the specific CAD actions and the subtask context. Execution is where the intended transformation of the CAD model is implemented through concrete CAD actions and operations.
In this thesis, execution can be described in terms of recurring CAD operation clusters that provide functional context for low-level CAD actions. These include navigating the CAD application, referencing, drawing, constraining, dimensioning, manipulating the view, and deleting. Although these clusters are introduced as intermediate-level subtasks and process units, they also help interpret execution-phase demands: for example, manipulating view typically couples motor control with ongoing perception (P+M), whereas drawing and dimensioning often require tighter coupling of cognition, perception, and motor control (C+P+M). Execution phases can also include brief pauses or micro-episodes of checking and decision-making; when such pauses become substantial and functionally shift from “doing” to “checking”, they are treated as transitions into inspection.

4.2.3.3 Inspection

During the *inspection* phase, an engineering designer evaluates the result of the executed actions and verifies correctness by comparing the current CAD model state with the requirements provided by the task instructions or by the intended intermediate goal. Inspection is therefore primarily perceptual–cognitive and is modelled as continuous activity of P and C operators across the phase. Motor activity (M) may occur in parallel, for example when *navigating* the CAD system, *manipulating view* through *zoom in/out* and *rotate* CAD actions, or simply *moving a cursor* where the attention is focused.

Inspection is introduced explicitly because it captures a recurrent and meaningful function in CAD modelling: ensuring that geometry, dimensions, and relations match the intended design before proceeding. This phase is also central for understanding rework and iterative refinement. When inspection reveals discrepancies, designers may return to acquisition (to reinterpret requirements or revise the plan) and/or re-enter execution (to correct geometry or constraints). As a result, inspection is not assumed to occur only at the end of a unit task; it may appear as shorter episodes embedded between execution segments, depending on the designer's strategy, uncertainty, and task complexity.

4.2.4 Mapping P, C, and M operators to CAD actions, CAD operations, and unit task parts

The model of engineering designers' performance is conceptualised as a sequence of CAD actions that are grouped into CAD operations and situated within unit task operations and unit task phases. P, C, M, and R operators, representing engineering designer's embodied cognition and CAD system response, are then mapped to these sequences using a rule-based procedure. Conceptually, the theoretical model summarises in Figure 44 the basic mapping logic at the level of CAD actions and the unit task phases (specifies what is mapped to what). Operationally, the method specifies how the mapping is implemented in coding through detailed decision rules, which are provided in Table 60 (Section 5.2.3).

The purpose of the mapping is to translate observable CAD behaviour (coded CAD actions and their aggregation into CAD operations and unit task operations) into a theory-consistent representation of embodied cognition, expressed as the time-varying engagement of perceptual (P), cognitive (C), and motor (M) resources, together with CAD system response periods (R). In other words, the hierarchical task structure and process layers defined in previous sections specify what engineering designers do (and at which goal/task level), whereas the operator mapping specifies what information-processing resources are engaged while doing it. This mapping is essential because subsequent analyses quantify CAD performance not only through outcome and process metrics, but also through resource engagement that can be compared across engineering designers and CAD tasks, and later related to CL.

UNIT TASK PHASE	Acquisition	Execution										Inspection									
CAD EVENT		Moving a cursor	Clicking on	Hovering over	Typing a value	Entering a command	Rotating	Zooming in/out	Dragging	CAD system response	(Pause)	Moving a cursor	Clicking on	Hovering over	Typing a value	Entering a command	Rotating	Zooming in/out	Dragging	CAD system response	(Pause)
OPERATOR	C			C							C	C	C	C	C	C	C	C	C		C
	P			P			P	P			P	P	P	P	P	P	P	P	P		P
		M	M		M	M	M	M	M			M	M		M	M	M	M	M		
										R										R	

Figure 44 Mapping the operators to CAD actions and unit task phases

The mapping rules for P and C are grounded in the assumptions of Card et al. [92] and therefore function as heuristics, not as a direct measurement of internal cognitive states. To make the mapping reproducible and interpretable, the rules explicitly incorporate CAD context in two ways. First, unit task phase context is used because the same observable CAD action can correspond to different information-processing demands depending on whether it supports preparing/understanding (acquisition), doing (execution), or checking (inspection). Second, subtask cluster (CAD operation) context is used where applicable (e.g. view manipulation during execution), because certain CAD actions (e.g. rotating, zooming, dragging) are closely tied to perceptual monitoring and evaluation of the 3D CAD model and therefore imply different P/C involvement depending on how they are performed and for what functional purpose.

Finally, the mapping does not explicitly parameterise individual differences in knowledge, experience, or expertise even though these factors can affect activity of the C operator through differences in how chunks are organised [92]. Instead, such differences are expected to manifest in the resulting operator time allocations and sequences (e.g. more or fewer P+C episodes, different unit task phase distributions, different frequencies of inspection or rework), as well as in the associated behavioural measures (durations and distributions of CAD actions/CAD operations/unit task phases). In this way, the mapping provides a common representational layer that enables comparison while preserving meaningful variability in behaviour.

The mapped operator sequences yield quantitative descriptors of resource engagement at multiple hierarchical levels. These descriptors provide the basis for computing utilisation-based indicators of CL, because CL is conceptualised as demand placed on limited information-processing resources. The next section therefore builds directly on this operator-based representation to introduce the theoretical model of CL, specifying how CL is estimated from operator utilisations and, later, how these estimates can be integrated with EEG-derived information about processing intensity.

4.3 Theoretical model of CL in CAD

As reviewed in Section 2.4.1.3, prior cognitive modelling studies typically quantify CL/WL from the model's internal activity using either capacity utilisation (how "busy" a module/subnetwork is) and/or activation time (how long it remains engaged) [112], [114], [116]–[121]. The present model adopts the same core principle (CL is linked to the temporal occupation of processing resources) but operationalises utilisation from the P, C, and M operators of the CAD performance model, which correspond to the perceptual, cognitive, and motor systems of embodied cognition. Because utilisation is fundamentally a time/occupancy measure, the model further distinguishes utilisation (duration of engagement) from information-processing intensity (effective processing cost per unit time), which is introduced via multipliers at the level of task segments and experimental conditions. In this context, utilisation denotes the proportion of an observed period in which a given system is actively engaged in information processing when performing a task, i.e. how busy that system is over time [114]. This view aligns with established conceptualisations of WL as the ratio of resources required to resources available, where required resources are reflected in the time needed to process information within the time available, thereby linking WL to the temporal occupation of processing resources [115].

Modelling CL in CAD therefore builds upon quantifying WL as a utilisation *U* of the P, C, and M resources required to accomplish CAD task goals and thus solve the task [113]. The first assumption of the model is that the ratio of the total time the P, C, and M operators are active in an observed period *T* (i.e. their utilisation *U*) can serve as an index of overall WL [112]. Therefore, Equation (33) represents the contribution of each embodied cognition system *i* to WL based on its activity A_i within the observed period *T* (i.e. its utilisation U_i).

$$\mathrm{WL}_i = U_i = \frac{\int_0^{\mathrm{T}} A_i(t) \cdot dt}{\mathrm{T}}, \qquad U_i \in [0,1], \qquad i = P, C, M. \tag{33}$$

The total WL across systems is then calculated as their sum, according to Equation (34):

$$\mathrm{WL} = \sum_i \frac{\int_0^{\mathrm{T}} A_i(t) \cdot dt}{\mathrm{T}} = \sum_i U_i, \qquad U_i \in [0,1], \; i = P, C, M. \tag{34}$$

Here, *T* may represent the entire CAD task, unit tasks, intermediate-level subtasks, or a predefined moving window.

To reflect the multi-component nature of WL, operators P, C, and M are treated as separate resource systems. The utilisation of M resources is primarily linked to physical component of WL, whereas CL is linked primarily to the utilisation of the C and P systems. This motivates the second assumption; the effect of system activity on CL differs for P, C, and M resources, which is represented through base weights w_i in Equation (35), thus adapting Equation (35) to estimate CL:

$$\mathrm{CL} = \sum_i w_i \cdot U_i\,, \qquad U_i \in [0,1], \qquad i = P, C, M. \tag{35}$$

In the present model, P activity is assumed to contribute less to CL than C activity (e.g. $w_P = 1$; $w_C = 2$), while M activity is not assumed to contribute to CL under normal circumstances ($w_M = 0$). The assigned base weights w_i are assumed to be constant and their values can be defined based on the previous literature (e.g. [112]), primarily to enforce a clear, interpretable distinction in the contribution of each embodied cognition system to CL. Importantly, these weights should be understood as relative scaling factors rather than universal constants; they *translate time a system is engaged* into *expected contribution to CL,* and their numerical values may depend on task context. For example, Jo et al. showed in a related utilisation-based WL formulation that multiple weight combinations can yield similarly good agreement with NASA TLX (as a controlling variable), indicating that weights are best treated as plausible priors rather than uniquely determined parameters [112]. Consequently, in the CAD context w_i are adopted primarily for their theoretical ordering ($w_C > w_P > w_M$), while their exact magnitudes remain open to CAD-specific corroboration.

While w_i differentiates the contribution of systems, it still treats each period of engagement as equally "costly". This motivates the third assumption of the proposed model: CAD task segments differ in information-processing intensity because they require processing different amounts and types of information (chunks), and different cognitive operations on that information. Specifically, *amount* refers to the number of information elements that must be simultaneously encoded, maintained, compared, or updated (i.e. the WM burden), whereas *type* refers to the nature of the information and operations involved (e.g. interpreting and integrating task geometry during acquisition, evaluating consistency during inspection, recalling references, or performing predominantly low-level visual search during navigation/manipulating). Because two task segments can show similar utilisation yet differ in the number/complexity of chunks processed per unit time, utilisation alone is insufficient for capturing assumed segment-level differences in information-processing intensity [114]. Therefore, additional information-processing weights ω_s are introduced to scale base-weighted utilisation by segment *s* (Equation (36)):

$$\mathrm{CL}_s = \sum_i \omega_s \cdot w_i \cdot \frac{\int_0^{\mathrm{T}_s} A_i(t) \cdot dt}{\mathrm{T}_s} = \sum_i \omega_s \cdot w_i \cdot U_{i,s}\,, \qquad (36)$$
$$U_{i,s} \in [0,1], \qquad i = P, C, M.$$

In this formulation ω_s is a dimensionless relative multiplier that operationalises the "amount and time of information" assumption; it does not claim to measure chunk counts directly, but instead provides an interpretable proxy for relative intensity across segments. The rationale for ω_s is that segments that 1) require maintaining and integrating more task-relevant information (more chunks) and/or 2) invoke more demanding cognitive operations (e.g. analysis, evaluation, generation) should contribute more to CL than segments dominated by low-level operations such as routine visual search, even if both occupy similar time. In the present work, an initial ordering of unit task phases and intermediate-level subtasks by assumed intensity (shown in Table 56) is derived from utilisation results extracted from the first empirical study (averaged utilisations across the participants and task segments) and assumed cognitive interpretation of chuck demands per segment (e.g. acquisition/inspection requiring broader task information integration; dimensioning requiring entity/type/value integration; navigation primarily involving icon and location information).

Table 56 Proposed information-processing intensity rank of segments, based on utilisations

Segment	Description/Explanation	$\sum_i w_i \cdot U_i$	Rank
Acquisition	• Chunks: All design information that can fit the WM • High-level cognitive activity: Analyse, Understand, Create	3	1
Inspection	• Chunks: Sketched/modelled geometry, Design information related to sketched/modelled geometry (available by the task goal) • High-level cognitive activity: Analyse, Evaluate	3	2
Manipulating	• Chunks: Sketched/modelled geometry • Low-level cognitive activity: Visual search	2.85	3
Referencing	• Chunks: Reference feature/geometry • High-level cognitive activity: Recall • Low-level cognitive activity: Visual search	2.36	4
Navigating	• Chunks: Icon, Icon location • High-level cognitive activity: Recall • Low-level cognitive activity: Visual search	1.55	5
Dimensioning	• Chunks: Geometric entity/entities, Dimension type, Dimension value • High-level cognitive activity: Recall, Apply • Low-level cognitive activity: Visual search	1.49	6
Drawing	• Chunks: Design information (shape), Geometric entities, Relations between geometric entities • High-level cognitive activity: Recall, Apply • Low-level cognitive activity: Visual search	1.28	7
Constraining	• Chunks: Design information (shape), Geometric entities, Relations between geometric entities, Constraint icon, Constraint icon location • High-level cognitive activity: Recall, Apply • Low-level cognitive activity: Visual search	1.07	8
Deleting	• Chunks: Sketch/CAD feature • High-level cognitive activity: Recall, Apply • Low-level cognitive activity: Visual search	0.99	9

However, since utilisation is fundamentally a time/occupancy measure, this initial ordering should be corroborated using psychophysiological measurements that are sensitive to processing intensity beyond duration. EEG is well suited for this purpose, as previous work (see Section 2.4.3.1) showed that brain activity changes systematically with changes in cognitive demands, supporting its interpretation as a psychophysiological indicator of information-processing intensity. This suggested modelling approach, combining time-based resource engagement with psychophysiological indicators to strengthen inferences about momentary processing demands, has precedent in prior work that links cognitive modelling with EEG-based measures of CL [218]–[221]. Taken together, adding EEG-based weights to utilisation equations is theoretically justified as a principled way to combine two complementary components of CL estimation: 1) duration of resource engagement, captured by utilisations of P, C, and M operators, and 2) intensity of information processing, captured by EEG-informed weights. In practical terms, ω_s can be interpreted as an intensity gain that scales time-based utilisation into an estimate of effective CL within each task segment, consistent with the thesis' goal to track CL dynamically across task segments or moving windows throughout the CAD modelling process.

Furthermore, as literature review showed, many factors could affect CL (Section 2.4.2) and engineering designer's performance (Section 2.2.1) in CAD activities. In principle, all characteristics of the four key elements of CAD performance and CL (an engineering designer, a CAD task, a CAD system, and an environment) present the potential factors influencing these two constructs (Figure 38). However, modelling the full range of interacting subject-, task-, and method/tool-related factors in a single coherent framework is not realistically achievable, nor would it remain interpretable for CAD research and practice [92]. Therefore, the proposed model follows an intentionally minimal strategy, prioritising dominant influences that are theoretically motivated and empirically supported [102]. Within the scope of this thesis, the model explicitly considers two main task-related factors: 1) design information format (2D-orthographic vs. 3D-isometric) and 2) task complexity (low vs. high geometric and CAD model complexity).

Regarding information format, additional cognitive resources are assumed to be required to mentally transform and integrate information from three orthographic views into a coherent mental 3D model, compared to isometric projection that already provides spatial cues [35], [201]. This mental model must either be maintained in the WM or reconstructed repeatedly whenever external information is re-acquired, which is expected to increase CL; this effect is expected to be strongest in acquisitions, but it may also influence other segments (e.g.

inspection when the model must be checked against the goal, or referencing when geometry must be re-interpreted). Similarly, higher complexity is assumed to increase CL and results in a division of the task into more subgoals due to the enlarged number of interacting elements engineering designers must simultaneously process as they are logically related [13], [107]. Although complexity effects are also expected to be most pronounced during acquisition (encoding and structuring task information into CAD modelling intentions), they can plausibly affect later segments (e.g. inspection and constraint application) as the number of relations to verify and maintain in the WM increases.

To incorporate these factors explicitly, the model introduces condition-level intensity gains G that scale the contribution of utilisation to CL. This follows established practice in cognitive modelling of WL, where additional influences are incorporated by modifying processing capacity, re-weighting model components or by introducing multiplicative modifiers that capture systematic changes in processing demands under different conditions (e.g. [112], [116]–[119]). Analogously, in the present model, information format and task complexity can be represented by their own gain parameter, applied separately depending on which experimental condition is being modelled, while keeping the base operator weights unchanged. Accordingly, for a given factor, a segment-specific condition multiplier G_s can be adapted as suggested by Equation (37) and applied directly to the weighted utilisation-based contribution in the specific segment (Equation (38)):

$$G_s = (1 + d_{condition,s}) \tag{37}$$

$$CL_s = G_s \cdot \omega_s \cdot (w_P U_{P,s} + w_C U_{C,s}) \tag{38}$$

Here, w_P and w_C represent the base weights for the perceptual and cognitive operators, $U_{P,s}$ and $U_{C,s}$ are their utilisations in segment s, while $d_{condition,s}$ is a dimensionless parameter capturing the relative increase in information-processing intensity per unit utilisation time attributable to the specific condition being analysed (i.e. $d_{format,s}$ for information format variations, or $d_{complexity,s}$ for complexity variations, respectively). The subscript s indicates that these gains can, in principle, affect any task segment; however, based on the theoretical rationale above, acquisition is expected to show the largest effects for both factors. In the absence of segment-level evidence, G_s can be treated as constant across segments, while still maintaining the theoretical expectation that acquisition is the primary locus of impact.

Therefore, the factors are treated as separate, explicitly parameterised influences that enter the CL formulation through distinct gain; either $d_{complexity,s}$ or $d_{format,s}$, estimated from their corresponding experimental contrasts. The two-level operationalisations (orthographic vs.

isometric; low vs. high complexity) should be understood as controlled experimental instantiations of broader constructs (representational transformation demands and intrinsic element-interactivity demands). In the present model, the gains are interpreted as net effects of the manipulation under controlled conditions, while other potential influences are held constant by experimental design (same CAD system, comparable task type, same participants, same procedure) and explicitly acknowledged as out of scope for the present model.

The values of $d_{complexity,s}$ and $d_{format,s}$ are proposed to be calibrated using EEG, because utilisation primarily reflect the duration/occupancy of processing resources, whereas EEG-derived TRP reflects changes in brain activity relative to baseline and is widely used as an indicator of changes in CL (see Section 2.4.3.1). This provides the rationale for linking TRP to the intensity-gain term; if two task conditions yield comparable utilisation but reliably different TRP, the difference is interpreted as a change in processing intensity. Consequently, differences in TRP between conditions provide a principled basis for scaling utilisation-based CL estimates with gains $d_{condition,s}$ interpreted as an empirical calibration coefficient that maps observed psychophysiological CL differences onto a multiplicative adjustment of the utilisation-based processing costs.

In practical terms, $d_{format,s}$ (and analogously $d_{complexity,s}$) can be defined from the standardised effect size of the TRP difference between the relevant conditions, computed at the task level or, if available, per task segment. This mapping does not assume a universal physiological law; rather, it is a monotonic calibration rule that a larger condition effect on the CL indicated by psychophysiological measures implies a larger intensity gain applied to the utilisation-based estimates. This logic is consistent with prior modelling work that validated utilisation/activation-based CL indices against subjective CL measures and psychophysiological indicators [112], [116]–[119]. Importantly, the same prerequisite applies here as for deriving the information-processing weights ω_s; both the segment-level intensity weighting and the condition-based gains ($d_{condition,s}$) depend on first identifying an EEG feature that serves as a reliable EEG-based indicator of CL in the given CAD context. Accordingly, the next section prescribes the overall method, within which one key step is to identify the EEG-based CL indicator (Section 5.3.2) and use it to estimate EEG-informed parameters (including ω_s and $d_{condition,s}$) from the relevant differences.

This page intentionally left blank

5 A METHOD FOR MEASURING AND ANALYSING COGNITIVE LOAD IN CAD ACTIVITIES

Chapter 2 showed that CAD performance has been assessed using heterogeneous metrics and definitions. Output-based measures of the final CAD model (e.g. completeness) are widely used because they are intuitive and accessible, but they provide only an end-state snapshot and cannot explain how performance unfolded (Section 2.2.2.1). Process measures are increasingly used (e.g. time, number/type of features, number of sketches), yet they often remain cumulative summaries that do not explicitly capture the temporal structure of CAD modelling or the cognitive demands associated with different segments of the process (Section 2.2.2.2). In parallel, Chapter 2 also established that understanding CAD modelling as an HCI activity benefits from human performance modelling perspectives grounded in embodied cognition (perceptual, cognitive, motor processing) and task decomposition (Section 2.3). However, the reviewed work provides limited consolidation of how such perspectives can be synthesised and translated into CAD performance and CL models that remain interpretable and usable for analysing realistic CAD modelling sessions.

This chapter therefore prescribes a method for measuring and analysing CL in CAD activities. It builds on insights from the literature review (Chapter 2) and the first empirical study (Chapter 3), which supported the feasibility and suitability of the individual constructs here integrated in the method. Additionally, the method is grounded in the theoretical models introduced in Chapter 4, which define the key elements and relationships needed to represent, measure, and relate CAD performance and CL. Finally, the method is shaped by the objectives and the main purpose defined in Section1.3 (Table 3), ensuring that the prescribed steps, required inputs, and produced outputs remain coherent, usable, and suitable for application and validation in realistic CAD modelling settings. In line with these objectives, the method is designed to capture dynamic changes in CL across the CAD modelling sequence and relate them to task segments, to rely on a minimal yet robust and clearly specified set of metrics and steps, and to enable systematic relating of CL to CAD performance considered both as a process and as CAD modelling outputs, while remaining understandable and reproducible for engineering design researchers.

5.1 Method assumptions, input requirements, and scope

The method for measuring CL in relation to CAD performance requires a multimodal approach, combining analytical methods with empirical ones, where the first grounds upon cognitive modelling while the latter includes performance assessment, subjective measurement, and psychophysiological measurement methods. Here prescribed method thus adopts a process-based approach to measuring and analysing CL in relation to CAD performance that combines NASA TLX questionnaire (subjective assessment), cognitive modelling (performance modelling and task analysis frameworks), and EEG (psychophysiological measurement) methods.

This triangulation is motivated by the CL literature, which emphasises that CL is a covert construct inferred indirectly and is best captured through complementary methods that differ in temporal granularity and intrusiveness [13], [102], [103], [106]. In CAD, this is particularly important because CAD modelling is a lengthy, cognitively complex HCI activity in which information-processing demands are expected to fluctuate across segments rather than remain constant over the full task execution [5], [9]–[11], [15], [16], [123].

The method builds on three theoretically grounded assumptions; they are described below.

1) Engineering designers' embodied cognition and HCI during CAD can be approximated with a simple cognitive architecture such as Model Human Processor (MHP) and its perceptual, cognitive, and motor (P, C, and M) systems.

This assumption follows from the human performance modelling literature reviewed in Section 2.3. MHP is intentionally simplified and designed for estimating and explaining performance in interactive tasks under limited information-processing capacity, decomposing processing into P, C, and M systems that can be related to observable interaction [7], [84], [92]. This is advantageous in CAD because CAD modelling tasks are complex and highly variable, and richer architectures (e.g. ACT-R variants) typically require stronger assumptions and more detailed cognitive specifications (e.g. production rules) that are difficult to justify and validate in realistic CAD settings [79], [83]. Therefore, adopting MHP provides a theoretically valid cognitive structure while keeping the modelling burden aligned with what can be supported by available behavioural evidence in CAD tasks [7], [92].

This assumption is also consistent with the CAD performance literature reviewed in Section 2.2, which frames CAD as a process in which design information is progressively transformed into CAD representations through interaction with a CAD system [8]. Process-oriented CAD models similarly imply that CAD modelling behaviour can be described across levels of abstraction, but they do not specify reusable cognitive constructs for interpreting why particular sequences are demanding [28], [29]. The MHP-based P, C, and M structure complements these approaches by providing a principled information-processing operators that can be mapped to HCI in CAD activities.

2) Changes in CL are reflected in the utilisations of P and C operators (systems).

This assumption follows from the resource-based definition of CL adopted in this thesis (demands relative to limited processing capacity [83]) and from CL conceptual models distinguishing a task-centred component (intrinsic/effective/mental load) from an avoidable component influenced by the individual, representation format, and environment (extraneous/ineffective/mental effort) [13], [102], [107]. In CAD terms, these components manifest largely through changes in visuospatial perception, attention allocation, and CAD action selection; processes that are primarily perceptual and cognitive [11], [170], [189].

More explicitly, cognitive models grounded in cognitive architectures provide a formal rationale for using utilisation as a CL proxy. Across the reviewed modelling studies, CL/WL is computed from capacity utilisation (how busy a module/subnetwork is over time) and/or activation time (how long a module remains engaged), directly reflecting the idea that higher load corresponds to a higher proportion of limited processing resources being occupied [112]–[118]. This logic is also consistent with broader theoretical formulations of CL/WL as ratios of required to available resources and as time required relative to time allowable for information processing [119]–[121]. Therefore, translating this to CAD, the method assumes that when CAD modelling becomes more demanding (e.g. due to added visual search, increased inference effort or increased verification/rework), engineering designers will show increased engagement of P and C processing, which can be approximated as higher estimated utilisation of P and C operators [108], [112]–[115], [119]–[121].

This is also consistent with process-based CAD performance research that highlights that differences are often expressed as how time is allocated across process segments (e.g. more checking/revision, different distributions of actions/segments), not only as longer total duration [33], [58], [63]. Operator utilisations thus provide a segment-aligned, interpretable bridge between CL theory and process-based CAD performance descriptors [112], [115].

3) Changes in CL are reflected in EEG signals.

This assumption follows from the psychophysiological measurement literature in which CL is inferred from correlates and EEG is widely used because it enables continuous monitoring with high temporal resolution, which is critical when CL is expected to vary across time and segments [103], [170]. EEG-based CL research in cognitively controlled tasks repeatedly shows systematic changes in oscillatory activity with increasing demands (commonly in theta/alpha/beta power and their ratios), supporting the plausibility that EEG signals reflect changing information-processing demands [149], [154], [157]–[160].

For CAD-related and design tasks, EEG studies further support this premise while motivating caution; CL-sensitive EEG features are more diverse and can behave differently than in controlled tasks (including direction of effects), implying that transfer cannot be assumed without CAD-specific feasibility checks and validation. Accordingly, the method does not assume a universal EEG-based CL indicator. Instead, it assumes that weaker but theoretically justified premise that EEG contains information sensitive to changes in information-processing demands during CAD modelling and can therefore complement subjective and process-based indicators with a temporally informative trace [147], [149].

Taken together, these three assumptions justify why the method combines NASA TLX, operator-based cognitive modelling, and EEG rather than relying on a single measure. NASA TLX is widely used as a subjective method in CL/WL research and is frequently used for validation in cognitive modelling studies [112], [114], [116], [117]. However, subjective ratings typically yield task-level summaries and do not localise when CL increases/decreases throughout the task [101], [102]. Operator utilisations provide interpretable, segment-aligned behavioural proxies grounded in resource-based CL theory and cognitive modelling formulations of CL/WL from utilisation/activation [112]–[115], [119]–[121]. EEG provides continuous psychophysiological sensitivity needed to capture temporal fluctuations in CAD activities, while still requiring CAD-specific validation of candidate features. This triangulation directly addresses the gaps highlighted in Chapter 2; CAD performance research rarely integrates temporally structured process analysis with temporally informative CL measurement, and CAD-specific evidence for EEG-based indicator remains narrow and demands careful validation.

The applicability of the proposed method is intentionally bounded by the scope of its theoretical foundations and the practical constraints identified in the literature review and the first empirical study. First, it targets CAD modelling tasks rather than open-ended design problems; the method is grounded in human performance modelling and operator-based task descriptions (MHP/KLM style) and therefore focuses on HCI-driven information processing during CAD modelling, not on design problem solving mechanisms such as creativity, reframing, or broader design-space exploration. Second, it is scoped to feature-based solid modelling in conventional desktop CAD environments (monitor, keyboard, mouse), because the prescribed segmentation structure and operator mappings are derived for such interaction paradigms and would require redefinition for other CAD modalities (e.g. surface modelling, VR/AR, gesture-based systems). Third, the method is intended for description and analysis of CAD performance and associated CL rather than prescriptive "best-process" recommendations. This reflects a key conclusion from Chapter 2; even among skilled users, CAD modelling exhibits high process variability because there are often many valid ways to produce the same geometry and feature structure, which limits the feasibility of defining a single normative process without strong task- and context-specific constraints [12], [29], [61].

Finally, the prescribed method requires several datasets as inputs, reflecting its multimodal nature and the need to align process-sensitive performance descriptors with temporally informative CL measures. The required inputs are summarised in Table 57 and span: 1) EEG data that enable non-invasive, time-resolved monitoring of brain activity during CAD modelling, 2) behavioural interaction data capturing keyboard and mouse activity and interface transitions, 3) CAD modelling outputs that support outcome-based performance assessment (e.g. geometric entities/relations, sketches, and features), 4) CAD modelling process recordings (timestamped screen captures) that enable temporal segmentation and mapping of CAD events to unit tasks and their phases, and 5) subjective CL ratings (NASA TLX) used as a criterion measure for validation. Together, these datasets provide the minimum evidence needed to 1) describe CAD modelling as a HCI process, 2) quantify operator utilisations aligned with that process structure, and 3) synchronise behavioural and psychophysiological traces to the same CAD modelling timeline, which is essential if CL is expected to fluctuate unevenly across CAD task segments rather than remain constant over the whole task.

Table 57 Input data requirements to the method

Data type	**Practical requirements**
EEG data	• Non-invasively capture temporal dynamics of brain activity so that it can be used during CAD task performance • Spatial resolution at which EEG is captured allows observing the changes in CL; signal should be captured from the entire cortex, from at least 14 spatially distributed channels to allow calculating suggested EEG features • Sampling frequency allows observing changes in CL; sampling frequency of at least 128 Hz is required to analyse spectrum up to 45 Hz • EEG data must be timestamped to allow task segmentation, TRP calculation, and alignment with behavioural data and data related to CAD modelling process • EEG data must be gathered from the baseline tasks in addition to the CAD modelling tasks to allow TRP calculation
Behavioural data	• Contain information about keyboard strokes, required for building a HCI script • Contain information about mouse movements and button clicks, required for building a HCI script • Contain information about transitions between the windows/screens if there is more than one available to the study participants, required for distinguishing unit task phases • Data must be timestamped to allow segmentation and alignment with EEG data

CAD modelling outputs	• Contain information about geometric (sketch) entities and relations, required for quantifying outputs • Contain information about sketches, required for quantifying outputs • Contain information about sketch definition, required for quantifying outputs • Contain info about CAD features, required for quantifying outputs
CAD modelling process	• Timestamped screen captures are required to allow segmentation from the CAD events to unit tasks
Subjective CL assessment data	• Data reporting self-reported subjective CL scores is required for validation

Table 57 also clarifies practical requirements that follow from the measurement principles reviewed in Chapter 2 and the conducted empirical study (Chapter 3). For EEG, sufficient spatial coverage (e.g. full-cortex coverage with distributed channels) and sampling frequency are required to compute frequency-domain features within the bands of interest and to preserve temporal sensitivity. Timestamping and baseline recordings are necessary to support synchronisation, task-related normalisation, and segmentation approaches used in EEG-based CL analysis. For behavioural and process data, timestamping is similarly essential because the method relies on aligning HCI events and task segments to operator utilisation estimates and EEG signal as well as on calculation task segment durations.

5.2 Steps to model and analyse engineering designer's CAD performance

A process-based model of engineering designer's CAD performance is built in three steps depicted in Figure 45: 1) analysing HCI during CAD modelling to generate a timestamped HCI script of CAD events, 2) analysing a structure of a CAD task to generate a timestamped hierarchical decomposition into CAD actions, CAD operations, unit task operations, and information-processing phases (unit task phases), and 3) mapping the P, C, and M operators to generate the process-based cognitive model of CAD performance.

There are two input data requirements for building the cognitive model of CAD performance: 1) screen capture video of the CAD modelling process, and 2) behavioural data that includes information on a keyboard strokes, mouse movements and button presses. The video should specify in detail the engineering designer's interaction sequence with a CAD system throughout the execution of the predefined CAD modelling tasks, and observable transformation of CAD model information throughout the process, while behavioural data provides data on engineering designers' interaction with HCI tools.

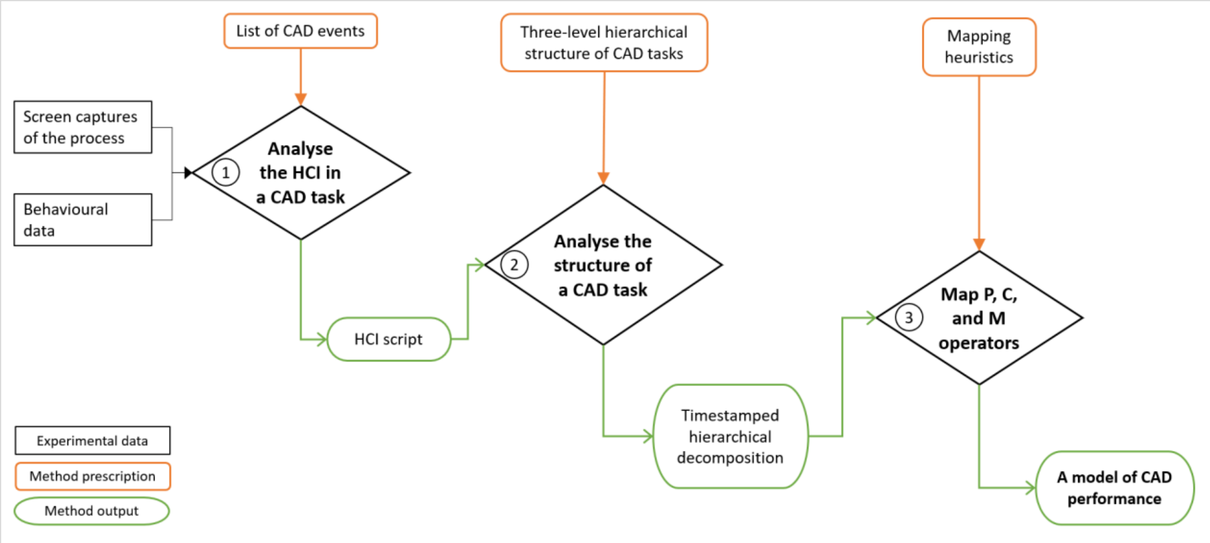

Figure 45 Steps to build a model of CAD performance

The list of CAD events, the three-level hierarchical structure of CAD tasks, and the mapping heuristics are used as input prescribed by the method; they are defined and explained as part of the theoretical model of CAD performance in Section 4.2. They are referred to in this section as needed, but are not reproduced here again in full to avoid repetition.

5.2.1 Analyse HCI to create a HCI script

The first step in analysing a CAD task is to extract CAD events through the event-based coding activity of the screen captures. The CAD events are listed and explained in Table 58; the list includes CAD actions, CAD system responses, and pauses.

Table 58 List of CAD events

CAD event	Description/Example
*Moving a cursor	Moving a cursor on the screen.
*Clicking on an icon/feature/textbox	Clicking a mouse button to select an icon/feature the cursor is positioned over.
*Hovering over (pointing to) an icon/feature/point	Positioning a cursor over an icon/feature/point without selecting it.
*Typing a value	Striking keyboard key(s) to input a value.
*Entering (striking) a command	Striking keyboard command key, such as Escape, Delete, Enter, Undo/Redo, etc.
*Rotating a view	Rotating a model view about the specified point by the specified angles in the directions of the screen X and Y axes [56].
*Zooming in/out	Modifying the zoom factor for the model view [56].
*Dragging a geometric entity	Pressing a mouse button to select a geometric entity and moving a mouse to change its position in the virtual global coordinate system.
CAD system processing HCI input	Waiting for a CAD system to process (complete) action requested through the input to HCI tools and update CAD application.
Pause	No visible CAD action not CAD system processing.

*Star marks CAD events that are also CAD actions

The final HCI script contains the stream of CAD events, with specified name, and start and end time, as exemplified with an excerpt shown in Table 59. CAD events occur in a strict sequence.

The exception is a CAD system response that may occur in a parallel with some of the CAD actions (e.g. *moving a cursor* or *hovering over* an icon/feature).

Table 59 An example of a HCI script

N	CAD event	Begin Time	End Time
1	Moving a cursor	41:11.2	41:12.9
2	Hover over Top plane	41:12.9	41:13.6
3	Click on Top plane	41:13.6	41:13.6
4	CAD system response	41:13.6	41:14.0
5	Moving a cursor	41:14.0	41:15.6
6	Hover over Sketch tab	41:15.6	41:16.3
7	Click on Sketch tab	41:16.3	41:16.3
8	Moving a cursor	41:16.3	41:16.8
9	Hover over Sketch	41:16.8	41:17.0
10	Click on Sketch	41:17.0	41:17.0
11	CAD system response	41:17.1	41:17.5
12	Moving a cursor	41:17.1	41:17.5
13	Pause	41:17.5	41:19.7
14	Moving a cursor	41:19.7	41:20.7

The purpose of this first step is to convert the raw CAD modelling process into an event-based, time-referenced description of *what happened when*, capturing both engineering designer CAD actions and system delays. In that way, the developed HCI script represents CAD actions engineering designers perform to accomplish a CAD modelling task, pauses, and CAD system response periods, and establishes the temporal backbone required for later structuring, operator mapping, and utilisation calculations.

The HCI script is later used as the input for 1) grouping CAD events into task segments (Section 5.2.2), 2) mapping CAD events to P, C, and M operators with correct timing and concurrency (Section 5.2.3), and 3) computing task segment durations and operator activity durations required for utilisations and CL estimation (Section 5.3.1).

5.2.2 Analyse the structure of a CAD task to create a timestamped hierarchical decomposition into subtask clusters and unit tasks

This step consists of four coding rounds. The first three follow the goal/task hierarchy and its aligned process analysis layers, while the fourth codes unit task phases, which are orthogonal to the hierarchy, at the level of unit tasks.

1) Low-level process coding (CAD events → CAD actions)

The raw CAD event stream is first coded into CAD actions, creating a consistent, interpretable representation of basic interactions (e.g. moving a cursor, clicking, rotating, typing).

2) Intermediate-level process/goal coding (CAD actions → CAD operations)

Next, CAD actions are grouped into clusters of CAD operations through which intermediate-level subtasks/subgoals are reached (e.g. navigating CAD application, drawing, dimensioning). The list of subtask clusters and their definitions is provided in Table 53.

3) High-level goal coding (CAD operations → unit task operations)

The intermediate-level subtasks are then grouped into unit task operations through which unit tasks (high-level subtasks/subgoals) are reached (see Section 4.2.2.1 for detailed description).

4) Unit task phases coding (unit tasks → acquisition, execution, inspection)

Finally, each unit task is coded into acquisition, execution, and inspection phases (see Table 55). Importantly, these phases do not constitute another layer of the goal/task hierarchy nor a separate process analysis layer; instead, they capture recurring information-processing modes that can occur within unit tasks regardless of the specific operations used.

Figure 46 depicts a coding example of one CAD operation (*Drawing*), which consists of CAD actions (*Moving a cursor* and *Clicking on an icon/feature/textbox*). CAD actions were coded from the stream of CAD events, which in this case consists of these two CAD actions and pauses between them. The layer of unit task operations shows that *Drawing* belongs to *Sketching* unit task, and *Execution* as its information-processing phase. Two figures below the timeline in Figure 46 represent the start and the end state of the CAD system interface and the built 3D model.

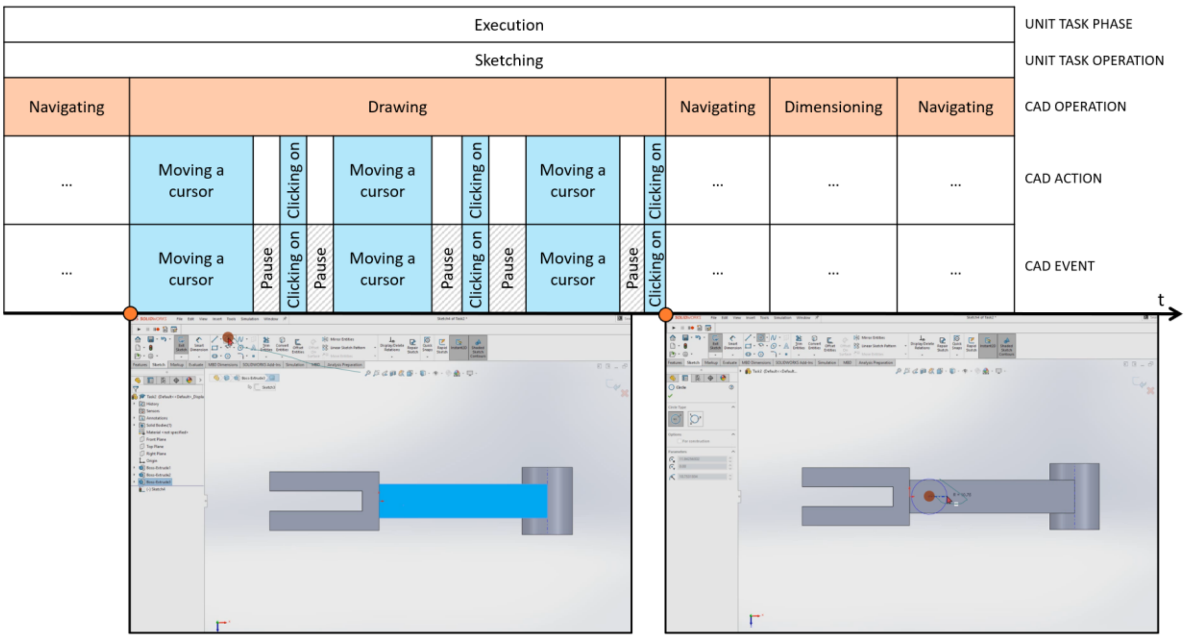

Figure 46 An example of a decomposition

The purpose of this step of the method is therefore to transform a fine-grained CAD event stream into a goal- and process- aligned structure that is interpretable, comparable across participants, and suitable for segment-level quantification. Individual CAD events (and CAD actions) are often too granular and numerous to support meaningful interpretation of CAD performance and CL. Grouping is therefore used to 1) align observed CAD actions with intermediate goals (subtasks), 2) reduce sensitivity to small differences in how individual participants execute the same subtask at the CAD action level, 3) enable aggregation of operator activity into comparable segments, and 4) define the segment set needed later for segment-wise EEG comparison and information-processing weight assignment (Section 5.3.3).

The output of this step is a timestamped hierarchical decomposition spanning CAD events and CAD actions, CAD operations (subtask clusters), and unit task operations, augmented with unit task phases. This decomposition defines the task segments over which in later steps 1) operator utilisations are computed (Section 5.3.1), 2) EEG-based CL indicator is aggregated and compared across segments (Sections 5.3.2 and 5.3.3), 3) information-processing weights are assigned (Section 5.3.3), and 4) CL is visualised against the CAD modelling process at multiple abstraction levels (Sections 5.3.4 and 5.3.5).

5.2.3 Map the motor, perceptual, and cognitive operators to the CAD events

CAD actions observed during CAD modelling are supported by perceptual, cognitive, and motor processing. The corresponding operator set (P, C, M) and CAD system response operator (R), together with their associated operations, are listed in Table 52. The purpose of this step is to convert the coded CAD event stream into an interpretable representation of engineering designers' embodied cognition by specifying which operators are active and for how long.

Figure 44 presents the conceptual mapping structure at the level of CAD events and unit task phases. Table 60 operationalises this structure within the method by providing detailed heuristics for assigning P, C, M, and R to CAD events, while additionally accounting for contextual cues (e.g. CAD actions and subtask clusters) and execution characteristics (slow vs. fast execution). The rules are grounded in psychological assumptions stemming from Card et al. [92] and therefore function as heuristics that guide mapping rather than direct measurements of internal cognitive states.

Table 60 Heuristics for assigning P, C, M, and R operators

Operator	CAD action/Subtask cluster/Unit task phase	Heuristic
R	CAD system response	Entire duration of every CAD system response event
M	Moving a cursor	Entire duration of the event.
	Clicking on	Entire duration of the event.
	Typing	Entire duration of the event.
	Entering a command	Entire duration of the event.
	Rotating	Entire duration of the event.
	Zooming in/out	Entire duration of the event.
	Dragging	Entire duration of the event.
P	Zooming in/out as a part of a Manipulating subtask in the execution part of a unit task	Target identification is assumed.
	Rotating as a part of a Manipulating subtask in the execution part of a unit task	Target identification is assumed if rotation is fast.
P + C	Entire duration of an acquisition part of a unit task	Analysis/Evaluation/Understanding/Creating is assumed.
	Entire duration of an inspection part of a unit task	Analysis/Evaluation is assumed.
	Pause between subtasks during an execution part of a unit task	Cognitive and perceptual processing is assumed. This reflects (typically short) transitions between the subtasks. If longer in duration, the pause belongs either to Acquisition or Inspection, not to Execution.
	Before and after every click on event	Assumed target identification and evaluation after it (mapping the location of an icon as a target and the cursor). Assumed attention shift after the selection.
	Before, during, and after every typing event	Assumed attention shift (between the screen and the keyboard) and target identification (keys).
	Before and after every entering a command event	Assumed attention shift (between the screen and the keyboard) and target identification (command keys).
	During every hover over event	Target identification and evaluation are always assumed. Analysis/Evaluation is assumed if hovering is long.
	Moving a cursor	Analysis/Evaluation is assumed if the cursor movement is slow.
	Moving a cursor	Visual search is assumed if the cursor movement is fast but intermittent with brief hovering over icons/features.
	Dragging	Analysis/Evaluation is assumed if dragging is slow.
	Rotating as a part of a Manipulating subtask in the execution part of a unit task	Analysis/Evaluation is assumed if the rotation is slow.
	Zooming in/out as a part of a Manipulating subtask in the execution part of a unit task	Analysis/Evaluation is assumed if zooming in/out is slow.

Operator parallelism is permitted when different processors are involved and when no output–input dependency exists between concurrent operators, consistent with CPM-GOMS assumptions [94] and experimental demonstrations by Gong and Kieras [96]. This principle underlies the concurrent assignments in Table 60 (e.g. “P+C” during acquisition/inspection and “P+C” episodes around actions that require target identification or evaluation).

In the mapping, M is assigned to the full duration of any physically executed CAD action (mouse/keyboard/view manipulation/dragging), and R is assigned to the full duration of identifiable CAD system response periods in which the engineering designer must wait. P and C are assigned according to phase- and action-dependent heuristics. For example, acquisition and inspection are modelled as continuous P+C engagement because they primarily consist of perceptual encoding coupled with interpretation/evaluation. During execution, P and C are assigned more selectively; they are positioned around actions that require target identification, attention shifts, evaluation of intermediate results, or goal-directed search (e.g. before/after selections, during extended hovering, during slow cursor movements, or during slow/controlled view manipulation). Short pauses between subtasks within execution are treated as brief P+C transitions; when such pauses become longer and functionally shift toward preparation or verification, they are classified as acquisition or inspection rather than execution, consistent with the unit task phase definitions.

Because “slow vs. fast” and “long vs. short” execution is used as a cue in the heuristics, these distinctions must be operationalised consistently in the coding procedure (e.g. relative to the observed distribution of durations for a given action type and participant, and/or using predefined cut-offs established in the exploratory study) and applied uniformly across participants and CAD action types. The intent is not to claim that a particular duration threshold uniquely determines cognition, but to apply a reproducible rule for when a CAD action is treated as primarily motoric versus when it reflects additional cognitive–perceptual processing (e.g. evaluation or visual search).

Figure 47 depicts an example of mapping the operators to the CAD events whose coding was shown in Figure 46. In this case, P and C operators are assigned to the pauses, while M operators are assigned to the CAD actions. No additional P and C operators are assigned to CAD actions since the execution of *Moving a cursor* was fast and straightforward in this example.

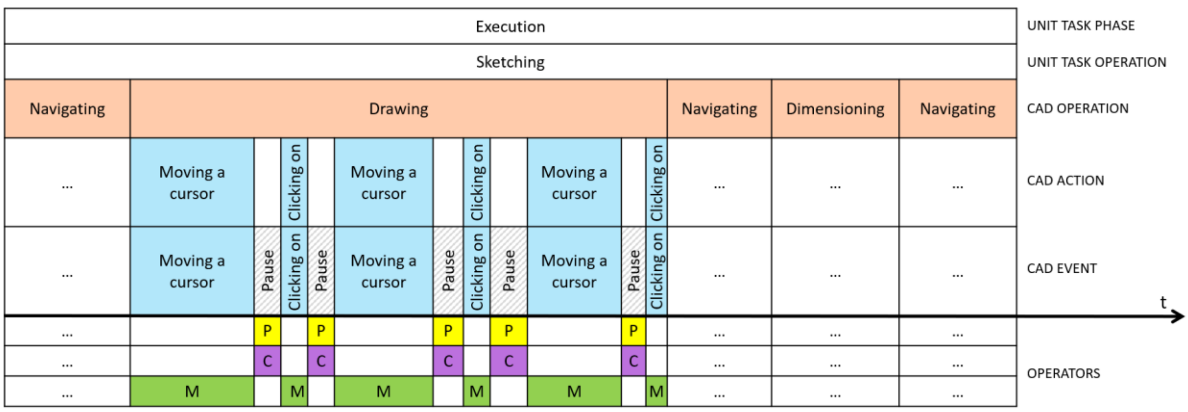

Figure 47 An example of mapping the operators

The output of this step is a process-based cognitive model of CAD performance; a time-aligned operator stream that specifies the sequence and concurrency of P, C, and M operators across the CAD modelling session. The operator timeline is the direct input for calculating operator activity durations and utilisations (Section 5.3.1).

5.2.4 Visualise the model of CAD performance

Multi-layered task performance timeline (raster plot) is suggested for the integrated model visualisation (Figure 48), where individual CAD events/actions, subtask clusters, unit task phases, unit tasks, and P, C, and M operators are aligned along a common temporal axis to represent their sequence, concurrency, and proportions during the CAD modelling session. Similar visualisation type was used in previous related CAD studies (e.g. [33], [58], [222]). This multi-layered task performance timeline becomes the reference frame onto which dynamic CL traces are overlaid (Section 5.3.4), enabling interpretation of CL fluctuations in relation to the process-based CAD performance.

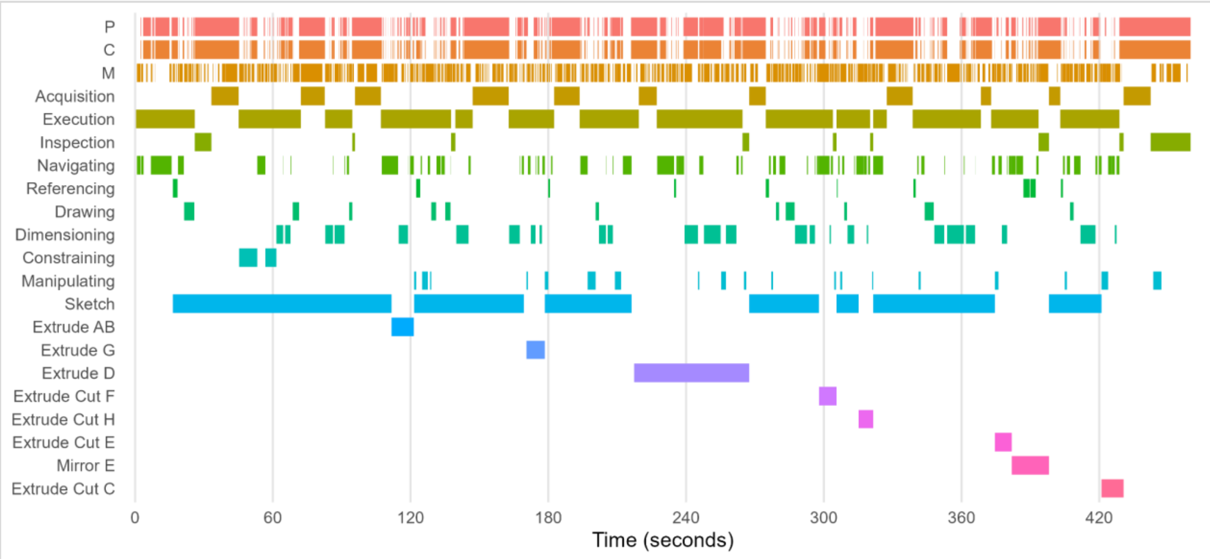

Figure 48 An example of the model of CAD performance

5.3 Steps to model and analyse dynamic changes in CL

Once the model of CAD performance is available, the model of CL is built upon it in four steps visualised in Figure 49. First, the utilisation of the P, C, M operators is calculated for a given period (segment) from the model of CAD performance (step 1 in Figure 49, detailed in Section 5.3.1). Second, EEG-based CL indicators applicable for the CAD context are identified (step 2 in Figure 49, detailed in Section 5.3.2). Then in the third step, the information-processing weights are assigned to the utilisations of the P and C operators, based on the differences in brain activity captured via identified EEG-based CL indicator (detailed in Section 5.3.3). The utilisations and weights are input to the suggested equation for calculating CL in the fourth step (detailed in Section 5.3.4). Finally, visualising dynamic changes in CL over a model of CAD performance gives a model of CL in a CAD modelling activity (detailed in Section 5.3.5).

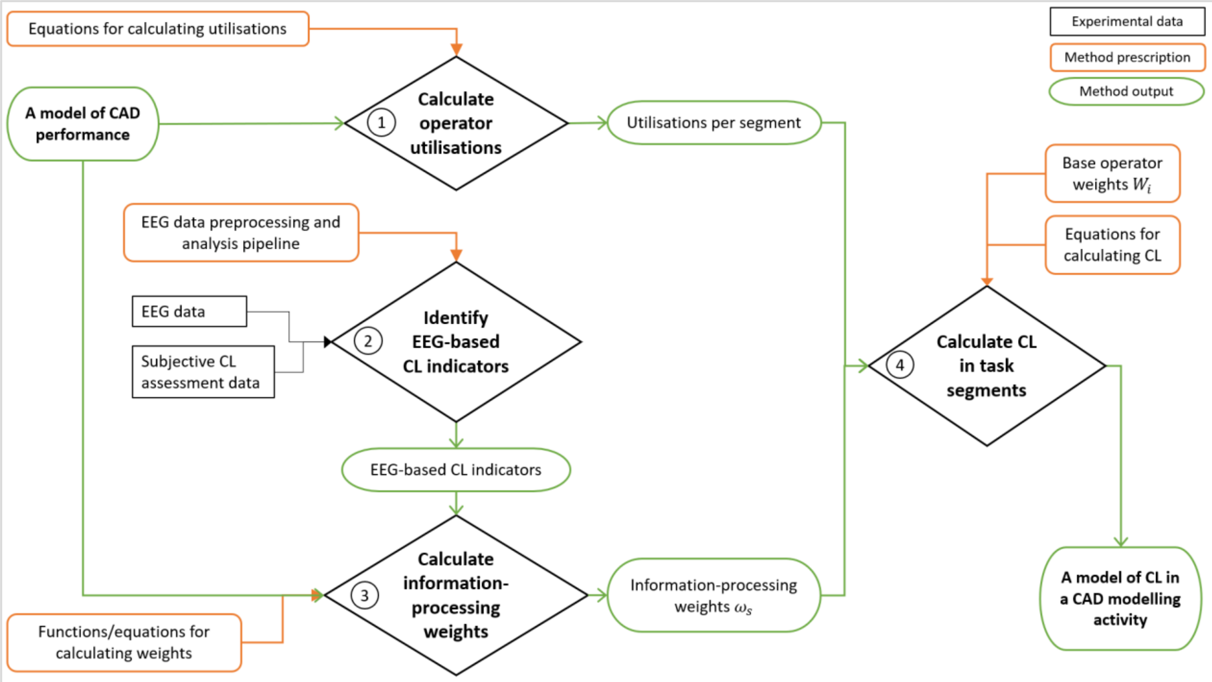

Figure 49 Steps to build a model of CL in CAD

In addition to the model of CAD performance (the output of steps described in Section 5.2), there are two input (experimental) data requirements for building the model of CL: 1) EEG data, and 2) subjective CL assessment data.

5.3.1 Calculate utilisations of the operators

In the first step, distribution of perceptual and cognitive operators' utilisation in each task segment *s* is calculated as a ratio of total time when a particular operator *i* is active $A_i(t)$ during a task segment (unit task/unit task phase/intermediate-level subtask) and the total duration T_s of the particular segment *s* [115], as shown in Equation (39):

$$U_{i,s} = \frac{1}{\mathrm{T}_s} \int_0^{\mathrm{T}_s} A_i(t) \cdot dt \,, \qquad (39)$$

T_s is duration of a task segment s

$$U_{i,s} \in [0,1], \qquad i = P, C$$

The purpose of this step is to quantify how long the CAD modelling process engages P and C resources relative to time available within a segment, thus allowing comparison across segments and engineering designers even when segments differ in absolute duration.

Furthermore, dynamic changes in operator's utilisation can be calculated using the same equation, but with a moving window. Segment-wise utilisation provides discrete values for unit tasks/subtasks; however, CL is expected to vary continuously and may change within long task segments or across short transitions. A moving window produces a continuous, time-references utilisation trace that can be 1) aligned with EEG-based time series, 2) used to localise peaks/troughs in processing demands, and 3) overlaid on the CAD performance timeline for process interpretation (Section 5.2.5). Considering the duration of unit task, unit task phases (see Table 55), and CAD operations (see Table 53), here suggested parameters of a moving window are its duration of 10 seconds with a step of two seconds. A 10-second time window is a practical compromise between stability of utilisation estimates (operator activity typically occurs to avoid highly volatile ratios) and temporal resolution (still sensitive to changes across CAD modelling segments). The 2-second step provides a smooth trace via overlapping windows while preserving sensitivity to transitions between CAD actions and intermediate-level subtasks. The window length and step values are therefore chosen to balance robustness and interpretability given typical durations of unit task phases and CAD operations (Table 55 and Table 53), and to support time-aligned comparison with EEG-based signal.

5.3.2 Identify EEG-based CL indicators applicable for the CAD context

As an initial step toward identifying EEG-based CL indicators applicable for the CAD context, a list of the most promising EEG features for distinguishing CL levels was extracted from the literature. This approach helped narrow down the vast number of potential EEG features that could describe CL to nine of the most frequently used features across studies (Table 61). These nine EEG features were selected because they were reported as sensitive to variations in CL by at least two different relevant studies (presented in Section 2.4.3). The selected EEG features are categorized into two groups: frequency band TRPs, and the ratios of frequency bands' TRPs.

Table 61 List of EEG features emerging from literature

	Frequency band	**Cortical area**	**References**
Frequency band TRP	Theta (4-8 Hz)	Entire	[164], [168], [176]
		Frontal	[25], [126], [130], [137], [141], [168], [172], [175]
	Alpha (8-13 Hz)	Entire	[136], [137], [141], [164], [172], [176]
		Parietal	[25], [168], [175], [176]
		Occipital	[25], [168], [176]
		Pz	[172], [173], [176]
	Beta (13-30 Hz)	Entire	[25], [141], [164], [176], [223]
Ratio	Theta/Alpha	Entire	[103], [171], [174], [175]
		Fz theta; Pz alpha	[133], [170], [177]

To use EEG as a CL-sensitive input in the prescribed method, EEG features shall be treated as candidate indicators that require context-specific validation before being used for process-level CL interpretation in CAD. This requirement follows from two well-established premises. First, CL is a covert construct inferred indirectly, and triangulation across complementary measures is recommended because different methods may capture different dimensions of CL and operate at different temporal granularities [13], [101]–[103], [105]. Second, the EEG literature shows that CL-sensitive features identified in cognitively controlled tasks do not transfer uniformly to more complex HCI tasks, including CAD-related activities; the set of sensitive features and even the direction of effects can vary with underlying processing and mechanisms used to solve the task. Consequently, the method does not assume a universal EEG-based CL indicator. Instead, it requires an empirical step that identifies which feature(s) are suitable for the specific CAD task and setup to which the method is applied. Therefore, which of EEG features from Table 61, if any, may best serve as the EEG-based CL indicators should be explored through the experimental study designed primarily for this purpose. Experimental tasks in such a study should be CAD tasks that intentionally impose distinct CL on engineering designers. Other variables should be controlled and kept the same in all the experimental tasks to reduce confounding effects. In addition to the CAD tasks, EEG should be captured during the baseline tasks, which typically require participants to remain still and observe a simple object (e.g. a cross or a dot) presented at the center of the screen. An alternative may be any other controlled task. Participants of the experimental study should be engineering designers with similar characteristics – primarily concerning their CAD experience and expertise. Experimental tasks should be validated with common subjective assessment method, such as NASA TLX to confirm the difference in CL levels they impose on engineering designers. Similarly, NASA TLX is suggested for a validation of the suitability of the EEG features that may significantly differ between the conditions; this is typically done with the correlation analysis.

To conclude, the best candidate for the role of EEG-base CL indicator is defined operationally rather than universally based on whether and how well it satisfies the following two criteria within an experimental study. First, it shows reliable differences between CAD tasks intentionally designed to impose at least two levels of CL (i.e. the EEG feature discriminates conditions with meaningful effect size, not only statistical significance). Second, it aligns with an established CL criterion (most commonly NASA TLX) at the appropriate level of aggregation, demonstrating the expected association across participants and/or conditions when such convergence is theoretically warranted.

The selected EEG-based indicator is then computed per task condition for validation, and per task segment (from the hierarchical decomposition) to test segment differences and derive information-processing weights (Section 5.2.3) and/or to provide an independent dynamic CL trace for triangulation (Section 5.2.4).

5.3.2.1 EEG data pre-processing pipeline

Here proposed pre-preprocessing script (see Figure 50) is inspired and adapted from the pipelines used in prior EEG studies within engineering design, such as those described by Li et al. [180] and Vieira et al. [150]. It is detailed for pre-processing in MATLAB employing the EEGLAB toolbox [202] although the same approach with the corresponding functions can be applied in other applications.

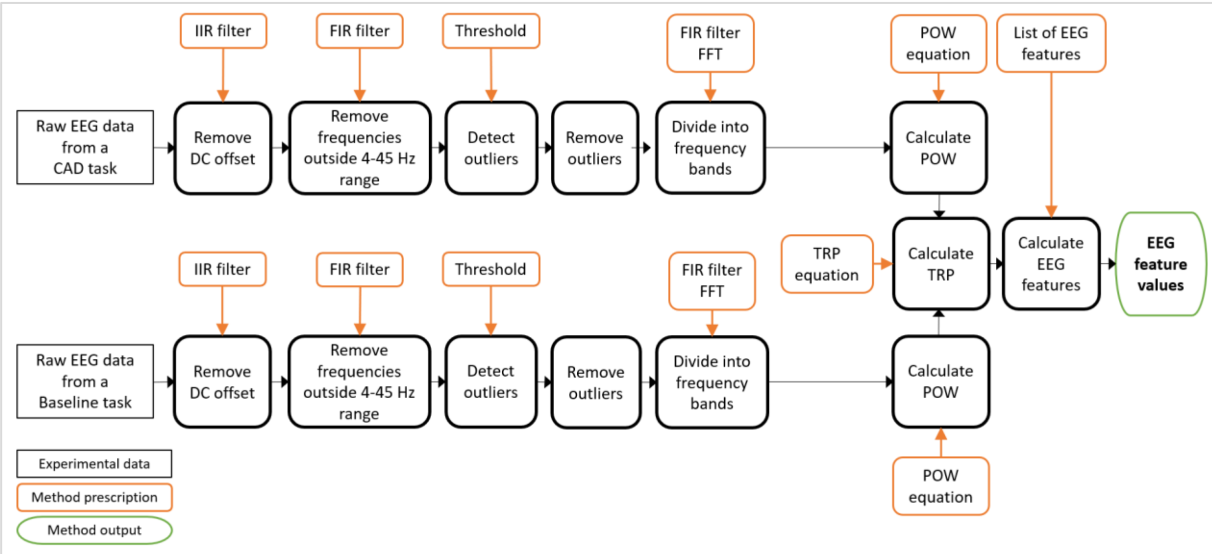

Figure 50 EEG data pre-processing pipeline

In the first step, the DC offset, which is specific to the EEG device employed in the experiment, is removed. This can be done with an infinite impulse response (IIR) filter [202].

After that, frequencies outside the 4-45 Hz range are filtered out with a finite impulse response (FIR) filter [202]. The 4-45 Hz frequency range is suggested for three reasons. First, this range of frequency bands can be analysed even with the sampling frequency of 128 Hz, according to the Nyquist-Shannon theorem [193]. Second, delta band (0.5-4 Hz) is excluded from the analysis as it gets highly contaminated by muscle movements [180], which are unavoidable part of the CAD tasks as their execution requires participants to use HCI tools. Third, the range of frequency bands to which the further analysis focuses depends on the construct in the focus of the investigation. When considering CL, the analysis is proposed to focus on theta (4-8 Hz), alpha (8-13 Hz), and beta (13-30 Hz) frequency bands since they emerged from literature as EEG features indicative of changes in CL (see Table 61).

In the next step, epochs of EEG signals contaminated by artifacts such as eye blinks, heartbeat, electrical interference, and muscle movements are identified. First, EEG signal is searched for outliers whose EEG amplitudes exceed a threshold of ±100 µV, set in accordance with the observations from clinical research on differences between normal and contaminated EEG signals [224]. Second, any one-second epoch (shifted by 1/128 seconds) containing at least one amplitude surpassing the threshold across the channels is subsequently discarded.

Following the removal of outliers, EEG data is divided into three frequency bands: theta (4-8 Hz), alpha (8-13 Hz), and beta (13-30 Hz), using FIR filtering and Fast-Fourier Transformation [148].

The absolute power (POW) of the EEG signals then can be computed as the magnitude-squared mean (M) of the band-filtered values (expressed in microvolts squared; μV^2). These pre-processing steps are conducted for each CAD and baseline task.

Next, task-related power (TRP) is calculated by subtracting the log-transformed POW average for a subject j at an electrode i during a baseline task (preceding the CAD task) from the log-transformed POW average for the same subject j at the same electrode i during the CAD task, following Equation (32) [225]. Given the TRP equation, positive TRP values reflect an increase in relative POW during the execution of the CAD task, while negative TRP values signify a relative decrease in POW.

Finally, EEG features are calculated from the TRP values for each task and the participant.

5.3.2.2 EEG data analysis pipeline

Data analysis includes both descriptive and inferential statistics, as suggested in Figure 51. Measures of descriptive statistics are calculated to summarize and describe the values of the EEG features in all task conditions. Moreover, inferential tests are conducted to quantitatively compare differences in EEG feature values between the task conditions. This is done by calculating the significance of statistical differences (expressed through p-values) and effect sizes of the calculated differences, using the appropriate statistical tests.

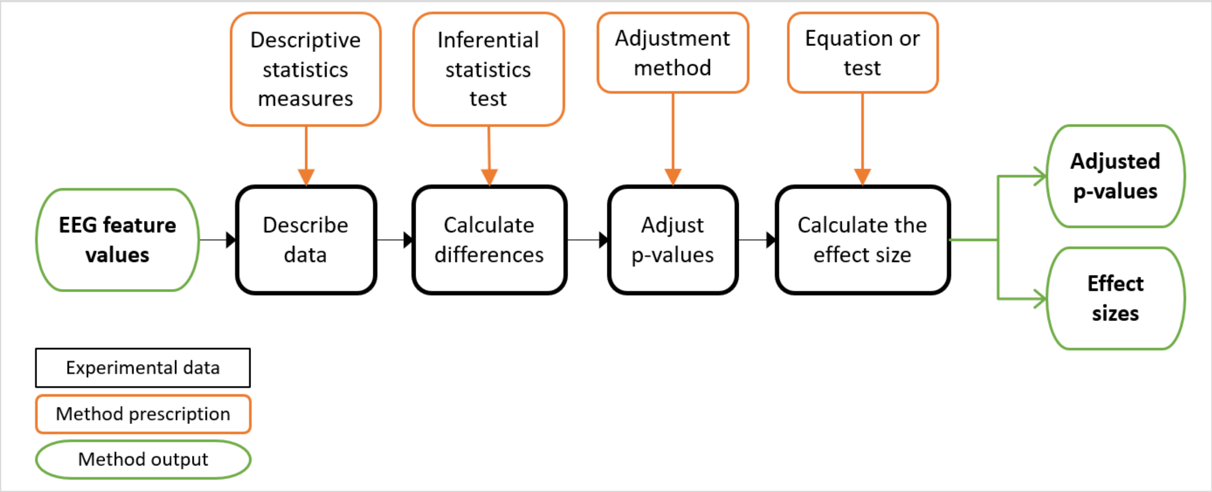

Figure 51 EEG data analysis pipeline

5.3.2.2.1 Describe EEG data

Descriptive statistics encompasses the calculation of the M and the Med as measures of central tendency, as well as SD and MAD as measures of variability for each experimental variable. Additionally, data description includes testing the assumptions on normality of a distribution, which guides the selection of inferential tests in the next step of the analysis. Assumption of data normality is typically examined by Shapiro-Wilk normality test ($p > 0.05$).

5.3.2.2.2 Calculate differences in EEG features between the task conditions

EEG feature values are compared using inferential statistics to test their sensitivity to changes in CL imposed by CAD tasks (and influencing factors). EEG features are considered sensitive to changes in CL if their values significantly differ between the task conditions ($p \leq 0.05$).

There are two main decisions to be made when selecting the appropriate inferential test for comparing the distributions, as suggested in Figure 52:

- Between-subject or within-subject tests;
- Parametric or non-parametric tests.

Distributions of EEG feature values are suggested to be compared participant/wise, using the repeated measures (within-subjects/paired) type of inferential tests. This suggestion is based on the premise that an absolute value of brain activity may differ between the participants (due to various factors), which is here addressed by relative comparison of the brain activity of each participant across the conditions.

Inferential statistics distinguishes between parametric and non-parametric tests. Validity of the parametric statistical tests depends on the distribution of data; these tests assume that and are valid when the data are normally distributed. Normality is often examined with Shapiro-Wilk test and confirmed when it results in $p > 0.05$. If the normality assumption is violated ($p \leq 0.05$), non-parametric version of tests should be employed instead. It is to be noted that normality is tested on a single sample of paired differences between the compared data sets (instead of two individual data sets) when conducting the within-subjects (paired/repeated measures) analysis. Additionally, the assumption of equal variances between groups does not exist in cases with two task conditions. However, sphericity assumption, which assumes that variances of the paired differences of repeated measures are equal, should be checked when comparing distributions in three or more conditions. This is usually done with Mauchly's test and corrected with a Greenhouse-Geisser correction if violated.

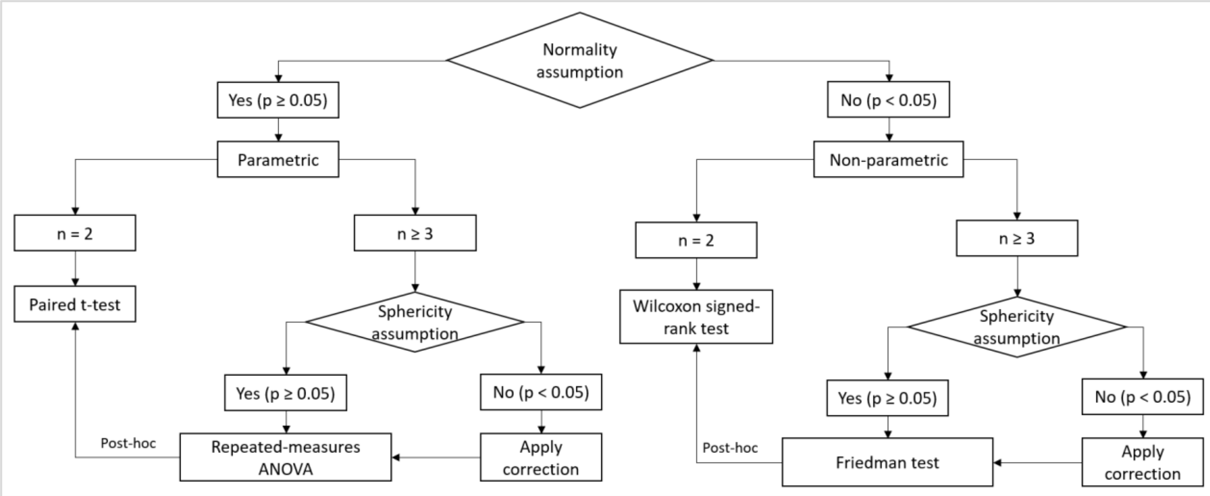

Figure 52 Selection of an inferential test

5.3.2.2.3 Adjust p-values

The obtained p-values should be corrected due to the relatively large number of tests conducted on the same data set (nine EEG features; Table 61) to mitigate the probability of a type I error (which increases with the number of tests). One of the commonly used and the most conservative corrections is the Bonferroni adjustment, which divides the initially used significance level (typically 0.05) by the number of conducted tests.

5.3.2.2.4 Calculate effect sizes

Effect size represents the strength (magnitude) of the difference between the variables, which stands for the effect of an independent variable on a dependent variable. In that sense, effect sizes of the differences in EEG feature values are analysed to identify those features that may serve as prominent EEG-based indicators in CL in the CAD context; larger effect size thus implies higher suitability of the particular EEG feature for indicating CL changes.

Calculation of the effect size of tested differences is associated with the particular test employed for the analysis. Effect sizes for the paired t-test are calculated using Cohen's d, representing the number of standard deviations between the means of the compared variables. The effect size for the Wilcoxon test is calculated as z-statistic divided by the square root of the sample size.

5.3.2.3 NASA TLX data analysis pipeline

NASA TLX data analysis pipeline follows the steps described by the EEG data analysis pipeline (see Figure 51). Descriptive statistics measures are calculated for the overall WL and scores of six individual NASA TLX dimensions across the task conditions. Inferential tests are then conducted to quantitatively compare differences in NASA TLX scores. These scores are compared between the conditions to validate the experimental tasks by statistically confirming differences in overall WL and CL imposed on the participants. Effect sizes of the differences in NASA TLX scores between the task conditions are then (qualitatively) compared across the NASA TLX dimensions. This qualitative comparison should point to the primary contributor to the differences in the WL and CL between the task conditions.

5.3.2.4 Correlation analysis

Correlation analysis measures the strength of a relationship between two continuous variables; Pearson's correlation is used as a parametric test, and Spearman's rank correlation as its non-parametric alternative. Importantly, repeated-measure correlation analysis [226] should be used when calculating correlations using data from the same participants but in two task conditions to prevent violating the assumption of independence of observations.

Within the proposed method, correlation coefficients are computed between NASA TLX components and EEG features to identify the best candidate for describing the CL variations imposed by CAD tasks. To be considered the best candidate, an EEG feature has to 1) significantly correlate with the perceived level of CL, as self-reported by participants on the mental demand component of the NASA TLX, and 2) have the highest correlation coefficient.

If, in a given setting, EEG features do not correlate with NASA TLX, this outcome shall be interpreted as diagnostically informative rather than as an automatic invalidation of the method. NASA TLX provides a retrospective, task-level summary across multiple WL dimensions, whereas EEG reflects time-varying psychophysiological activity that may be segment-specific and may not map one-to-one onto subjective ratings. Therefore, lack of convergence may indicate that 1) the tested EEG feature is not CL-sensitive in the given CAD context, 2) the manipulation changed demands in a way that affects physiology without being strongly represented in retrospective self-report (or vice versa), or 3) noise sources (e.g. movement or ocular artefacts) reduced correspondence. Under such conditions, the method prescribes iterative refinement: revising the candidate feature set, strengthening or re-validating the complexity manipulation, and/or adding additional validation evidence from process-based indicators (e.g. increased inspection/rework or altered operator utilisations) to support interpretation.

5.3.3 Calculate information-processing weights

Once the most promising EEG feature that may serve as the EEG-based CL indicator is identified, its values can be calculated for each unit task phase and CAD operation (intermediate-level subtask) cluster (CAD task segments available from the model of CAD performance) to explore potential differences in their information processing demands. Steps for exploring differences in EEG feature values among the task segments and calculating information-processing weights based on them are visualised in Figure 53.

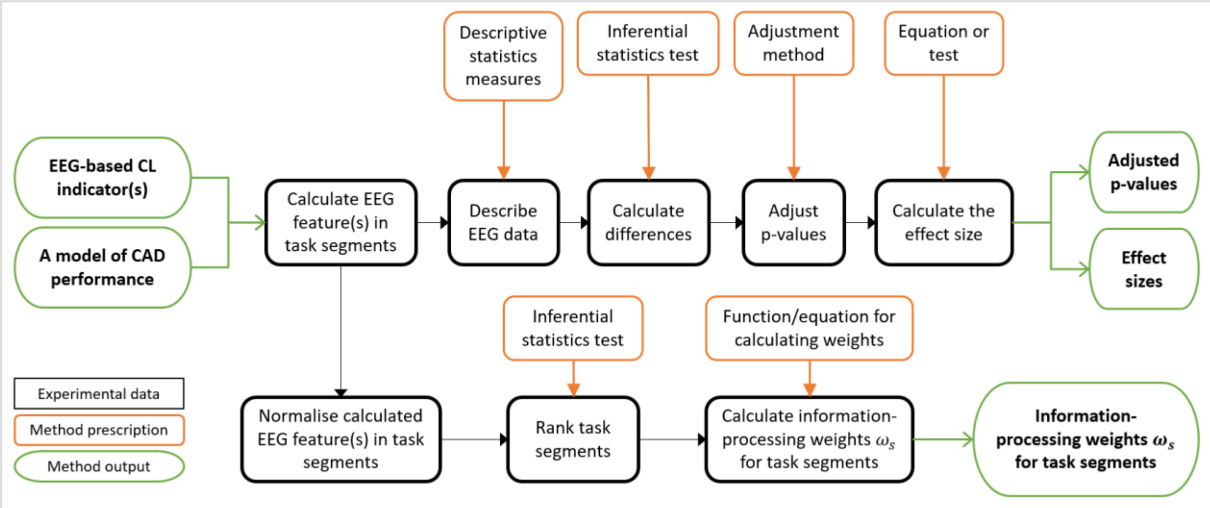

Figure 53 Steps for calculating information-processing weights

If EEG feature values differ significantly, this suggests that segments impose distinct information processing demands that should be reflected when modelling CL. In the proposed method, this is operationalised by assigning segment-specific weights to the utilisation of the P and C operators, so that identical operator utilisation can contribute differently to estimated CL depending on where in the CAD modelling process they occur (to which task segment they belong). This approach follows the broader CL premise that demand varies over time and is unevenly distributed across CAD activities, and that psychophysiological measures, especially EEG, are useful when such temporal variations are expected in complex HCI tasks such as CAD modelling.

Because EEG absolute EEG feature magnitude can vary across individuals (e.g. due to stable inter-individual differences and scaling related to measuring), the method normalises EEG feature value $x_{i,s}$ within participant i across task segment s. This ensures that segment weights reflect each segment's position relative to the participant's own distribution (rank and distance from the participant's typical level). This within-participant normalisation is consistent with the broader EEG practice focusing on task-related differences and within-participant changes rather than absolute values [205]. Accordingly, $x_{i,s}$ is converted to a within-participant robust z-score using the participant's Med and MAD, computed over the set of task segments s included in the analysis (see Equation (40)):

$$z_{i,s} = \frac{x_{i,s} - Med_i}{1.4826 \cdot \mathrm{MAD}_i}. \tag{40}$$

The constant 1.4826 is a correction factor that makes MAD a consistent estimator of the standard deviation under the assumption of normally distributed data. Since the raw MAD corresponds to approximately 0.6745 standard deviations in a normal distribution, multiplying MAD by 1/0.6745≈1.4826 places the robust scale estimate on a standard deviation scale.

In that way, values of EEG features are normalised (standardised) within each participant, while their magnitude and internal (participant-wise) ranking, essential for defining the weights, are preserved. Influence of a direction (positive or negative) is defined "locally"; a robust z-score is positive if higher than participant's typical (Med) EEG feature value, otherwise is negative, regardless of the "raw" EEG feature sign. After that, for each segment s the single value of the robust z-scores can be calculated as a M across the participants. This dataset, if necessary, can be trimmed (e.g. at 10% or 15%) to mitigate the potential influence of extremes.

The final robust z-scores (one for each segment) then can be converted to weights. One suggested way to accomplish this conversion is using *softmax* function [227], described with Equation (41):

$$\omega_s = \frac{e^{z_s}}{\sum_{u \in s} e^{z_u}}, \qquad \omega_s > 0,\ \sum_{s \in S} \omega_s = 1,\ s = \{\text{task segment}\}, \tag{41}$$

where s is the set of task segments, and u is the counter variable. In that way, differences in EEG feature(s) are converted into positive weights that sum in one (1), and are assigned to each task segment within the chosen set (e.g. unit task phases or intermediate-level subtasks). The value of the weights is driven by the position of each task segment (its value of EEG feature) relative to the individual participant's own distribution of that EEG feature values – its rank and distance above or below the participant's typical level (median of EEG feature values across the task segments).

Additionally, the consistency/heterogeneity of the rankings across the participants can be calculated using the Kendall's W, where high W (closer to one) implies consistent ranks and stable global weights, while low W (closer to zero) implies heterogeneity and participants' disagreement in the rankings. Statistical significance is assessed using the chi-square test. Only participants with complete data for all considered task segments should be included in this analysis.

These weights are in the following step multiplied with P and C utilisation values in the CL equation to compute weighted utilisation-based CL, enabling process-aligned CL estimates that account for both duration of engagement (utilisation) and segment intensity (EEG-based weighting).

5.3.4 Calculate CL in task segments

Dynamic changes in CL can be calculated and visualised using three complementary indicators: 1) EEG-based indicators (Equation (42)), 2) utilisations of P and C operators (Equation (43)), and 3) weighted utilisations (Equation (44)), where utilisations from 2) are multiplied by information-processing weights ω_s, empirically extracted from differences in EEG features identified as EEG-based CL indicators.

$$\mathrm{CL}_s^{\mathrm{EEG}} = Med\ (\mathrm{TRP}_s) \tag{42}$$

$\mathrm{TRP_s}$ is a TRP of selected EEG feature in a task segment s

$$\mathrm{CL}_s^{\mathrm{U}} = \sum_i w_i \cdot \frac{\int_0^{\mathrm{T}_s} A_i(t) \cdot dt}{\mathrm{T}_s} = \sum_i w_i \cdot U_{i,s} \tag{43}$$

w_i is a base weight of an operator i, $\quad i = P, C$

T_s is duration of a task segment s

$U_{i,s} \in [0,1]$

$$\mathrm{CL}_s^{\mathrm{U}\omega} = \sum_i \omega_s \cdot w_i \cdot \frac{\int_0^{\mathrm{T}_s} A_i(t) \cdot dt}{\mathrm{T}_s} = \sum_i \omega_s \cdot w_i \cdot U_{i,s} \tag{44}$$

ω_s is information − processing weight of a task segment s

w_i is a base weight of an operator i, $\ i = P, C$

$U_{i,s}$ is utilisation of an operator i, $\ i = P, C$, $\quad U_{i,s} \in [0,1]$

T_s is duration of a task segment s

EEG-based indicators provide a temporally sensitive psychophysiological trace that can reveal fluctuations in information-processing demands even when behavioural durations are ambiguous, but candidate EEG features must be validated for the CAD context. Operator utilisations provide an interpretable, process-aligned behavioural proxy by quantifying how long P and C processing is engaged within defined CAD task segments. They directly link CL estimates to the CAD modelling process structure (unit task phases and intermediate-level subtasks) and can be computed from behavioural and segmentation data even when EEG is unavailable or noisy. However, utilisations primarily reflect duration of engagement, not necessarily the intensity of processing within that duration. The weighted-utilisation indicator integrates both perspectives by scaling utilisation-based CL with segment-level information-processing weights empirically derived from EEG, thereby accounting for both processing duration and segment intensity. Although the weighted utilisations are expected to be most informative when EEG-derived segment differences are stable and interpretable, the other two indicators remain necessary for triangulation and diagnosis of disagreements. CL can be analysed at multiple hierarchical levels (e.g. intermediate-level subtasks), unit task phases and can also be computed continuously using a moving window, whose length and step should be selected with respect to typical segment durations to balance stability and temporal resolution. Equation (44) describes how to calculate CL in each task segment s, while Equation (45) summarises these values into the CL of an entire CAD task. In the letter case, each CL_s is multiplied by time weight to account for potential variations in duration of task segments.

$$CL_{Task} = \sum_{s} \frac{T_s}{T_{Task}} \cdot CL_s = \sum_{s} \frac{T_s}{T_{Task}} \sum_{i} \omega_s \cdot w_i \cdot U_{i,s} \tag{45}$$

T_s is duration of a task segment s

T_{Task} is duration of an entire CAD task

ω_s is information − processing weight of a task segment s

w_i is a base weight of an operator i, $\quad i = P, C$

$U_{i,s}$ is utilisation of an operator i in s, $\; i = P, C,\; U_{i,s} \in [0,1]$

5.3.5 Visualise the model of CL

A model of CL is suggested to be visualised with a timeline representing time on x-axis and dynamic changes in CL on y-axis, in relation to the model of CAD performance at selected abstraction level (considering process layers and task/goal levels). Such a visualisation enables process-level interpretation by showing when CL changes and which CAD modelling segments (intermediate level subtasks/unit tasks parts) those changes correspond to. For example, the CL time series can be visualised and interpreted at the level of unit task phases (Figure 54), moving windows of different lengths (e.g. 10 s and 2 s; Figure 55), or faceted across subtask clusters (Figure 56).

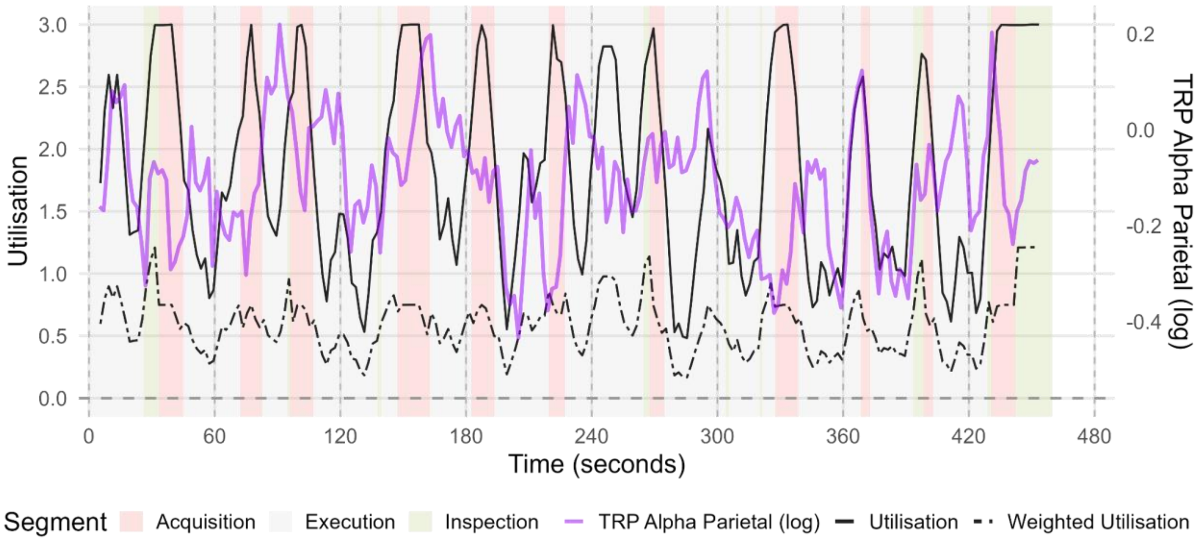

Figure 54 An example of the model of CL at the level of moving windows (10s, 2s)

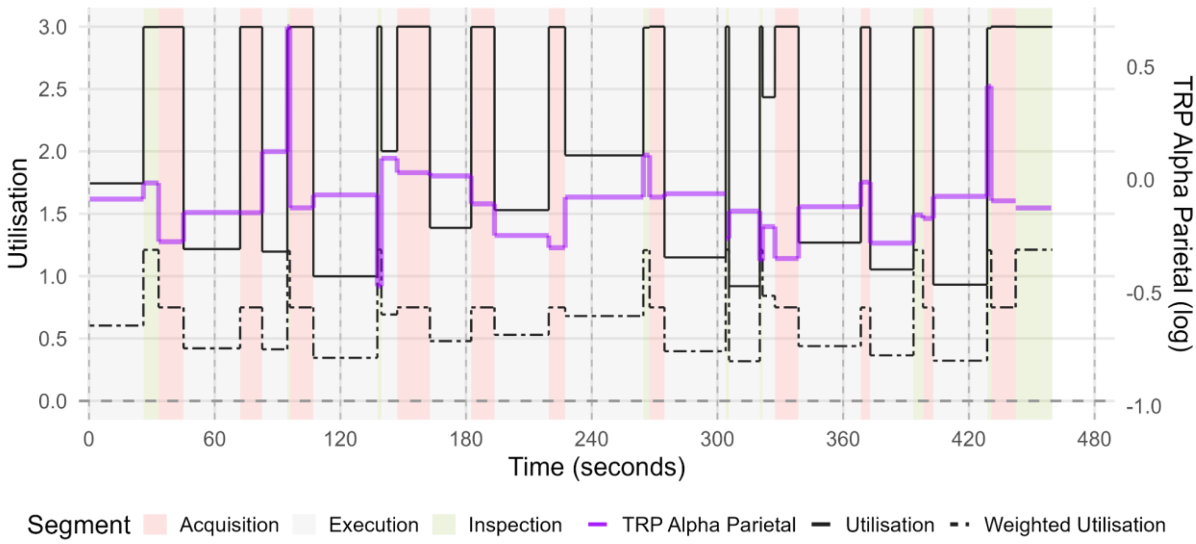

Figure 55 An example of the model of CL at the level of unit task phases

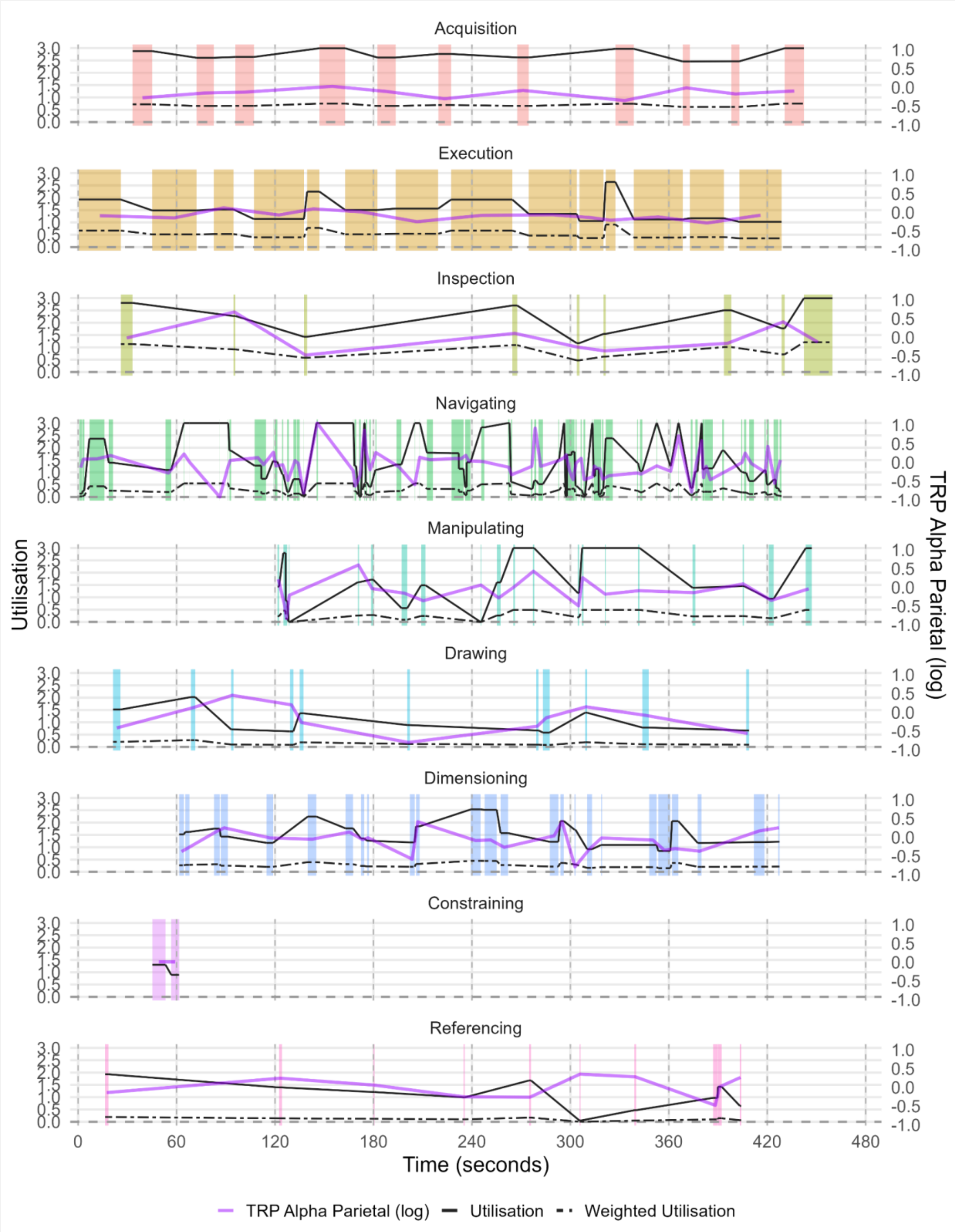

Figure 56 An example of the model of CL faceted across the subtask clusters

This page intentionally left blank

6 THE SECOND EMPIRICAL STUDY

The sixth chapter reports the second empirical study (as a part of Descriptive Study II), which applies and evaluates the proposed method for measuring and analysing CL in CAD modelling tasks performed by professional engineering designers. It therefore operationalises the method's full workflow, linking the process-based CAD performance model, utilisation-based metrics, and psychophysiological evidence.

Accordingly, the study first identifies EEG features the most suitable as CL indicators for CAD tasks of varying complexity (Section 6.2). The selected EEG-based indicator is then used to derive information-processing weights for unit task phases and subtasks (Section 6.3), thereby translating psychophysiological evidence into parameters that can be integrated into the utilisation-based equation introduced in Chapter 5. In parallel, the study builds a CAD performance model for the CAD modelling task as a concrete instantiation of the theoretical model of engineering designers' performance in CAD activities (Section 4.2) and the prescribed modelling procedure (Section 5.1). This operationalisation produces the operator utilisations required by the method and establishes the hierarchical segmentation into unit task phases and intermediate-level subtasks on which subsequent CL analyses are grounded. Once utilisations and EEG-derived information-processing weights are available, the study evaluates how CL estimates behave at multiple levels. Section 6.5 examines how utilisation-based CL indicators relate to subjective CL (NASA TLX), and it compares participant-specific utilisation estimates to simplified global utilisations motivated by their practicality for modelling and visualisation. Section 6.6 then shifts from overall task CL scores to a segment-level perspective by analysing differences in CL across unit task phases and intermediate-level subtasks and quantifying their contributions to overall CL. Section 6.7 extends this analysis to the temporal domain by modelling dynamic CL time series from EEG, utilisations, and weighted utilisation and quantifying their alignment across participants. Finally, Section 6.8 relates utilisation-based CL indicators to overall CAD performance to test the research hypothesis. In this way, the second experimental study and Descriptive Study II provide the empirical basis required to evaluate whether the method functions as intended and under which conditions its indicators and modelling steps provide the most informative view of engineering designers' CL and its relation to performance in CAD activities.

6.1 Study motivation, objectives, and research questions

The second experimental study uses internal task complexity as the primary influencing factor for CL. This choice follows from the literature review conclusions; while subject-related factors (e.g. experience and expertise) and method/tool-related factors (e.g. interface characteristics, representation formats, support methods) can strongly modulate CL, their effects are often context-specific, interacting, and occasionally conflicting across studies. In contrast, task-focused studies provide comparatively clearer and more consistent support that CL increases with internal task complexity, i.e. with the inherent information-processing demands imposed by the task itself. This aligns with generic CL theories (Section 2.4.1.1) that separate an inherent, task-centered component of CL (intrinsic/effective load) from an avoidable component driven by the person and environment (extraneous/ineffective load). For CAD, internal task complexity corresponds to what is inherently demanding about the CAD modelling problem; such as geometric and feature complexity, the number of interacting elements, and the need to coordinate multiple constraints and representations. Avoidable CL is shaped by interface characteristics, representation formats, and workflow choices that amplify visual search, WM demands, rework, or redundant CAD actions.

This experimental study therefore varies internal task complexity to provide a controlled, theoretically grounded way to elicit systematic CL differences while keeping the broader CAD environment, HCI modality, and format of representing engineering information constant. This is important for two reasons. First, the literature review showed that EEG indicators do not transfer uniformly from controlled laboratory setting to CAD-related tasks; thus, EEG-based CL measures should be treated as candidates requiring CAD-specific feasibility checks and validation. Varying task complexity provides a theoretically grounded validation reference; if an EEG feature is to be considered a viable CL indicator for CAD, it should be sensitive to CL changes induced by differences in internal task complexity. Second, the method this thesis proposes combines process-based performance modelling with temporally informative CL measurement (Section 5.3). Internal task complexity offers a suitable validation because it is expected to influence not only overall task-level CL but also the distribution of CL across recurring CAD modelling segments (unit task phases and subtasks), which is precisely what the CAD performance model and utilisation-based metrics aim to capture.

Accordingly, the second experimental study was designed around CAD tasks that intentionally differed in complexity, enabling the method to be operationalised and evaluated under two systematically different demand conditions. The first objective of the study was to identify EEG features indicative of changes in CL imposed on engineering designers during CAD tasks, in accordance with the step prescribed in Section 5.3.2 by the proposed method. Rather than assuming that commonly used EEG-based CL indicators generalise to CAD, the study empirically tests the most promising candidates emerging from the literature (Table 61) and selects those that are sensitive to experimentally induced complexity differences. This objective addresses the following research question:

2.1 *Which EEG feature, suggested by the literature as CL-sensitive, is the best candidate for describing CL variations imposed by CAD tasks of distinct complexity?*

The second objective was to operationalise CAD performance modelling steps (Section 5.2) by instantiating the theoretical CAD performance model (Section 4.2) through a concrete model of the low-complexity task, following the prescribed modelling steps (Section 5.2). The low-complexity task was selected because its scope is comparable to the relatively simple, well-specified part-modelling tasks used in the first empirical study and better aligns with the practical requirements of human performance modelling than the high-complexity task. This step is essential because it produces the hierarchical segmentation (unit task phases and intermediate-level subtasks) and the operator utilisations that the method uses to estimate CL from observable CAD actions. Using the EEG-based CL indicators identified in Section 6.2, Section 6.3 then derives information-processing weights for the unit task phases and intermediate-level subtasks, enabling an integration of EEG into utilisation-based CL equation (prescribed in Section 4.3 and Section 5.3.4). This objective addresses the following research question:

2.2 *Do unit task phases and subtasks differ in their information-processing intensity, as measured by EEG-based CL indicators?*

The third objective was to examine whether the weighted utilisations improve the method's ability to reflect engineering designers' perceived CL. Section 6.4 therefore compares CL calculated from weighted utilisations against CL calculated from non-weighted utilisations by relating both to NASA TLX ratings (its mental demand dimension). This addresses the following research question:

2.3 *How does CL calculated from weighted utilisations relate to self-reported subjective CL when compared to non-weighted utilisations?*

Because one of the key objectives of the method is that it is usable for engineering design researchers, Section 6.4 evaluates not only the most informative, participant-specific implementation, but also a simplified approach based on global utilisations (computed as across-participant median utilisation values for each task segment). Therefore, Section 6.4 compares global and participant-specific (median) utilisations as a test of the method's simplicity-fidelity trade-off. This addresses the following research question:

2.4 *How does CL calculated from global utilisations relate to self-reported subjective CL when compared to participant-specific median values?*

The fourth objective was to move beyond overall CAD task CL scores and examine CL at the level where the method is expected to provide unique insight; the unit task phases and intermediate-level subtasks defined by the CAD performance model. Section 6.6 quantifies segment-level differences in CL and their contributions to overall CAD task CL, and compares the implied ordering of segments across utilisation-based indicators and the selected EEG-based indicator. This section addresses the following three research questions:

2.5 *Do unit task phases/subtasks differ in CL implied by utilisation-based indicators?*

2.6 *What is the contribution of each unit task phase/subtask to overall CAD task CL?*

2.7 *Is the rank order of CL across task segments preserved when using different indicators?*

The fifth objective was to evaluate whether the indicators provide similar temporal distribution of CL during CAD activity performance. Section 6.7 models dynamic CL time series calculated from 1) EEG-based CL indicator, 2) non-weighted operator utilisation, and 3) weighted utilisations, and then quantifies their alignment. This addresses the following research question:

2.8 *Do dynamic changes in CL align when implied by EEG-based CL indicators, utilisations, and weighted utilisations?*

Finally, the sixth objective was to connect the method's CL estimates to CAD modelling task outcomes through which CAD performance is assessed. Section 6.8 therefore examines the relationship between CL calculated from utilisation-based indicators and overall CAD performance assessment metrics, thereby directly addressing the hypothesis introduced in Chapter 1 by answering the following research question:

2.9 *What is the relationship between CAD performance outcomes and CL in CAD tasks?*

The following section describes the methodological choices that enabled answering these research questions and implementing the method's workflow in a controlled CAD modelling experiment, providing the empirical basis for the method validation discussed in Chapter 7.

6.2 Methodology

Methodology through which the research was conducted, and the research questions were answered follows in the next sections.

6.2.1 Participants

The study recruited 24 mechanical engineers (two females and 22 males) to participate in the experimental study. Their age, professional engineering experience, CAD proficiency, as well frequency of using CAD software for generating 3D models and technical documentation are summarized in Table 62 using median (Med), median absolute deviation (MAD), and range. All participants were right-handed, had normal or corrected-to-normal vision, and did not report any neurological disorder. In addition, all the participants were familiar with the basics of CAD modelling in SolidWorks since they successfully completed the same CAD course as part of their studies.

Table 62 ES#2: Demographics, prior experience, and frequency of conducting CAD activities

	Med	MAD	Range
Age [years]	27	1.11	25 - 31
Professional engineering experience [months]	21	14.95	11 - 77
Self-reported CAD modelling skill [non-existent (1) – advanced (5)]	3	0.69	2 - 5
Self-reported SolidWorks proficiency [non-existent (1) – advanced (5)]	3	0.71	1 - 5
Performing engineering activities related to design [never (1) – every day (5)]	2.5	0.71	1 - 5
Using CAD software for technical documentation and 3D modelling [never (1) – every day (5)]	3	0.85	1 - 5

The participant selection followed the same logic as in the first experimental study (Section 3.2.1); engineering designers with prior SolidWorks modelling experience were intentionally recruited to ensure that task performance reflects skilled CAD system use rather than effects of learning the software, to reduce variability due to novice exploration, and to support comparability between empirical studies (if needed) by targeting participants with similar characteristics. This rationale also aligns with human performance modelling assumptions (MHP and KLM/GOMS [7], [92]), which are most applicable for users who can execute learned methods fluently, allowing observed differences in performance and CL measures to be attributed more directly to the experimental conditions rather than to baseline CAD proficiency.

6.2.2 Experimental procedure

Participants signed an informed consent at the beginning of the experimental procedure, as depicted in Figure 57. After that, they filled out a questionnaire covering demographics, educational background, their professional experience, and CAD proficiency. In the next step, they received instructions for filling out the NASA TLX questionnaire, including the description of its components.

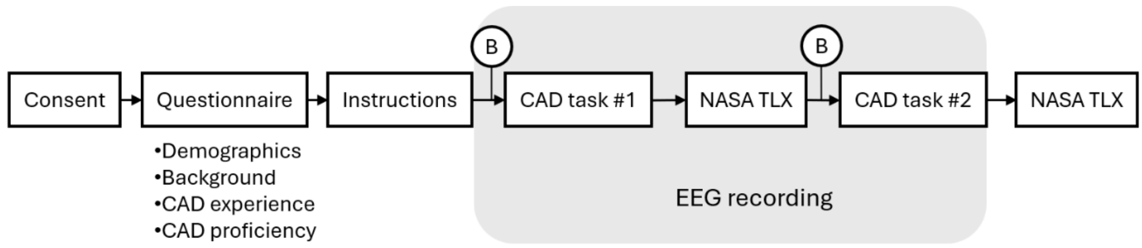

Figure 57 ES#2: Experimental procedure

Participants then proceeded with the CAD modelling tasks, time-limited to 15 minutes each. This time limit was used to bound maximum task exposure across participants and conditions and to reduce confounding effects of highly variable task durations. It also supports interpretation of EEG-based indicators as reflecting information-processing intensity per unit time (rather than total time spent), by avoiding extended periods of slower performance (e.g. acquisition/visual search noticed in the first empirical study; Section 3.4) that can attenuate contrasts in EEG signals when averaged over longer task segments.

All participants completed both tasks. However, they were divided into two groups with reversed task sequences to control for and mitigate any potential influence of task order on CL results. Before each CAD task, participants completed a baseline task, marked as “B” in Figure 57. In the baseline task, participants were asked to stare for two minutes at the black cross presented in the centre of the monitor screen with the white background. The EEG signals recorded during these baseline tasks were in further pre-processing steps used for normalisation purposes. After completing each CAD task, participants filled out the NASA TLX questionnaire. EEG signals were recorded continuously from the start of the first baseline task to the end of the last CAD task (highlighted in grey in Figure 57). This procedure allowed for a single placement of the EEG device, minimizing protocol length and reducing the risk of electrode misalignment. The average duration of the entire experiment was approximately 75 minutes.

6.2.3 Experimental setup

A high-performance computer powered the monitor screen (1920 x 1080 pixels; 60 Hz) on which stimuli, instructions, and SolidWorks application were presented and available to the participants (see Figure 58). A standard keyboard and a mouse served as interaction devices. The experimental protocol was run via iMotions platform. Participants advanced through the experiment and switched between the stimuli and SolidWorks application using a designated keyboard key. The screen, participants' faces, and the audio were recorded throughout the experiment. EEG data were collected using a 14-channel Emotiv EPOC+ device with an integrated amplifier. Continuous EEG signals were captured from the following electrodes: AF3, F7, F3, FC5, T7, P7, O1, O2, P8, T8, FC6, F4, F8, AF4. Their positions aligned with the international 10-20 system [228]. Two reference sensors used as the reference in the later data pre-processing steps were at P3 and P4 locations. Captured EEG signals were sent to the high-performance computer via wireless connection (Bluetooth Low Energy). The sampling frequency used to collect EEG data was 128 Hz, which is adequate for analysing frequency bands relevant to the study, according to the Nyquist-Shannon theorem [193].

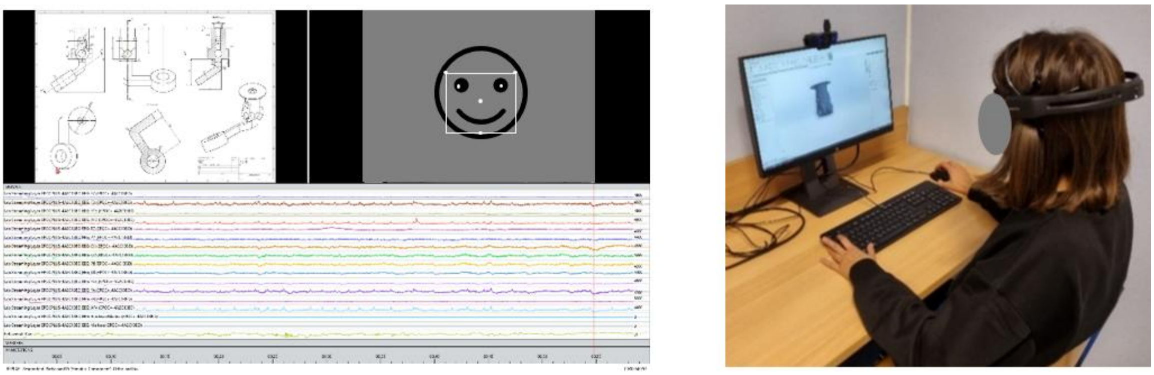

Figure 58 ES#2: Experimental setup

6.2.4 Collected experimental data

Data collected within the second experimental study is summarised in Table 63, together with the main purpose for collecting that data.

Table 63 Description of collected experimental data in the ES#2

Data type	Data description	Purpose
Questionnaire	Answers to a set of questions	Demographics, experience, expertise
PSVT	Answers to 12 questions, 3 parts of the test	Spatial ability
EEG data	14 channels, 128 Hz	EEG-based CL indicator(s)
NASA TLX scores	6 components (scores), comparison	Overall CL Experimental task validation Validation of the indicators
Behavioural data	Keyboard strokes, mouse clicks, window/screen transition	Model of CAD performance
CAD modelling outputs	CAD models, CAD log files	CAD performance assessment
CAD modelling process	Screen captures (videos)	CAD performance assessment Model of CAD performance

6.2.5 CAD modelling tasks

The study included two experimental tasks, in which participants were instructed to generate 3D CAD models of components using SolidWorks. The components were presented to participants in technical drawings, using both orthographic and isometric projections in both conditions. The complexity levels of the CAD modelling tasks varied intentionally. These levels were defined based on the overall geometric complexity of the components and the complexity of their CAD models. The components are presented in Figure 59.

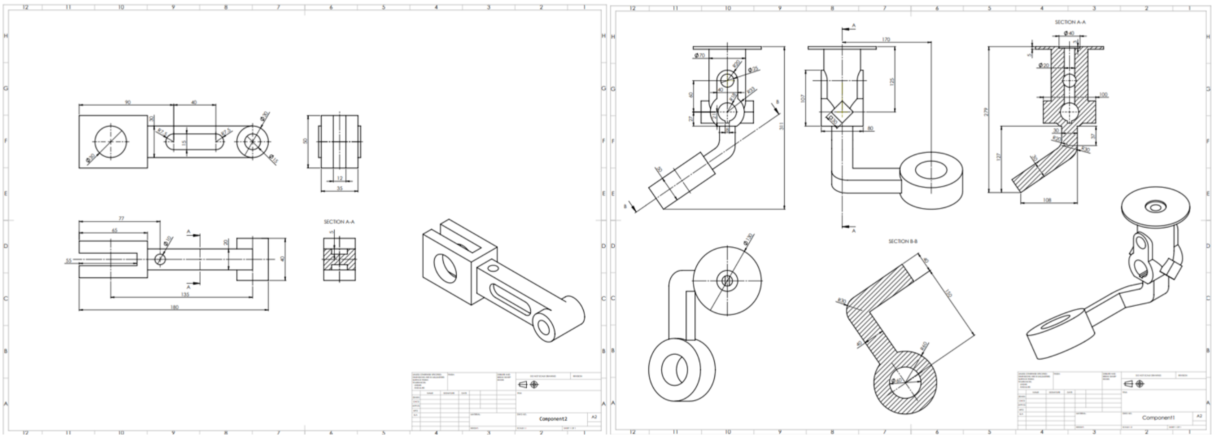

Figure 59 Low-complexity component (left) and high-complexity component (right)

Two proxies for the overall geometric complexity were the number of surfaces and dimensions, as suggested by Johnson et al. [109]. Similarly, the number of CAD features required to build CAD models in SolidWorks served as a proxy for CAD model complexity [109]. The components were chosen under the assumption that tasks involving a greater number of interacting elements would place a heavier CL on the engineering designer's cognitive system [13]. Interacting elements in the context of CAD modelling tasks refer to the design characteristics (such as form and dimensions) of the components to be modelled. These characteristics were presented to the participants in the technical drawings. The complexity levels were defined relatively for the purposes of the experiment, as they have not been prescribed by literature. The low-complexity (LC) component (shown in Figure 59 on the left) comprised fewer surfaces (33), dimensions (21) and CAD features (7), compared to the high-complexity (HC) component, which consisted of 49 surfaces, 31 dimensions, and 11 CAD features (presented in Figure 59 on the right). To ensure that the predetermined task complexity effectively manipulated CL, the experimental design was further validated using the NASA TLX questionnaire, which is further explained in Section 6.1.7.

6.2.6 NASA TLX questionnaire

The NASA TLX questionnaire was selected as the subjective assessment method for validating the experimental design and EEG results since it has been widely used as a "gold standard" for estimating CL (and WL) across various setups and contexts [134], [229]. A literature review also indicated that it is the most commonly used subjective assessment method within engineering design research (see Table 18). Furthermore, NASA TLX is a multi-dimensional questionnaire consisting of six components through which it investigates WL: mental demand, physical demand, temporal demand, performance, effort, and frustration [181]. This multi-dimensional nature allows for both a comparison of the overall WL experienced while performing CAD tasks, as well as an analysis of the individual contributions of each component. Of particular relevance to this study is the mental demand component, which is commonly used as a proxy for CL [129], [134], [181]. Descriptions of the components, adopted from Hart and Staveland [181], are provided in Table 64 and were made available to the participants during the experiment.

Table 64 Descriptions of NASA TLX components, adopted from [181]

Component	Definition
Mental demand	How much mental and perceptual activity was required (e.g. thinking, deciding, calculating, remembering, looking, searching, etc.)? Was the task easy or demanding, simple or complex, exacting or forgiving?
Physical demand	How much physical activity was required (e.g. pushing, pulling, turning, controlling, activating, etc.)? Was the task easy or demanding, slow or brisk, slack or strenuous, restful or laborious?
Temporal demand	How much time pressure did you feel due to the rate or pace at which the tasks or task elements occurred? Was the pace slow and leisurely or rapid and frantic?
Performance	How successful do you think you were in accomplishing the goals of the task set by the experimenter (or yourself)? How satisfied were you with your performance in accomplishing these goals?
Effort	How hard did you have to work (mentally and physically) to accomplish your level of performance?
Frustration	How insecure, discouraged, irritated, stressed, and annoyed versus secure, gratified, content, relaxed, and complacent did you feel during the task?

Participants self-reported the WL experienced during each CAD task by rating the magnitude of each of the six components on a twenty-point scale. In addition to providing these ratings, participants also evaluated the contribution (weight) of each component to the overall WL. This evaluation involved completing 15 pair-wise comparisons of the six components, where participants selected the component that contributed more to the workload in each comparison. Weights were assigned to each component based on the number of times it was selected as more relevant, with weights ranging from 0 (not relevant) to 5 (the most important). The score for each component was determined by multiplying its rating by the assigned weight. Finally, the overall weighted WL for each participant in each CAD task was calculated by summing the weighted rating across the six components and dividing by 15.

6.2.7 Experimental task validation

The design of experimental tasks was guided by an assumption of different complexities between the LC and the HC task, grounded in the work of Johnson et al. [109]. First, this assumption is corroborated by the participants' performance outcome scores, indicated by completeness and duration. Completeness is quantified via geometric similarity measures (volume, surface area, and shape) and dimensional accuracy of 3D CAD models. Geometric similarly measures were extracted from Graderworks application [72] and they are based upon comparing each participant's model to a gold-standard reference created by the author. Dimensional accuracy was calculated as a ratio of the correctly modelled dimensions (manually extracted from 3D CAD models) and the total number of specified dimensions in a technical drawing. Such metrics therefore provide transparent interpretation and avoid subjective grading; values closer to one (1) imply closer match and better output quality. Finally, completeness metrics and duration are combined in one overall performance metric (Equation (46)) that reflects both effectiveness and efficiency:

$$Overall\ performance = \frac{V + SA + S + DA}{Duration} \quad (46)$$

V is Volume, $V \in [0,1]$

SA is Surface Area, $SA \in [0,1]$

S is Shape, $V \in [0,1]$

DA is Dimensional Accuracy, $DA \in [0,1]$

Results presented in Table 65 show that CAD modelling duration differed between the LC and HC task. On the one hand, majority of the participants finished the LC task in the given time, ranging from 7.04 to 15.00 minutes, with the M task completion time of 12.36 minutes, and SD of 2.77 minutes. On the other hand, all the participants used the entire available 15 minutes to complete the HC task, while none of them finished it in the given time. Additionally, overall performance as well as four individual completeness metrics were better in the LC than in the HC task.

Table 65 CAD performance outcome metrics

CAD performance metric	Task	M	SD	Med	MAD	Range
Volume [%]	LC	0.97	0.18	1.00	$1.70 \cdot 10^{-2}$	0.19-1.22
	HC	0.44	0.18	0.46	$9.40 \cdot 10^{-2}$	0.00-0.87
Surface area [%]	LC	0.94	0.14	1.00	$1.10 \cdot 10^{-2}$	0.37-1.02
	HC	0.45	0.18	0.47	0.13	0.00-0.93
Shape [%]	LC	0.84	0.18	0.87	0.17	0.33-1.00
	HC	$3.10 \cdot 10^{-2}$	0.13	0.00	0.00	0.00-0.64
Dimensional accuracy [%]	LC	0.82	0.25	0.95	$4.76 \cdot 10^{-2}$	0.14-1.00
	HC	0.34	0.22	0.39	0.16	0.05-0.76
Duration [min]	LC	12.36	2.77	12.90	3.12	7.04-15.00
	HC	15.00	0.00	15.00	0.00	-
Overall performance [%/min]	LC	0.31	0.12	0.30	$7.98 \cdot 10^{-2}$	0.11-0.57
	HC	$8.42 \cdot 10^{-2}$	$3.89 \cdot 10^{-2}$	$8.74 \cdot 10^{-2}$	$2.29 \cdot 10^{-2}$	$4.34 \cdot 10^{-2}$-0.21

Second, the adequacy of the experimental tasks was validated by analysing the differences in NASA TLX scores between CAD tasks of varying complexity. The NASA TLX confirmed a significant difference in the overall WL participants experienced while completing the two CAD tasks. Significant differences were found in five out of six NASA TLX components (effort, frustration, mental demand, performance, and temporal demand) when compared between the HC and LC tasks, with higher scores in the HC task. Physical demand was the only component perceived as similarly low in both conditions. For physical and temporal demand, the effect size of the differences was moderate. For all other components, including overall workload, the effect size was large, ranging from 0.52 to 2.29. Detailed numerical values are provided in Table 66, with significant differences highlighted in bold.

Table 66 Differences in NASA TLX scores between the HC and the LC task

NASA TLX component	CAD task	M	SD	Med	MAD	Statistic; p-value	Effect size
Overall workload	HC	14.80	2.63	14.67	2.08	**T = 11.22; p = 2.38 · 10^{-11}**	2.29
	LC	7.58	2.50	8.27	3.02		
Mental demand	HC	45.46	29.45	48.00	39.29	**V = 260; p = 2.19 · 10^{-4}**	0.77
	LC	16.71	13.46	14.50	11.20		
Frustration	HC	30.75	25.91	24.00	20.76	**V = 210; p = 1.00 · 10^{-3}**	0.70
	LC	13.33	16.70	4.50	6.67		
Effort	HC	40.83	22.93	43.50	19.27	**V = 241.5; p = 2.00 · 10^{-3}**	0.64
	LC	22.17	12.63	18.00	15.57		
Performance	HC	38.08	21.22	36.00	26.69	**V = 240; p = 1.10 · 10^{-2}**	0.52
	LC	22.17	19.99	17.00	10.38		
Temporal demand	HC	62.75	33.78	74.00	31.14	**V = 222; p = 4.10 · 10^{-2}**	0.42
	LC	43.54	27.74	49.50	40.03		
Physical demand	HC	4.08	12.05	0.00	0.00	V = 105; p = 5.80 · 10^{-2}	0.33
	LC	5.21	3.53	0.00	0.00		

Therefore, obtained NASA TLX results align with the intended design of the experimental tasks, aimed to impose different levels of CL by varying their complexity. The largest effect size was observed for perceived mental demand, compared to the other NASA TLX components (presented in Table 66). These results identify mental demand, and consequently CL, as the main contributor to the overall WL. Hence, NASA TLX effectively validated the design of the experimental tasks.

6.2.8 EEG data pre-processing

The gathered raw EEG data were pre-processed in MATLAB employing the EEGLAB toolbox [202]. The pre-processing script followed steps prescribed in Section 5.3.2, while this section provides details on the used functions and explicit values where suitable. In the first step, the DC offset, specific to the Emotiv EPOC+ device employed in the experiment, was removed with an infinite impulse response (IIR) filter (0.16 Hz first-order high-pass filter). After that, frequencies outside the 4-45 Hz range were filtered out with a finite impulse response (FIR) filter, using the EEGLAB function "pop_eegfiltnew", hardcoded to a Hamming window [202]. The analysis focused on 4-45 Hz frequency range; bands outside that range were removed in the third step. In the next step, epochs of EEG signals contaminated by artifacts were identified and removed. Following the removal of outliers, EEG data were divided into theta (4-8 Hz), alpha (8-13 Hz), and beta (13-30 Hz) frequency bands using FIR filtering and Fast-Fourier Transformation [148]. The absolute power (POW) of the EEG signals was then computed for each CAD and baseline task. Finally, task-related power (TRP) was calculated for each participant and CAD task.

This TRP calculation method was applied to nine EEG features identified from the literature. TRP was computed for the entire cortex and four specific cortical areas: frontal, parietal, occipital, and Pz, across theta, alpha, and beta bands (see Table 61). EEG features reflecting brain activity over the entire cortex included signals captured from all 14 channels of the EEG device. For other features, TRP was calculated using signals from channels located within the cortical area of interest. The specific channels used in the calculation of each EEG feature are outlined in Table 67.

Table 67 List of considered EEG features

<table>
<tr><th></th><th>Frequency band</th><th>Cortical area</th><th>Included channels</th></tr>
<tr><td rowspan="7">Frequency band TRP</td><td rowspan="2">Theta (4-8 Hz)</td><td>Entire</td><td>All 14 channels</td></tr>
<tr><td>Frontal</td><td>F7, F3, F4, F8</td></tr>
<tr><td rowspan="4">Alpha (8-13 Hz)</td><td>Entire</td><td>All 14 channels</td></tr>
<tr><td>Parietal</td><td>P7, P8</td></tr>
<tr><td>Occipital</td><td>O1, O2</td></tr>
<tr><td>Pz</td><td>mean (P7, P8)</td></tr>
<tr><td>Beta (13-30 Hz)</td><td>Entire</td><td>All 14 channels</td></tr>
<tr><td rowspan="2">Ratio</td><td rowspan="2">Theta/Alpha</td><td>Entire</td><td>All 14 channels</td></tr>
<tr><td>Fz theta; Pz alpha</td><td>mean (F3, F4); mean (P7, P8)</td></tr>
</table>

6.3 Identifying EEG-based CL indicators for the CAD context[5]

As the first part of the analysis, differences in the values of the nine EEG features emerged from the literature as the promising CL indicators (listed in Table 61) were calculated and compared between two experimental conditions, depending on the task complexity and associated CL levels (low and high) of the experimental CAD tasks. Therefore, task complexity and its associated CL level served as the independent variable, whose influence was explored on the values of EEG features as a dependent variable. This part of the analysis considered data from two complementary methods for measuring CL: EEG as a psychophysiological measurement method and NASA TLX questionnaire as a subjective measurement method. NASA TLX here provided means for validating and comparing the relevance of the observed EEG features in distinguishing between CL levels in the CAD context.

6.3.1 Data analysis

Data analysis was performed using R version 2023.03.1+446. Descriptive statistics encompassed the calculation of the M and the Med as measures of central tendency, as well as standard deviation (SD) and MAD as measures of variability. These measures were calculated for NASA TLX scores across both task conditions (HC and LC CAD tasks) for each of the six components (see Table 64), as well as for overall WL. Additionally, measures of descriptive statistics were calculated to summarize and describe the values of the nine EEG features (listed in Table 67) in both the HC and the LC CAD task.

Inferential tests were then applied to answer two complementary questions that reflect the dual aim of the method: 1) identifying EEG features that can discriminate CL states (LC vs. HC) imposed by CAD tasks of different complexity and 2) selecting an EEG feature that is suitable for dynamic CL tracking during CAD modelling (i.e. across task segments or moving windows), consistent with the thesis' process-based view of CAD performance and CL. First, paired comparisons between HC and LC were conducted for NASA TLC and EEG features to test whether task complexity induced systematic differences at the group level. NASA TLX comparisons served to validate the manipulation by confirming that HC tasks elicited higher overall WL and CL and by inspection which TLX component contributed most strongly to the difference. EEG feature comparisons served to examine sensitivity to the complexity-driven CL

[5] This section is based on:
Lukačević et al. (2025) Identifying the electroencephalography features for measuring cognitive load in computer-aided design. *Journal of Mechanical Design*, 147 (12), 121403.
doi: 10.1115/1.4068746

manipulation; an EEG feature was considered a candidate CL indicator if it differed significantly between HC and LC. In addition to statistical significance, the corresponding effect sizes were examined to rank features by their ability to distinguish CL levels under the CAD task complexity manipulation.

A paired t-test with Bonferroni adjusted p-value was employed when the assumptions of normality (tested by Shapiro-Wilk normality test; $p < 0.05$) and equity of variances (tested by Levene test; $p < 0.05$) were met. Effect sizes for the paired t-test were calculated using Cohen's d, representing the number of standard deviations between the means of the compared variables. When normality assumptions were violated, the Wilcoxon signed-rank test with Bonferroni adjusted p-values was employed instead of the paired t-test. The effect size for the Wilcoxon test was calculated as the Z statistic divided by the square root of the sample size. Measures of descriptive statistics, significant differences (adjusted p-values), and associated effect sizes are coupled with the test statistic values; marked with T for the t-test and V for the Wilcoxon signed-rank test.

Second, to evaluate whether candidate EEG features relate to subjective CL, repeated-measures correlation (rmcorr) [226] was computed between mental demand NASA TLX component and EEG features. It tested the monotonic relationship between the variables but within the participant while considering the task condition (LC vs. HC), thus accounting for the paired (repeated measures) experimental task design. In other words, rmcorr evaluated whether, within the participant, higher EEG feature values in the HC condition relative to the LC condition tend to coincide with higher self-reported mental demand, while removing stable between-participant differences. Importantly, rmcorr does so using the original LC and HC observations rather than a single HC-LC score.

Finally, these analyses were corroborated and aligned with the intended dynamic use of the indicator through robust nonparametric regression using Theil-Sen estimation [230]. Because NASA TLX is obtained once per condition and is therefore not time-resolved, dynamic suitability could only be evaluated in the sense of condition-driven change; an EEG feature is considered better suited for dynamic CL estimation if its Δ(HC-LC) reliably covaries with the corresponding Δ(HC-LC) in NASA TLX mental demand rating across participants. Accordingly, Theil-Sen models tested linear relationships between Δ EEG feature (HC-LC) and Δ mental demand (HC-LC), thereby removing stable individual baselines and directly targeting whether the complexity manipulation affects coupled changes in CL, considering both the psychophysiological and subjective responses. This change-based regression can be viewed as the validation step for “dynamic” suitability in the present dataset, as it evaluates whether within-participant differences in the magnitude of complexity-induced change in the EEG feature relate to within-participant differences in the magnitude of complexity-induced change in subjective mental demand. Within this thesis, only linear regression models were tested as the simplest option due to the limited and exploratory dataset (n = 22).

6.3.2 Results

EEG results are divided into three parts. First presented are the differences, tested using statistical tests, in the nine EEG features between the CAD tasks, along with their effect sizes. These results are illustrated graphically using box plots. In addition, results are detailed numerically within the tables, including measures of descriptive statistics, adjusted p-values, associated effect sizes, and test statistics values (marked with T for the t-test, and V for the Wilcoxon signed-rank test).

The second part of the results details the relationships between EEG features and NASA TLX scores, tested using the repeated-measures correlation. For each tested EEG-NASA TLX pair, the results express the strength of their relationship with the repeated-measures correlation coefficient (r), together with its 95% confidence interval (CI), the corresponding error degrees of freedom (df), and the associated p-value. To support interpretation and transparently illustrate the within-participant relationships, the results are accompanied by a scatterplot in which each participant's repeated observations are shown in a unique colour and connected by parallel within-participant trend lines, highlighting the common within-individual association estimated by r.

Finally, the third section presents the developed Thiel-Sen regression models, used as a robust nonparametric alternative to ordinary least squares for modelling the relationship between EEG features and subjective CL. In line with robust reporting practice, the slope (b) is reported as the Med of all pairwise slopes, and the intercept (a) is reported as a robust location estimate (based on Med, positioning the line through the central data density). Model fit is summarised using a robust error metrics, reported here as the median absolute error (MAE). Uncertainty in the estimated trend is communicated via a bootstrapped 95% confidence interval (CI) for the slope b.

6.3.2.1 Differences in EEG features between the tasks

The analysis revealed six EEG features that significantly differed when compared between the HC and the LC task: alpha TRP, alpha (occipital) TRP, alpha (Pz) TRP, alpha (parietal) TRP, theta/alpha TRP, and beta TRP. These six features are, therefore, identified as sensitive to varying complexity of CAD tasks and associated changes in CL. Significant differences are marked with a star in Figure 60 and bolded in Table 68. EEG features are, in Table 68, listed in descending order based on the effect size of the differences, suggesting that the top features are the best candidates for distinguishing CL levels in CAD tasks.

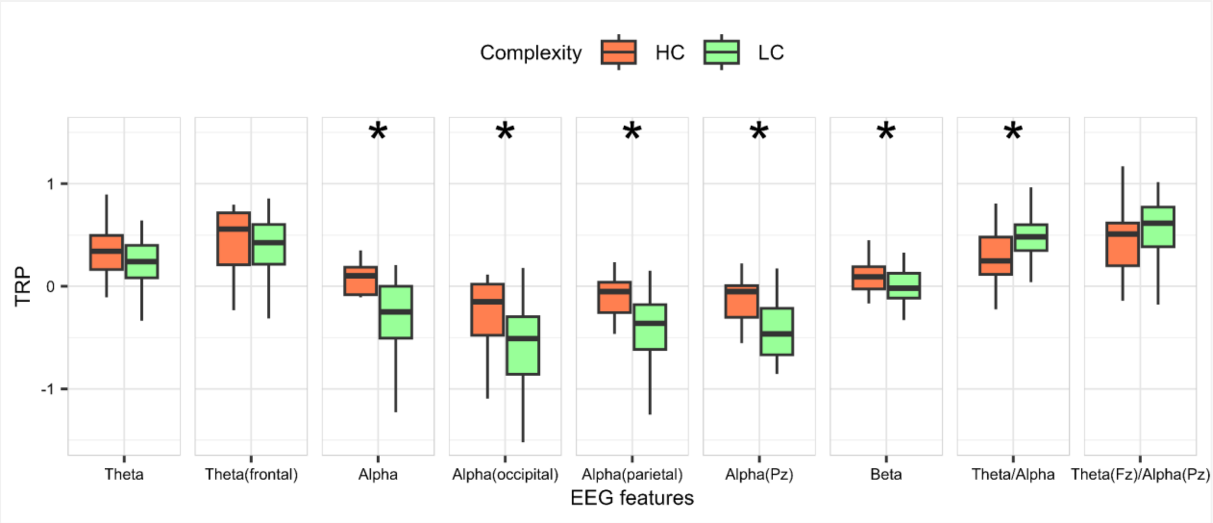

Figure 60 Differences in EEG features between the HC and the LC task

The largest effect size (0.85) was found for alpha TRP across the entire cortex, followed by alpha TRP in different cortical areas (occipital, Pz, parietal). This highlights the critical role of alpha TRP in distinguishing CL levels, outperforming both ratios and the other frequency bands (beta and theta). The theta/alpha ratio reflected differences in CL between the tasks only when considering the entire cortex. Theta TRP showed similar results in both tasks, regardless of spatial distribution, suggesting that theta is not a key EEG feature in distinguishing CL levels in the CAD context.

The results further suggest that increases in TRP are relevant indicators of differences in CAD task complexity and associated CL levels. Notably, alpha, beta, and theta TRP increased with task complexity (and CL), regardless of the observed spatial region. When considering the entire cortex, alpha and beta TRP increased in the HC task and decreased in the LC task. However, when specific cortical areas (occipital, parietal, and Pz) are observed, alpha TRP decreased from baseline levels, but the decrease was smaller in the HC than in the LC task. In contrast, theta TRP increased in both CAD tasks relative to baseline, with a larger increase in the HC task, regardless of spatial distribution. Finally, both calculated ratios decreased as task complexity increased, aligning with the changes observed in theta and alpha as individual frequency bands.

Table 68 Differences in EEG features between the HC and the LC task

EEG feature	CAD task	M	SD	Med	MAD	Statistic; Adjusted p-value	Effect size
Alpha	HC	$-9 \cdot 10^{-3}$	0.31	0.10	0.20	**V = 295; p = 1.19 · 10^{-6}**	**0.85**
	LC	-0.28	0.37	-0.25	0.39		
Alpha (occipital)	HC	-0.29	0.40	-0.15	0.31	**V = 279; p = 5.33 · 10^{-5}**	**0.75**
	LC	-0.60	0.48	-0.51	0.39		
Alpha (Pz)	HC	-0.19	0.33	$-5.2 \cdot 10^{-2}$	0.23	**V = 279; p = 5.33 · 10^{-5}**	**0.75**
	LC	-0.47	0.44	-0.46	0.34		
Alpha (parietal)	HC	-0.16	0.31	$-5.1 \cdot 10^{-2}$	0.23	**V = 272; p = 1.75 · 10^{-4}**	**0.71**
	LC	-0.41	0.40	-0.36	0.34		
Theta / Alpha	HC	0.33	0.38	0.25	0.31	**V = 42; p = 1 · 10^{-3}**	**0.63**
	LC	0.53	0.40	0.48	0.19		
Beta	HC	0.10	0.39	$9.4 \cdot 10^{-2}$	0.16	**V = 249; p = 4 · 10^{-3}**	**0.58**
	LC	$-2.2 \cdot 10^{-2}$	0.30	$-1.8 \cdot 10^{-2}$	0.19		
Theta (Fz) / Alpha (Pz)	HC	0.51	0.60	0.51	0.32	V = 83; p = 5.60 · 10^{-2}	0.39
	LC	0.62	0.52	0.62	0.28		
Theta	HC	0.33	0.26	0.34	0.24	T = 1.59; p = 0.13	0.33
	LC	0.25	0.25	0.24	0.24		
Theta (frontal)	HC	0.47	0.39	0.56	0.33	T = 1.44; p = 0.16	0.29
	LC	0.38	0.31	0.43	0.31		

6.3.2.2 Correlations between EEG features and NASA TLX scores

Repeated-measures correlation was used to quantify within-participant associations between NASA-TLX mental demand and EEG features across the two task conditions (LC and HC), thereby accounting for the paired experimental design. The results (Table 69) show that mental demand was most strongly and consistently associated with Alpha TRP, with significant positive relationships observed for Alpha TRP captured over the entire cortex, as well as for Alpha TRP extracted from occipital, parietal, and Pz regions.

Table 69 Repeated-measures correlation results between mental demand and EEG features

EEG feature	r	p-value	95% CI	df
Alpha	**0.78**	$\mathbf{3.61 \cdot 10^{-6}}$	**0.56, 0.90**	**23**
Alpha (occipital)	**0.71**	$\mathbf{6.62 \cdot 10^{-5}}$	**0.44, 0.86**	**23**
Alpha (Pz)	**0.70**	$\mathbf{8.99 \cdot 10^{-5}}$	**0.43, 0.86**	**23**
Alpha (parietal)	**0.62**	$\mathbf{9.31 \cdot 10^{-4}}$	**0.30, 0.82**	**23**
Theta / Alpha	**-0.62**	$\mathbf{9.46 \cdot 10^{-4}}$	**-0.82, -0.30**	**23**
Beta	**0.47**	$\mathbf{1.86 \cdot 10^{-2}}$	$\mathbf{8.81 \cdot 10^{-2}}$**, 0.73**	**23**
Theta (Fz) / Alpha (Pz)	-0.46	$2.19 \cdot 10^{-2}$	-0.72, $-7.47 \cdot 10^{-2}$	23
Theta	0.24	0.26	-0.17, 0.58	23
Theta (frontal)	0.27	0.20	-0.14, 0.60	23

The rmcorr visualisations (Figure 61) illustrate significant relationships by showing participant-specific trajectories between LC and HC together with the common within-participant trend; Alpha TRP variants show clear positive slopes across participants, Theta/Alpha shows an inverse relationship, and Beta exhibits a comparatively shallow positive trend. Overall, the repeated-measures analysis corroborates EEG features that reflect Alpha TRP (across the entire cortex as well as across occipital, parietal, and Pz regions) as the most promising candidates for representing subjective CL variations during CAD modelling.

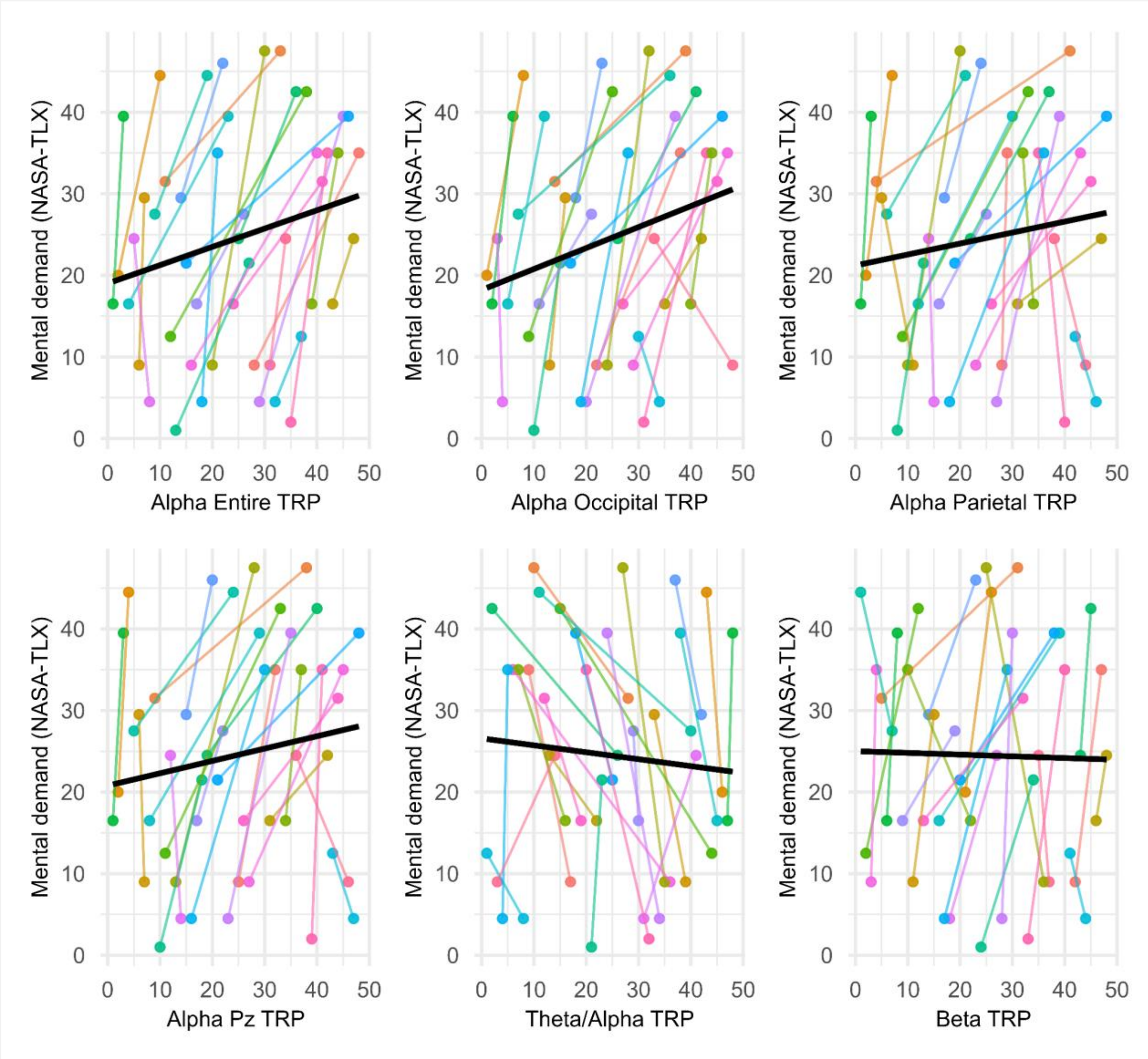

Figure 61 Repeated-measures correlations between NASA TLX scores and EEG features

6.3.2.3 Theil-Sen regression model for predicting subjective CL from EEG features

A Theil–Sen (Med-based) regression was used to evaluate whether condition-driven (CAD task complexity) changes in EEG features between the HC and LC tasks (Δ EEG feature = HC − LC) covary with the corresponding condition-driven changes in subjective CL captured by NASA TLX mental demand (Δ mental demand = HC − LC). Candidate indicators were compared primarily using the median absolute error (MAE) from LOOCV as a robust measure of how tightly Δ mental demand aligns with Δ EEG, while p-values were used to classify relationships as significant or non-significant. Among the alpha-band features, Occipital Alpha TRP showed the lowest MAE and was significant, followed by Parietal Alpha TRP, Alpha at Pz, and Alpha TRP from the entire cortex (see Table 70 and Figure 62).

Table 70 Theil-Sen regressions

EEG feature	Med intercept a	Med slope b	95% CI for b	p-value	MAE
Alpha (occipital)	**6.82**	**2.87**	**0.33, 5.10**	$\mathbf{4.19 \cdot 10^{-2}}$	**1.61**
Alpha (parietal)	**6.33**	**2.18**	**1.56, 6.33**	$\mathbf{7.89 \cdot 10^{-3}}$	**1.68**
Alpha (Pz)	**6.34**	**1.57**	**2.27, 7.56**	$\mathbf{8.36 \cdot 10^{-3}}$	**1.81**
Alpha	**6.09**	**3.02**	**0.29, 6.44**	$\mathbf{5.00 \cdot 10^{-2}}$	**1.84**
Theta (Fz) / Alpha (Pz)	6.87	-2.40	-4.36, -0.042	$4.63 \cdot 10^{-2}$	2.03
Theta / Alpha	6.56	0.00	-3.78, 5.71	1.00	6.56
Theta (frontal)	6.97	-3.02	-6.42, 0.24	$6.46 \cdot 10^{-2}$	2.02
Theta	7.00	0.00	-5.55, 3.84	0.52	2.00
Beta	7.00	0.00	-9.60, 4.53	0.53	1.56

In contrast, other EEG features either exhibited larger errors or were not significant; Theta/Alpha produced a near-zero slope and was non-significant, Theta and Frontal Theta were non-significant, and Beta was non-significant despite a relatively low MAE, suggesting that its change scores did not consistently align with changes in mental demand. Overall, the results indicate that TRP within alpha frequency band provides the most suitable basis for capturing complexity-driven CL changes, with Occipital and Parietal Alpha TRP offering the tightest significant change coupling (lowest MAE) and thus representing the strongest candidates for dynamic CL estimation in the CAD context.

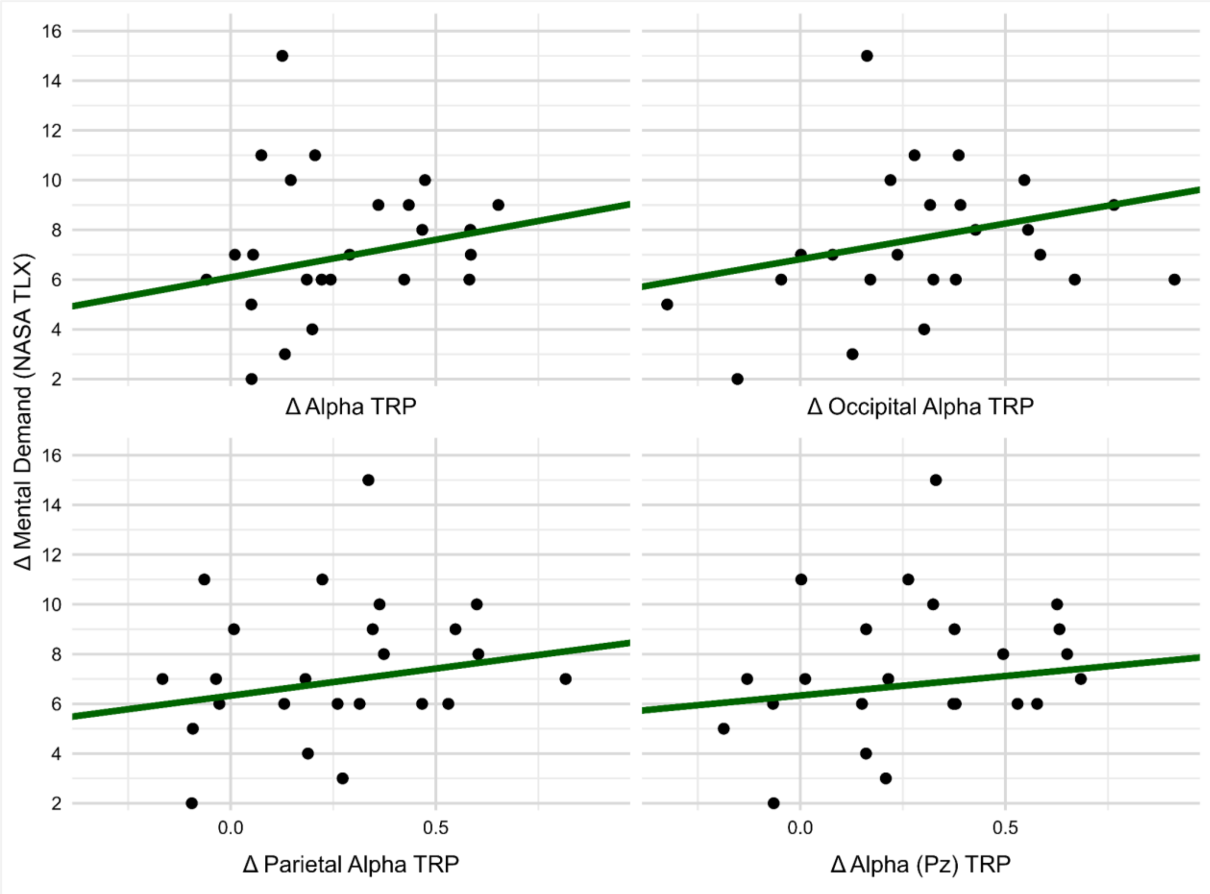

Figure 62 Theil-Sen regression models between EEG features and mental demand (NASA TLX)

6.3.3 Discussion

The presented research is the first attempt to study CL in the underexplored context of CAD tasks using two complementary methods: NASA TLX as a subjective measure and EEG as a psychophysiological measure. Using both methods enabled validation of findings from each and an exploration of their relationship. Moreover, analysing differences in NASA TLX component scores provided an important validation of the experimental task design, a step often overlooked in studies employing psychophysiological methods to explore CL. This is particularly important in the CAD context, where the high complexity of realistic tasks makes it challenging to design experiments where CL is the primary differentiating factor. Without confirmation from subjective measurement methods like NASA TLX, it would be difficult to argue that CL is indeed driving the differences in EEG signals. This methodological limitation also applies to the only similar study using EEG to investigate CL in CAD tasks of varying complexity, described by Anand et al. [223] in their unpublished preprint.

Previous studies have shown that the theta, alpha, and beta frequency bands are sensitive to variation in CL across different tasks (see Section 2.4.3). As a part of the second empirical study, nine EEG features based on POW changes in these three frequency bands (as listed in Table 61) were analysed as candidate EEG-based indicators of CL in the CAD context. The selection of these features was informed by previous research in both cognitive psychology and engineering design (see Section 2.4). As a result, this study tested a larger set of EEG features compared to similar studies and identified those most relevant to the CAD context based on the existing literature. Overall, the comparisons of the LC and HC task showed clear, systematic differentiation across multiple EEG features, with several effects also exhibiting large magnitudes (as shown in Figure 60 and Table 68). Across the candidate list, alpha emerged as the most consistently responsive frequency band. This interpretation is further supported by the repeated-measures correlation analysis (Table 70), which showed strong positive relationships between mental demand and alpha TRP: entire-cortex alpha ($r = 0.78$), occipital alpha ($r = 0.71$), Pz alpha ($r = 0.70$), and parietal alpha ($r = 0.62$). However, while these HC–LC contrasts and repeated-measures correlations establish group-level sensitivity and within-participant covariation, dynamic CL indicators must additionally capture condition-driven changes at the participant level. Because NASA TLX was obtained once per condition and is therefore not time-resolved, dynamic suitability was evaluated here using change-based analyses in which both EEG and NASA TLX ratings were expressed as Δ(HC−LC). The Theil–Sen regressions on Δ scores therefore provide the strictest test of which alpha feature best supports dynamic CL estimation in the present dataset. Rather than pointing to a single best EEG feature, the combined results of conducted analyses suggest that alpha TRP captured in parietal and occipital cortical regions provide a robust psychophysiological sensitivity to complexity-driven CL variations in CAD modelling. This finding aligns with previous studies that reported parietal and occipital alpha sensitivity to CL, but in cognitively controlled tasks (e.g. [25], [168], [175], [176]).

However, a closer examination of frequency bands' behaviour, especially alpha, suggests that EEG-based CL indicators from prior studies should not be directly transferred to the CAD context. Differences were observed when compared to studies involving both cognitively controlled and more complex CAD-related tasks. For example, previous research indicated that alpha TRP decreases with increasing CL [164], whereas this study found an increase in alpha TRP with increased CL in CAD tasks, regardless of spatial distribution. One possible explanation for the observed increase in alpha TRP may be related to cognitive characteristics of the CAD tasks and associated processing of visuospatial information required to perform them [178]. Aligned with that suggestion, increased alpha TRP was also found in the first experimental study (see Section 3.5) and previously reported by Lukačević et al. [178]. In addition, several other studies have reported an increased alpha TRP while solving standardised tests related to aspects of visuospatial thinking (e.g. [155], [157], [198]). Therefore, an interpretation of the observed increase in parietal and occipital alpha TRP may be grounded in the functional roles of these cortical areas and in the context-depended meaning of alpha activity [231]. CAD modelling is dominated by sustained visuospatial processing, goal maintenance, and repeated mental updating of spatial relations; accordingly, posterior cortex is expected to be heavily engaged, with the occipital area supporting visual information processing and the parietal area supporting visuospatial attention, spatial transformations, and top-down control of visual processing [155], [157], [165], [198], [232]. Importantly, alpha power in these regions is not a unitary more/less CL indicator [163]; in WM research, alpha TRP often decreases when decoding external input but increases during maintenance as a mechanism for gating sensory input and protecting internal representations from interference [233]. In more demanding CAD tasks, the balance of processing is assumed to shift toward longer and more intense periods of maintaining and manipulating an internal 3D representation, which would suggest stronger alpha synchronization in parietal and occipital areas. In addition, posterior alpha TRP increases have been interpreted as the brain suppressing visual processing when demands rise, effectively reducing sensitivity to ongoing visual input to protect and maintain task-relevant information in working memory [232]. Finally, parietal and occipital alpha increase can also be framed as a "task-shielding" mechanism; when internal processing demands rise, the system may suppress stimulus-driven reorienting and task-irrelevant input to stabilise ongoing reasoning, which is aligned with findings that parietal alpha TRP can increase when cognition requires stronger internal attention and reduced distraction from external stimuli [234]. Taken together, these perspectives suggest that the observed occipital and parietal alpha TRP increase in the HC CAD task is compatible with higher complexity-driven CL, but reflects not simply more activation, rather stronger top-down control and inhibition of irrelevant sensory processing in service of maintaining and manipulating richer visuospatial information. These explanations align well with the literature review's conclusion that EEG-based CL indicators are task-dependent (e.g. [130], [163]), and they provide a plausible explanation for why parietal and occipital alpha TRP may increase with complexity-driven CL rather than decrease in CAD modelling, where maintaining and manipulating a rich visuospatial mental model is central.

Other possible explanations include that an increase in alpha could be related to higher creativity, as suggested by Nguyen and Zeng [126]. However, such explanations seem to be less applicable in the presented study dealing with CAD modelling tasks. Furthermore, the increase in beta TRP with rising CL may be another CAD-specific pattern in EEG signals, consistent with prior CAD-related studies (Table 20) but contradicting findings from cognitively controlled tasks. A larger TRP increase in the HC task, regardless of spatial distribution, is consistent with both cognitively controlled and CAD-related studies, though statistical significance for theta TRP differences was not found. This lack of significance persisted even when focusing on frontal theta, which, along with parietal alpha, had been suggested as a key EEG-based CL indicator across various fields [164]. Further research is needed to explain these CAD-specific EEG patterns and the underlying cognitive processes in CAD tasks. Additionally, future studies should investigate whether these findings are representative of CAD tasks in general and if they remain consistent across other types of CAD tasks (e.g. related to assembly, motion and stress simulation and analysis, or design review). These initial findings thus pave the way for future studies aiming to explore more combinations of EEG features (e.g. frequency bands and cortical areas) using machine learning algorithms to better distinguish CL levels. Within the scope of this thesis, further work focuses on analysing dynamic changes in CL based on the values of parietal alpha TRP throughout the CAD modelling tasks and across its hierarchical levels. Occipital and parietal alpha TRP showed very similar behaviour across conditions, and both correlated strongly with subjective mental demand. Because these two posterior alpha features capture largely overlapping variance, the subsequent analyses use parietal alpha TRP as a single representative EEG-based CL indicator to avoid redundancy and keep the dynamic CL modelling simple and interpretable. This choice is supported by their strong association on change scores Δ(HC-LC) quantified by Spearman's $\rho = 0.83$ ($p < 0.05$), indicating that participant who showed larger occipital alpha TRP changes also tended to show larger parietal alpha TRP changes.

6.4 Assigning information processing weights to intermediate-level subtasks and unit task phases

Parietal alpha TRP, the EEG feature identified as the promising candidate indicator of CL for CAD modelling, was further analysed here to investigate where information processing intensity concentrates within the process. To enable segment-wise EEG comparison and assignment of information-processing weights, the CAD performance modelling procedure described in Section 5.2 was applied first to the LC task for all 22 participants. Specifically, this procedure produces 1) a timestamped HCI script of CAD events (Section 5.2.1), 2) a timestamped hierarchical decomposition of the modelling process into unit task phases and three hierarchical task/goal levels (Section 5.2.2), and 3) the operator mapping needed later for utilisation-based CL estimation (prescribed in Section 5.2.3, calculated for empirical data in Section 6.5). The resulting segmentation defines the task segments (with their start and end times) over which TRP can be aggregated and compared in a consistent and interpretable way.
The resulting CAD performance models (i.e. the full multi-layered timelines and operator streams for all 22 participants) are not presented here, because the aim of this section is not to compare CAD performance model structures across participants or conditions, but to use the derived segments as the basis for segment-wise TRP comparison and subsequent weight assignment. In this section, TRP is compared between the recurring unit task phases (*acquisition* and *inspection*), and across the intermediate-level subtasks (CAD operations) that operationalise *execution* (*navigating, manipulating, referencing, drawing, dimensioning, constraining, deleting*). The following analysis therefore supports the overarching goal of integrating two complementary perspectives: psychophysiological measurement (EEG) as an indicator of information-processing intensity and performance modelling (operator utilisation) as an indicator of resource engagement over time, enabling later computation of weighted utilisation-based CL estimates in Section 6.5.

6.4.1 Data analysis

EEG analysis was performed on the LC-task recordings of all 22 participants using the segment boundaries obtained from the CAD performance modelling procedure (Section 5.2). Concretely, the time-aligned decomposition provided start and end times for unit task phases (acquisition, execution, inspection) and for the intermediate-level subtasks that occur during execution (navigating, manipulating, referencing, drawing, dimensioning, constraining, deleting). These segment definitions were used here as the input for aggregating parietal alpha TRP to the process level of interest. In particular, median parietal alpha TRP values were calculated for each participant and each task segment (unit task phases and intermediate-level subtasks). The TRP values were referenced to each participant's baseline task and thus made comparable within and between participants.

Unit task phases are required to describe the CAD modelling process in full because they capture how information processing unfolds over time; preparing/understanding what to do (acquisition), carrying out the modelling work (execution), and checking/verifying intermediate results (inspection). In the present analysis, *acquisition* and *inspection* are treated at the unit task phase level because they represent broadly recurring preparation and verification modes that are comparatively consistent across unit tasks and CAD contexts. They are not further segmented because their analytic value lies precisely in being generic phases that can accompany many different CAD operations; subdividing them would tend to produce fine-grained categories that are harder (or impossible) to code reliably and less comparable across participants, tasks and studies.

In contrast, *execution* is not analysed as a single unit task phase but segmented into intermediate subtasks because it is too heterogeneous in CAD. Execution can involve fundamentally different CAD operator mixtures and functional demands depending on whether the engineering designer is navigating, referencing, drawing, constraining, dimensioning, manipulating the view, or deleting. For that reason, execution is further segmented and analysed through intermediate-level subtasks (CAD operations) that preserve functional meaning and are reusable across CAD tasks, improving interpretability and providing a cleaner basis for linking task segments to utilisation-based indicators.

The intermediate subtask level is explicitly prescribed in the proposed CAD performance model (Section 4.2) and measurement method (Section 5.2) as a generic, reusable abstraction; it is specific enough to preserve functional meaning (what type of CAD modelling work is being done), yet general enough to be applied across users, tasks, and studies without collapsing to software-specific commands (as CAD action layer) or task-specific geometry details (as unit task operation layer). This choice is also supported by the first empirical study (Section 3.4.4.5), where the CAD operation layer provided the clearest and most systematic discrimination between conditions, whereas the CAD action layer was more sensitive to low-level HCI strategy differences. Moreover, strategically important CAD knowledge is often "hidden" at exactly this layer; in how engineering designers organise and combine command types into meaningful CAD operations rather than at the keystroke/command level [98]. Analysing TRP at the level of intermediate subtasks therefore targets the same layer at which CAD modelling strategies and efficiency-relevant choices are expected to manifest. Prior process-oriented CAD models likewise separate primitive HCI events from higher-level operations and trear correct event-to-operation clustering as the key to interpretable analysis [28]. In that sense, the present analysis aligns with both the thesis' model (hierarchical task/process layers and orthogonal unit task phases) and the broader CAD literature that motivates this layer as the most informative and transferable basis for linking process structure, performance, and CL.

6.4.1.1 Calculating the differences in TRP values

Data analysis was performed using R version 2023.03.1+446. Descriptive statistics encompassed the calculation of the Med as a measure of central tendency and MAD as a robust measure of variability. These measures were calculated for alpha parietal TRP of 22 participants across unit task phases (*acquisition*, *inspection*) and the seven intermediate-level subtasks (*navigating*, *manipulating*, *referencing*, *drawing*, *dimensioning*, *constraining*, *deleting*). This segmentation corresponds to the model's prescribed reusable abstraction level and provides the segment set s used later in this section when computing information-processing weights.
Furthermore, Friedman test was applied to quantitatively compare differences in alpha parietal TRP between the unit task phases as well as between subtasks. The effect size of differences determined with the Friedman test was computed using Kendall's W. Post-hoc analysis involved pairwise comparisons between the unit task phases as well as the subtasks using the Wilcoxon signed-rank test with Bonferroni adjusted p-values. The effect size for the Wilcoxon test was calculated as the z-statistic divided by the square root of the sample size. Measures of descriptive statistics, significant differences (adjusted p-values), and associated effect sizes are coupled with the test statistic values; marked with F for the Friedman test and V for the Wilcoxon signed-rank test.

6.4.1.2 Calculating the information-processing weights

Information-processing weights were computed following the procedure prescribed in Section 5.3.3. A robust within-participant z-score of parietal alpha TRP was computed for each segment s and participant i according to Equation (40). This dataset was trimmed at 10% to mitigate the potential influence of extremes. The final robust z-scores (one for each segment) were converted to weights using the *softmax* function (see Equation (41)). Considered task segments s in this case were the unit task phases and the intermediate-level subtasks. Only participants with complete data for all segments (unit task phases and subtasks) were included in this analysis. To assess how consistently segments were ranked across participants based on parietal alpha

TPR, Kendall's coefficient of concordance (W) was computed. The corresponding significance of concordance was evaluated using the standard chi-square approximation (χ^2).

6.4.2 Results

Median parietal alpha TRP values of these unit task phases and intermediate-level subtasks are presented in Figure 63, while numerical values are available in Table 71.

Table 71 Descriptive statistics of parietal alpha TRP values for unit task phases and subtasks

Unit task phase/Subtask	Occurrences	Med	MAD	Range
Acquisition	22	-0.16	0.17	1.24
Execution	22	-0.12	0.19	1.23
Inspection	22	-0.15	0.15	1.19
Navigating	22	-0.12	0.16	1.25
Referencing	22	$-9.71 \cdot 10^{-2}$	0.17	1.19
Drawing	22	-0.11	0.17	1.19
Dimensioning	22	-0.14	0.14	1.11
Constraining	16	$-5.35 \cdot 10^{-2}$	0.17	1.52
Manipulating	22	-0.10	0.12	1.17
Deleting	4	$-3.96 \cdot 10^{-2}$	$9.91 \cdot 10^{-2}$	1.02

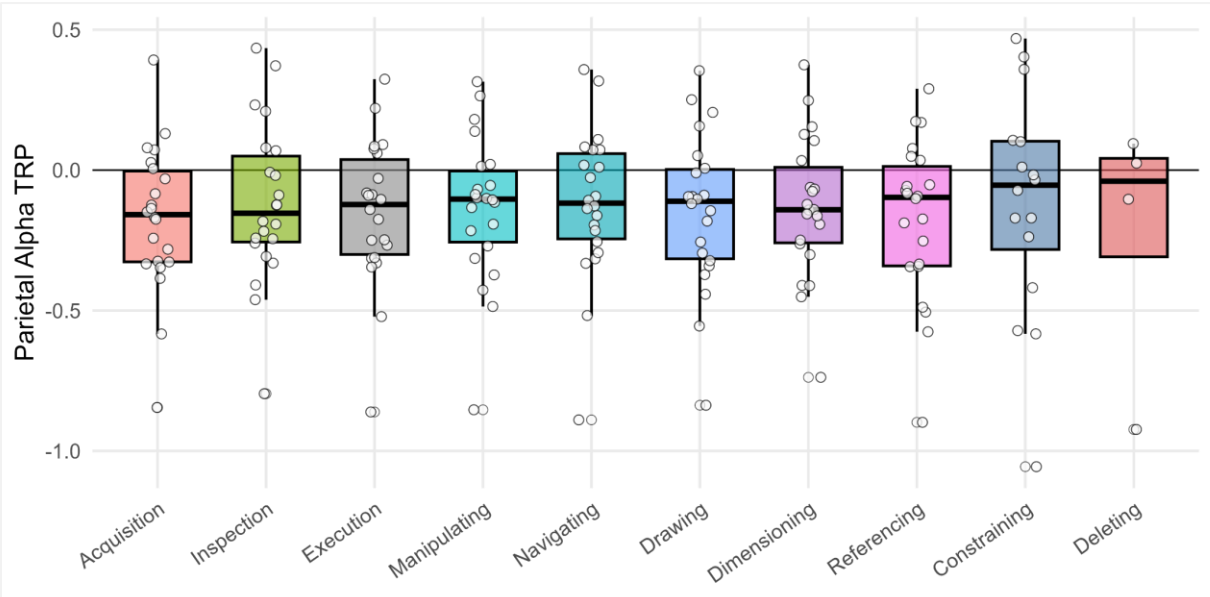

Figure 63 Median parietal alpha TRP across unit task phases and subtasks

Because *constraining* and *deleting* were not present for all participants (*constraining* was found in 16/22; *deleting* in 4/22 cases, according to Table 71), the subsequent analyses focused on five subtasks (*navigating, manipulating, referencing, drawing, dimensioning*) together with *acquisition* and *inspection*, yielding seven task segments with complete data. *Execution* was intentionally not included as a separate unit task phase at this level of analysis because it requires a non-overlapping segmentation; the intermediate-level subtasks constitute the decomposition of the *execution* phase, so including both *execution* and its constituent subtasks would double-count the same time intervals and confound comparisons. In contrast, *acquisition* and *inspection* were retained to preserve coverage of these two phases alongside the subtask-level decomposition of *execution*. The Friedman test across these seven segments showed no significant overall differences in parietal alpha TRP and a small effect size (Table 72). Post-hoc Wilcoxon comparisons likewise did not yield any differences that withstood Bonferroni correction for a multiple comparison (adjusted p > 0.05), indicating no reliable evidence that the analysed segments systematically elicited higher or lower parietal alpha TRP.

Table 72 Differences in parietal alpha TRP between acquisition, inspection, and subtasks

Test	**Parts**	**Statistic**	**Adjusted p-value**	**Effect size**
Friedman	All seven	9.88	0.13	$7.51 \cdot 10^{-2}$
Wilcoxon	Acquisition - Navigating	45	0.18	0.56
	Acquisition - Dimensioning	63	0.86	0.44
	Acquisition – Inspection	73	1	0.37
	Acquisition - Manipulating	69	1	0.40
	Acquisition – Drawing	87	1	0.27
	Acquisition - Referencing	124	1	$1.38 \cdot 10^{-2}$
	Inspection - Manipulating	123	1	$2.08 \cdot 10^{-2}$
	Inspection – Navigating	93	1	0.23
	Inspection - Drawing	139	1	$8.31 \cdot 10^{-2}$
	Inspection – Dimensioning	93	1	0.23
	Inspection – Referencing	162	1	0.24
	Manipulating – Navigating	102	1	0.17
	Manipulating – Drawing	153	1	0.18
	Manipulating – Dimensioning	124	1	$1.38 \cdot 10^{-2}$
	Manipulating – Referencing	158	1	0.22
	Navigating – Drawing	184	1	0.40
	Navigating – Dimensioning	141	1	$9.69 \cdot 10^{-2}$
	Navigating – Referencing	170	1	0.30
	Drawing – Dimensioning	105	1	0.15
	Drawing – Referencing	148	1	0.15
	Dimensioning - Referencing	167	1	0.28

For completeness and to demonstrate how EEG-derived intensity could be integrated into the utilisation-based model of CL, exploratory information-processing weights (ω) were computed for these seven task segments from robust within-participant z-scores mapped via the softmax function (Table 73). These weights represent a relative distribution of each participant's segment-wise TRP values (summing to one) rather than statistically confirmed phase differences. In line with the null inferential results (Table 72), agreement in phase ranking across participants was low (Kendall's W = $7.48 \cdot 10^{-2}$, χ^2 = 9.88, p = 0.13, df = 6), indicating that the implied ordering should be interpreted as a tendency rather than a stable ranking shared across participants.

Table 73 Information-processing weights assigned to acquisition, inspection, and subtasks

Unit task phase/Subtask	Information-processing weight (ω)
Navigating	0.183
Dimensioning	0.176
Manipulating	0.163
Inspection	0.149
Drawing	0.136
Referencing	0.101
Acquisition	$9.22 \cdot 10^{-2}$

6.4.3 Discussion

Since no significant differences were found between the analysed task segments, posed RQ 2.2 is not supported by the present data at these levels; no reliable segment-level differences were detected in information processing intensity between intermediate-level subtasks nor unit task phases, as measured by EEG-based CL indicator (parietal alpha TRP). This conclusion is supported by the non-significant Friedman tests, the non-significant Bonferroni-adjusted pairwise comparisons, and the low Kendall's W values, which together indicate that 1) segment differences are small relative to within- and between-participant variability and 2) participants do not share a stable ranking of segments by parietal alpha TRP.

The information-processing weights ω reported in Table 73 therefore should not be interpreted as stable, population-level parameters that quantify true intensity differences between segments. Instead, given the lack of statistical evidence for segment separation, these weights must be treated as exploratory, descriptive within-participant normalisations. In practical terms, ω expresses a participant's relative distribution of parietal alpha TRP across segments (a compositional profile summing to one) and is presented here primarily to demonstrate how a psychophysiological indicator could be integrated into the utilisation-based performance model when segment-level effects are empirically supported.

This caution is particularly important because the weight computation can be sensitive to analytical choices. The current procedure (robust z-scoring, trimming, and softmax mapping) is mathematically valid, but the resulting numerical weights can depend on the trimming threshold, the specific robust standardisation method (e.g. Med/MAD versus alternative robust estimators), and the softmax sensitivity. When group-level differences are weak and participant-specific ranking of segments vary (as suggested by the small effect sizes and low W), such transformation scan turn small fluctuations into seemingly precise weight differences. In addition, parietal alpha TRP showed bidirectional change patterns at the segment level; majority of the participants exhibited consistent increases relative to baseline, whereas other exhibited decreases (see Figure 61). This heterogeneity implies that “higher intensity” may not map monotonically onto a single signed direction of TRP change for all individuals. In principle, both stronger decreases and stronger increases could reflect elevated processing effort under different strategies, which would motivate alternative operationalisations (e.g. using absolute deviation from baseline or participant-specific sign conventions) and would directly affect the resulting segment ordering and weights. However, the present dataset does not provide sufficient evidence to justify such a redefinition, so the weights were computed using the signed TRP values as specified. Moreover, no segment-level subjective criterion is available for calibration in the present study because NASA TLX is obtained once per condition; consequently, the wights cannot be cross-validated against an external “ground truth” at the segment level. For these reasons, the values in Table 73 should be interpreted as tendencies rather than stable parameters, and they should not be used to make strong claims about which segment is “more intense” than another.

At the same time, the comparison between Table 56 and Table 73 remains conceptually informative because it highlights the distinction between two constructs that the thesis explicitly separates; utilisation (how much of a segment time is occupied by P/C activity) vs. information-processing intensity (how “costly” processing is per unit time, regardless of duration). Table 56 reflects a utilisation-based ranking derived from the CAD performance model, which conceptually elevates phases such as acquisition and inspection because they involve sustained engagement of perceptual and cognitive operators across large portions of time (assumed 100%). Table 73, in contrast, provides an exploratory ordering based on relative parietal alpha TRP variations. Here, *navigating* and *dimensioning* moved to the top (highest ω), suggesting that when these subtasks occur, they trigger relatively stronger information processing than what utilisations of P and C operators alone would imply (Table 56). In other words, information processing in these subtasks may be shorter but more intense (e.g. due to rapid recall of command locations/functions or maintaining spatial references). Conversely, *acquisition* drops to the bottom in Table 73, which is consistent with *acquisition* being time-dominant in utilisation but not necessarily the most intense per unit time in terms of brain activity. The fact that *manipulating* stayed near the top in both tables strengthens confidence that geometry manipulation is both perceptually and cognitively consuming as well as intense, whereas the reordering of *navigating* and *dimensioning* highlights subtasks that may be underestimated if we only look at operator utilisation. However, because this ordering is not statistically supported and agreement across participants is low, it should be treated as hypothesis-generating observation, not as evidence of stable segment-level intensity differences.

6.5 Relating utilisation-based and subjective CL assessments

This section evaluates how utilisation-based indicators of CL relate to subjective CL assessed with NASA TLX. Building on the previously identified EEG-based CL indicator and the derived information-processing weights, the analyses use the robust (Theil-Sen) linear regression to test whether utilisation-based indicators covary with self-reported CL at the CAD task level and whether they can support prediction of subjective CL. A specific objective is to examine whether utilisation-based indicators, initially computed at finer process levels, can be meaningfully summarised into global values while preserving (or strengthening) their association with NASA TLX mental demand, thereby supporting interpretable validation against as widely used subjective criterion.

In addition, this section assesses the value of incorporating EEG-derived information-processing weights into utilisation-based CL estimates. Conceptually, weighted utilisations operationalise the proposed integration of two complementary components: the duration of engagement of perceptual and cognitive resources captured by operator utilisation, and the information-processing intensity captured by EEG-derived weights. However, the preceding section showed that, in the present dataset, parietal alpha TRP did not yield statistically reliable differences between task segments and that agreement in segment ranking across participants was low. Therefore, the segment weights should be interpreted as information tendencies rather than stable parameters. For this reason, the validation is performed for both 1) raw utilisation-based indicators (without weights) and 2) weighted utilisation-based indicators, to determine whether the inclusion of EEG-derived weights improves correlation and explanation of NASA TLX variances beyond what is already captured by utilisation alone, while avoiding overinterpretation of potentially unstable segment-level weighting.

Because the method is intended to support both detailed participant-level analyses and practical, reproducible CL modelling, this section also evaluates a simplified approach based on global utilisations (across-participant medians per segment) and compares it to participant-specific (median) utilisation values. Within this framing, participant-specific median utilisations represent the most informative implementation of the method; they preserve individual differences in how engineering designers allocate perceptual and cognitive resources across CAD task segments and therefore maximise sensitivity to behavioural variability. However, they also increase practical burden, because they depend on participant-level CAD modelling outputs and may require more detailed segmentation and consistent annotation to remain comparable across subjects. In contrast, global utilisations (computed as across-participant median utilisation values for each segment) provide a deliberately simplified approximation; the same segment is assigned the same utilisation intensity for all participants, reducing the method to a stable "segment template" that can be reused for CAD modelling and visualisation. This simplification directly supports the usability objective in Table 3 by reducing the number of participant-specific parameters, improving transparency of the calculations, and making the approach easier to replicate across CAD tasks and studies.

Therefore, the comparison between global and participant-specific utilisations serves as an explicit test of the method's simplicity–fidelity trade-off. If global utilisations preserve similar relationships with subjective CL (NASA TLX) as participant-specific utilisations, then a researcher can use the simplified version when the goal is to obtain a robust, interpretable estimate of CL dynamics across typical CAD segments (without having to fully parameterise the model for each individual). If, however, global utilisations perform substantially worse, this would indicate that the method cannot be simplified without sacrificing validity and would imply that participant-specific CAD modelling is necessary for reliable CL estimation. In this way, analysis in this section operationalises the usability objective ("opt for the simplest metrics through which changes in CL can be reliably measured") as an empirical question rather than an assumption, which is consistent with the validation-driven methodology of the second experimental study.

6.5.1 Data analysis

First, the overall CL (at the level of the entire CAD modelling task) was computed using four utilisation-based indicators (Weighted Median U, Weighted Global U, Median U, Global U). The calculation was based on a non-overlapping segmentation of CAD performance into two unit task phases (*acquisition* and *inspection*) and the intermediate-level subtasks that represent the execution-level decomposition (*navigating*, *manipulating*, *referencing*, *drawing*, *dimensioning*, *constraining*, *deleting*). Overall CAD task CL (CL_{Task}) was then obtained by summing segment-level CL contributions weighted by each segment's relative duration within the task (as prescribed in Section 5.3.4). The computation of overall CL is described by Equation (47) for non-weighted utilisations and Equation (48) for weighted utilisations.

$$\mathrm{CL}_{Task}^{\mathrm{U}} = \sum_{s} \frac{T_s}{T_{Task}} \cdot CL_s = \sum_{s} \frac{T_s}{T_{Task}} \left(w_P U_{P,s} + w_C U_{C,s}\right), \tag{47}$$

T_s is duration of a task segment s

T_{Task} is duration of an entire CAD task

$w_P = 1,\ w_C = 2$

$U_{P,s}$ and $U_{C,s}$ are Med or global utilisations of P and C operators in s

$s \in S, S = \{$Acquisition, Inspection, Navigating, Manipulating, Referencing, Drawing, Dimensioning, Constraining, Deleting$\}$

$$\mathrm{CL}_{Task}^{\mathrm{U\omega}} = \sum_{s} \frac{T_s}{T_{Task}} \cdot CL_s = \sum_{s} \frac{T_s}{T_{Task}} \omega_s \left(w_P U_{P,s} + w_C U_{C,s}\right), \tag{48}$$

T_s is duration of a task segment s

T_{Task} *is* duration of an entire CAD task

ω_s is an information processing weight of a task segment s

$w_P = 1,\ w_C = 2$

$U_{P,s}$ *and* $U_{C,s}$ are Med or global utilisations of P and C operators in s

$s \in S, S = \{$Acquisition, Inspection, Navigating, Manipulating, Referencing, Drawing, Dimensioning, Constraining, Deleting$\}$

Here, Median U denotes a participant-specific utilisation summary (computed from that participant's segment-level utilisations), whereas Global U denotes a group-level utilisation value shared across participants (summarised across participants' segment-level utilisations). The global values are therefore computed once from the sample and then applied as fixed segment utilisation values; consequently, under the Global U formulation, between-participant variability in CL_{Task} is driven primarily by differences in segment durations/occurrences ($\frac{T_s}{T_{Task}}$) rather than by differences in segment utilisation levels. Used Global U and information-processing weight values are available in Table 74.

Table 74 Global utilisation values and information-processing weights

	Operator	Acquisition	Inspection	Navigating	Referencing	Drawing	Constraining	Dimensioning	Manipulating	Deleting
Global U	P	1	1	0.23	0.22	0.30	0.27	0.47	0.66	0.33
	C	1	1	0.23	0.22	0.30	0.29	0.46	0.36	0.33
Weight (ω)		$9.22 \cdot 10^{-2}$	0.149	0.183	0.101	0.136	-	0.176	0.163	-

Second, Theil–Sen regression models were developed to relate CL_{Task} values calculated using these four utilisation-based indicators and subjective CL assessed with NASA TLX (mental demand dimension). In these models, both "raw" utilisations and weighted utilisations (using EEG-derived information-processing weights) were considered as predictors. Theil–Sen regression was selected as a nonparametric, outlier-resistant alternative to ordinary least squares, providing a robust estimate of the monotonic linear trend between utilisation-based indicators and subjective CL without relying on normality assumptions.

6.5.2 Results

The resulting distributions of CL_{Task} based on four utilisation-based indictors are described in Table 75 with measures of descriptive statistics (M, SD, Med, MAD, and range) and visualised in Figure 64.

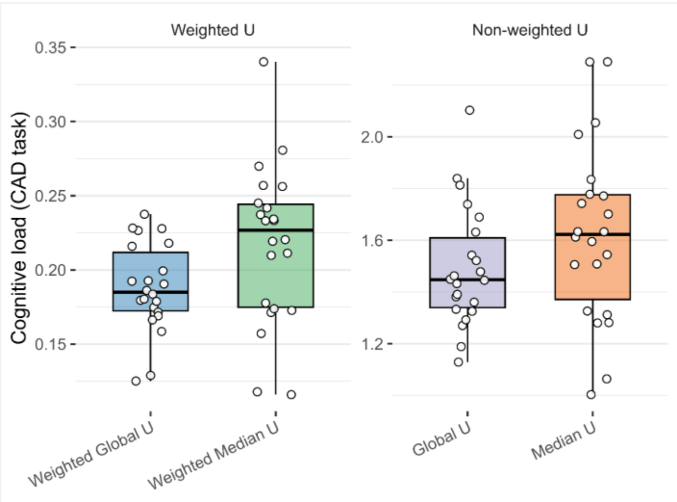

Figure 64 Overall CL values calculated from utilisation-based indicators

Table 75 Summary of overall CL values calculated from utilisations

Indicator	Range	Med	MAD	M	SD
Weighted Global	0.13 – 0.24	0.19	$2.2 \cdot 10^{-2}$	0.19	$3.00 \cdot 10^{-2}$
Weighted Median	0.12 – 0.34	0.23	$4.4 \cdot 10^{-2}$	0.22	$5.3 \cdot 10^{-2}$
Global U	1.13 – 2.10	1.45	0.18	1.49	0.23
Median U	1.00 – 2.29	1.62	0.27	1.63	0.34

A Theil–Sen regression was used to evaluate whether utilisation-based CL indicators covary with subjective CL assessed by the NASA TLX Mental Demand score across participants (one observation per participant in the analysed LC task condition). Four indicators were tested (Weighted Median U, Weighted Global U, Median U, Global U), and models were compared primarily using the median absolute error (MAE) from leave-one-out cross-validation

(LOOCV) as a robust measure of predictive tightness, while p-values were used to classify relationships as significant or non-significant (Table 76, Figure 65).

Table 76 Theil-Sen regression models: Utilisation-based indicators and Mental Demand (NASA TLX)

Utilisation-based indicator	Med intercept a	Med slope b	95% CI for b	p-value	MAE
Weighted Median U	6.04	6.10	4.46, 43.39	$2.14 \cdot 10^{-2}$	2.16
Weighted Global U	-0.73	39.23	12.83, 78.29	$1.47 \cdot 10^{-2}$	2.19
Global U	-5.92	8.20	6.40, 13.91	$2.31 \cdot 10^{-4}$	2.30
Median U	2.41	1.87	1.38, 6.55	$4.09 \cdot 10^{-3}$	2.93

All four indicators showed a significant positive association with Mental Demand (all $p < 0.05$). Among the weighted indicators, Weighted Median U yielded the lowest MAE, closely followed by Weighted Global U. Among the non-weighted indicators, Global U showed a comparatively low error (MAE = 2.30), whereas Median U produced the largest MAE but remained significant. Overall, the LOOCV results suggest that EEG-informed weighting provides a small reduction in prediction error relative to non-weighted variants, and that global utilisation summaries can approximate (or exceed) participant-specific Med summaries in this dataset, supporting their use for simplification when participant-specific utilisation estimates are unavailable.

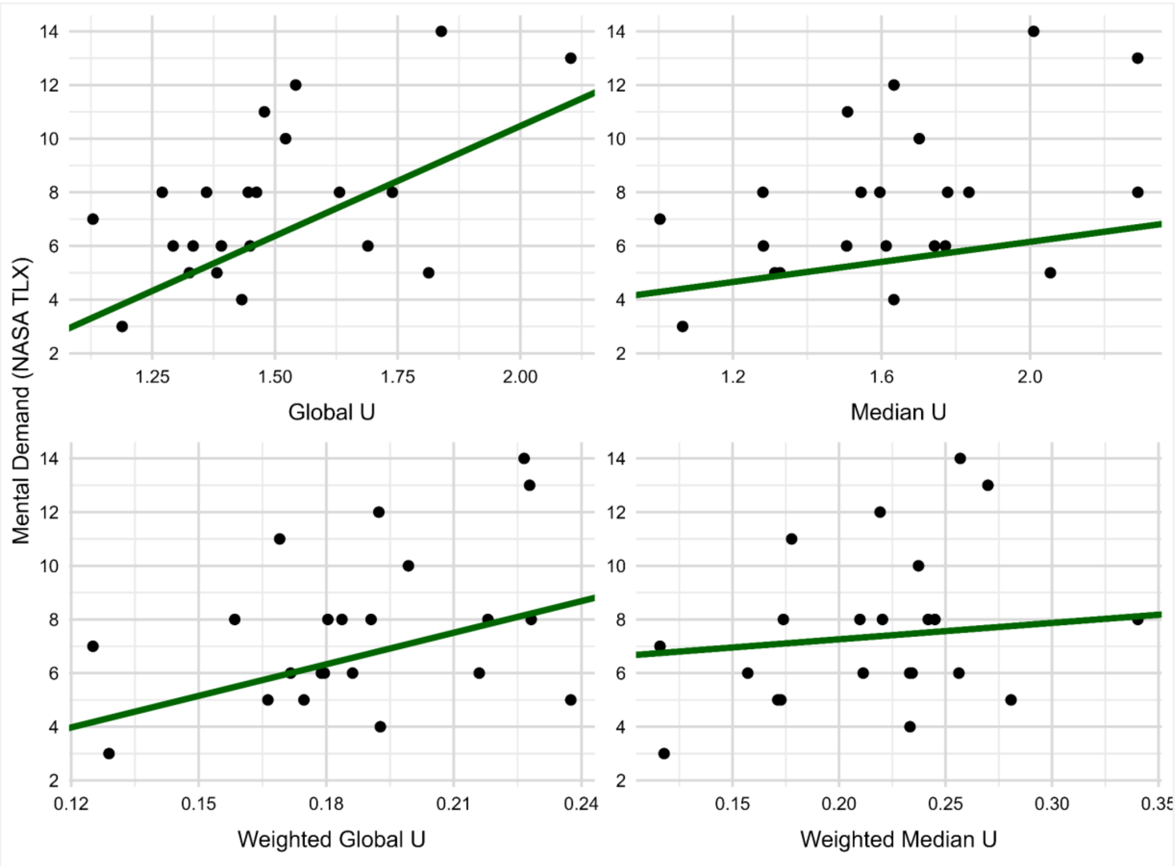

Figure 65 Visualisation of Theil-Sen regression models between utilisation-based indicators and NASA TLX scores

6.5.3 Discussion

The comparison of CL values calculated from utilisation-based indicator variants addressed two questions: 1) whether EEG-derived information-processing weighting adds value beyond raw utilisation, and 2) whether global (shared, participant-level) utilisation values can replace participant-specific utilisation summaries (median-based) for simplification.

This analysis assumed that aggregated subjective WL rating such as NASA TLX should show a monotonic and approximately linear relationship with overall P and C operator utilisation, because both aim to reflect how strongly processing resources are engaged during task execution [120]. In the present dataset, the Theil-Sen regressions support this interpretation; all four utilisation-based indicators exhibited a significant positive association with mental demand NASA TLX dimension, indicating that higher utilisation-based CL estimates correspond to higher self-reported subjective CL.

6.5.3.1 *How does CL calculated from weighted utilisations relate to self-reported subjective CL when compared to non-weighted utilisations?*

The results suggest that EEG-derived information-processing weights can yield incremental gains over raw utilisations in predicting NASA TLX Mental Demand, but the magnitude of improvement is modest, consistent with the limited stability of segment-level weights in this dataset. Using LOOCV MAE as the primary calibration criterion, the weighted indicators achieved slightly lower prediction error that their non-weighted counterparts. In particular, Weighted Median U yielded the lowest MAE, closely followed by Weighted Global U, while the best non-weighted alterative (Global U) showed a marginally higher MAE.

These results suggest that NASA TLX Mental Demand in this setting is driven primarily by sustained perceptual and cognitive engagement over the task, which non-weighted utilisation indicators capture alone. Weighting may still help by modestly correcting for segment-level differences in effective processing cost, but the improvement remains small, which is consistent with the earlier finding that EEG-derived information-processing weights are not stable at the segment level (low agreement across participants and non-significant segment differences). In other words, the weighting step appears directionally sensible as calibration refinement, but in this dataset it should be treated as supportive rather than decisive when relating to subjective CL.

6.5.3.2 *How does CL calculated from global utilisations relates to self-reported subjective CL when compared to median values?*

The median-versus-global comparison provides clearer guidance for potential simplification of the method. The non-weighted Median U showed substantially poorer calibration (higher MAE) than Global U, suggesting that, at least in this dataset, participant-specific Med aggregation does not improve (and may even reduce) the correspondence with subjective mental demand. For the weighted indicators, the difference between Weighted Median U and Weighted Global U was small (when comparing MAE), indicating that once a fixed segment weighting scheme is applied, using group-level utilisation summaries can approximate participant-specific summaries without sacrificing predictive accuracy.

Taken together, these results support using global utilisation values as a practical simplification when the goal is to compute utilisation-based CL indicators without requiring participant-specific utilisation estimates for each task segment. Weighting can be retained when slightly improved predictive fit is desired and when EEG-derived information-processing weights are available. However, given the limited stability of segment weights in this dataset, the main takeaway is that utilisation-based CL is robust even under simplified (global) aggregation, and weighting should be interpreted as an exploratory refinement rather than a necessary component.

6.6 Comparing CL between unit task phases/subtasks and their contribution to the overall CL

The proposed theoretical model of CL in CAD (Section 4.3) suggests that CL is not uniform across the CAD modelling process, but varies between recurring unit task phases (*acquisition*, *execution*, *inspection*) and functionally distinct intermediate-level subtasks within *execution* (e.g. *navigating*, *referencing*, *drawing*, *dimensioning*). Building on the preceding task-level analyses (Section 6.5), this section serves as a next validation step for the method; rather than evaluating whether utilisation-based CL measures distinguish overall CL between CAD tasks, it tests whether they can also reproduce meaningful process-level differences by localising CL to the segments prescribed by the CAD performance model. Specifically, this section first compares CL estimates between unit task phases and intermediate-level subtasks using both non-weighted and EEG-weighted utilisations. It then quantifies how much each segment contributes to the overall CAD-task CL once its time share is considered. Finally, in line with Section 6.5, this section evaluates whether the utilisation-based CL indicators support a simplified implementation when moving from participant-specific estimates (median utilisations) to global utilisations, now at the segment level and without NASA-TLX as an external comparison criterion. Specifically, the analysis examines whether median versus global utilisations (and their weighted versus non-weighted variants) produce a similar rank ordering of unit task phases and intermediate-level subtasks by implied CL (with parietal alpha TRP included as a comparison reference). If the ordering is preserved, global utilisations can be interpreted as a pragmatic approximation that retains the same qualitative conclusions about which unit task phases/subtasks are more versus less demanding, while reducing modelling and reporting complexity.

6.6.1 Data analysis

First, CL was calculated in each segment s for each participant using both median and global (across segment types and participants) values of utilisations (Equation (49)).

$$\mathrm{CL}_s^{\mathrm{U}} = w_P U_{P,s} + w_C U_{C,s} \tag{49}$$

$$w_P = 1,\ w_C = 2$$

$$U_{P,s} \text{ and } U_{C,s} \text{ are Med or global utilisations of } P \text{ and } C \text{ operators in } s$$

$$s \in S, S = \{\text{Acquisition, Inspection, Navigating, Manipulating, Referencing, Drawing, Dimensioning, Constraining, Deleting}\}$$

Second, these values were multiplied with information-processing weights ω_s assigned to each unit task phase and subtask s (Equation (50)).

$$\mathrm{CL}_s^{\mathrm{U}\omega} = \omega_s\left(w_P U_{P,s} + w_C U_{C,s}\right), \tag{50}$$

$$\omega_s \text{ is an information} - \text{processing weight of a task segment } s$$

$$w_P = 1,\ w_C = 2$$

$$U_{P,s} \text{ and } U_{C,s} \text{ are Med or global utilisations of } P \text{ and } C \text{ operators in } s$$

$$s \in S, S = \{\text{Acquisition, Inspection, Navigating, Manipulating, Referencing, Drawing, Dimensioning, Constraining, Deleting}\}$$

Third, absolute CL contribution (CLC) of each unit task phase and subtask was calculated by extending Equation (49) and Equation (50) to include a time share $\frac{T_s}{T_{Task}}$ of each segment, resulting in Equation (51) and Equation (52).

$$\mathrm{CLC}_s^{\mathrm{U}} = \frac{T_s}{T_{Task}}\left(w_P U_{P,s} + w_C U_{C,s}\right), \tag{51}$$

$$T_s \text{ is duration of a task segment } s$$

$$T_{Task} \text{ is duration of an entire CAD task}$$

$$w_P = 1,\ w_C = 2,$$

$$U_{P,s} \text{ and } U_{C,s} \text{ are Med or global utilisations of } P \text{ and } C \text{ operators in } s$$

$$s \in S, S = \{\text{Acquisition, Inspection, Navigating, Manipulating, Referencing, Drawing, Dimensioning, Constraining, Deleting}\}$$

$$\mathrm{CLC}_s^{\mathrm{U}\omega} = \frac{T_s}{T_{Task}} \omega_s \left(w_P U_{P,s} + w_C U_{C,s} \right), \tag{52}$$

T_s is duration of a task segment s

T_{Task} is duration of an entire CAD task

$w_P = 1,\ w_C = 2$

$U_{P,s}$ and $U_{C,s}$ are Med or global utilisations of P and C operators in s

$s \in S, S = \{\text{Acquisition, Inspection, Navigating, Manipulating, Referencing, Drawing, Dimensioning, Constraining, Deleting}\}$

Fourth, relative CLC of each segment (unit task phase or subtask) to the overall CL CL_{Task} was calculated as a ratio of a CL assigned to the segment and the overall CL, according to Equation (53):

$$Relative\ CLC_s = \frac{CLC_s}{CL_{Task}} = \frac{CLC_s}{\sum_{s \in S} CLC_s} \tag{53}$$

$s \in S, S = \{\text{Acquisition, Inspection, Navigating, Manipulating, Referencing, Drawing, Dimensioning, Constraining, Deleting}\}$.

Overall CL (CL_{Task}) was calculated in Section 6.4 according to Equation (47).

Once CL and CLC values were calculated across the segments and participants, Friedman test was conducted to quantitatively compare differences in their values between the unit task phases as well as between subtasks. This comparison was done for CL calculated from Med values of utilisations, while it was omitted for global values as they were the same for all the participants. Comparison of segments based on their CLC was done for all four utilisation-based indicators since time variable differentiates participants' results even in cases with global utilisation values. The effect size of differences determined with the Friedman test was computed using Kendall's W.

Finally, post-hoc tests involved pairwise comparisons between the unit task phases as well as the subtasks using the Wilcoxon signed-rank test with Bonferroni adjusted p-values. The effect size for the Wilcoxon test was calculated as the Z statistic divided by the square root of the sample size. Measures of descriptive statistics, significant differences (adjusted p-values), and associated effect sizes are coupled with the test statistic values; marked with F for the Friedman test and V for the Wilcoxon signed-rank test.

6.6.2 Results

Results are divided into two sections; first Section 6.5.2.1 presents CL associated with unit task phases and subtasks (Section 6.6.2.1) and then Section 6.5.2.2 shows differences in CLC of these task segments. The results are depicted visually using boxplots and detailed numerically in tables. In addition to CL based on utilisations, each plot is accompanied by CL indicated by parietal alpha TRP for the rank comparison purposes.

6.6.2.1 CL associated with subtasks

Friedman test showed that differences in the rankings of subtasks were significant when indicated both by Median U and its weighted version, as presented in Table 77.

Table 77 Friedman test: Differences in CL associated with acquisition, inspection, and subtasks

Indicator	F statistic	p-value	Effect size
Median U	27.40	$6.03 \cdot 10^{-4}$	0.18
Weighted Median U	23.40	$2.89 \cdot 10^{-3}$	0.15

In addition, pairwise post-hoc comparisons of CL (Table 78) showed that the segments significantly differed at the level of individual pairs. When CL was indicated by Median U, *acquisition* differed significantly from all other segments (*inspection* and the five subtasks). For Weighted Median U, *acquisition's* CL remained significantly different from *inspection* and from *referencing*, *drawing*, and *dimensioning*, while its differences with *navigating* and *manipulating* were no longer significant. Inspection's CL was significantly higher or lower than the CL of *referencing*, *drawing*, and *manipulating* for both indicators; when CL was expressed by Weighted Median U, *inspection* also differed from *navigating* and *dimensioning*. CL associated with *navigating* significantly differed from that during *referencing* and *drawing* for both indicators, and additionally from *inspection* when CL was indicated by Weighted Median U.
CL associated with *referencing* was distinct from *acquisition*, *inspection*, *navigating*, *dimensioning*, and *manipulating* irrespective of the CL indicator. Moreover, its CL differed from *drawing* when CL was based on Weighted Median U, whereas this contrast was not detected for Median U. Drawing's CL significantly differed from *acquisition*, *inspection*, *navigating*, and *dimensioning* across both indicators; in addition, Weighted Median U revealed further differences between *drawing* and *referencing* and between *drawing* and *manipulating*. *Dimensioning* showed significantly different CL compared with *acquisition*, *referencing*, and *drawing* for both indicators, and additionally with *inspection* when CL was indicated by Weighted Median U, while its contrasts with *navigating* and *manipulating* were not significant. Finally, *manipulating's* CL differed from *acquisition*, *inspection*, and *referencing* when

measured by Median U, but only from *inspection*, *referencing*, and *drawing* when CL was based on Weighted Median U; no significant differences were found between *manipulating* and *dimensioning* for either indicator. Across all significant pairwise contrasts, effect sizes were large (r ≥ 0.74) and in most cases very large (r ≥ 0.80), indicating substantial differences in CL between the segments.

Table 78 Significant pairwise differences in CL associated with acquisition, inspection, and subtasks

Indicator	Pair	V statistic	Adj. p-value	Effect size
Median U	Acquisition-Inspection	189	$6.05 \cdot 10^{-3}$	0.88
	Acquisition-Navigating	189	$6.05 \cdot 10^{-3}$	0.88
	Acquisition-Referencing	190	$5.15 \cdot 10^{-3}$	0.88
	Acquisition-Drawing	190	$5.15 \cdot 10^{-3}$	0.88
	Acquisition-Dimensioning	190	$5.15 \cdot 10^{-3}$	0.88
	Acquisition-Manipulating	190	$5.15 \cdot 10^{-3}$	0.88
	Inspection-Referencing	190	$5.15 \cdot 10^{-3}$	0.88
	Inspection-Drawing	185	$1.14 \cdot 10^{-2}$	0.83
	Inspection-Manipulating	183	$1.55 \cdot 10^{-2}$	0.81
	Navigating-Referencing	190	$5.15 \cdot 10^{-3}$	0.88
	Navigating-Drawing	188	$7.10 \cdot 10^{-3}$	0.86
	Referencing-Dimensioning	0	$5.15 \cdot 10^{-3}$	0.88
	Referencing-Manipulating	15	$4.96 \cdot 10^{-2}$	0.74
	Drawing-Dimensioning	0	$5.15 \cdot 10^{-3}$	0.88
Weighted Median U	Acquisition-Inspection	0	$5.15 \cdot 10^{-3}$	0.88
	Acquisition-Referencing	190	$5.15 \cdot 10^{-3}$	0.88
	Acquisition-Drawing	184	$1.33 \cdot 10^{-2}$	0.82
	Acquisition-Dimensioning	15	$4.96 \cdot 10^{-2}$	0.74
	Inspection-Navigating	188	$7.10 \cdot 10^{-3}$	0.86
	Inspection-Referencing	190	$5.15 \cdot 10^{-3}$	0.88
	Inspection-Drawing	190	$5.15 \cdot 10^{-3}$	0.88
	Inspection-Dimensioning	188	$7.10 \cdot 10^{-3}$	0.86
	Inspection-Manipulating	182	$1.80 \cdot 10^{-2}$	0.80
	Navigating-Referencing	190	$5.15 \cdot 10^{-3}$	0.88
	Navigating-Drawing	190	$5.15 \cdot 10^{-3}$	0.88
	Referencing-Drawing	13	$3.74 \cdot 10^{-2}$	0.76
	Referencing-Dimensioning	0	$5.15 \cdot 10^{-3}$	0.88
	Referencing-Manipulating	0	$5.15 \cdot 10^{-3}$	0.88
	Drawing-Dimensioning	0	$5.15 \cdot 10^{-3}$	0.88
	Drawing-Manipulating	2	$5.15 \cdot 10^{-3}$	0.88

At the subtask level, Weighted Median U and Median U indicated the highest CL for *dimensioning* and *manipulating*, followed by *navigating*. *Deleting* occupied a mid-range position, whereas *drawing*, *constraining*, and especially *referencing* tended to show lower median utilisation values (see Figure 66). Therefore, both Median U and Weighted Median U ranked subtasks from highest to lowest CL in a similar sequence (with *dimensioning* and *manipulating* at the upper end and *referencing*/*drawing* at the lower end), but the weighted metric produced a steeper gradient: differences between neighbouring subtasks were more pronounced in Weighted Median U than in Median U.

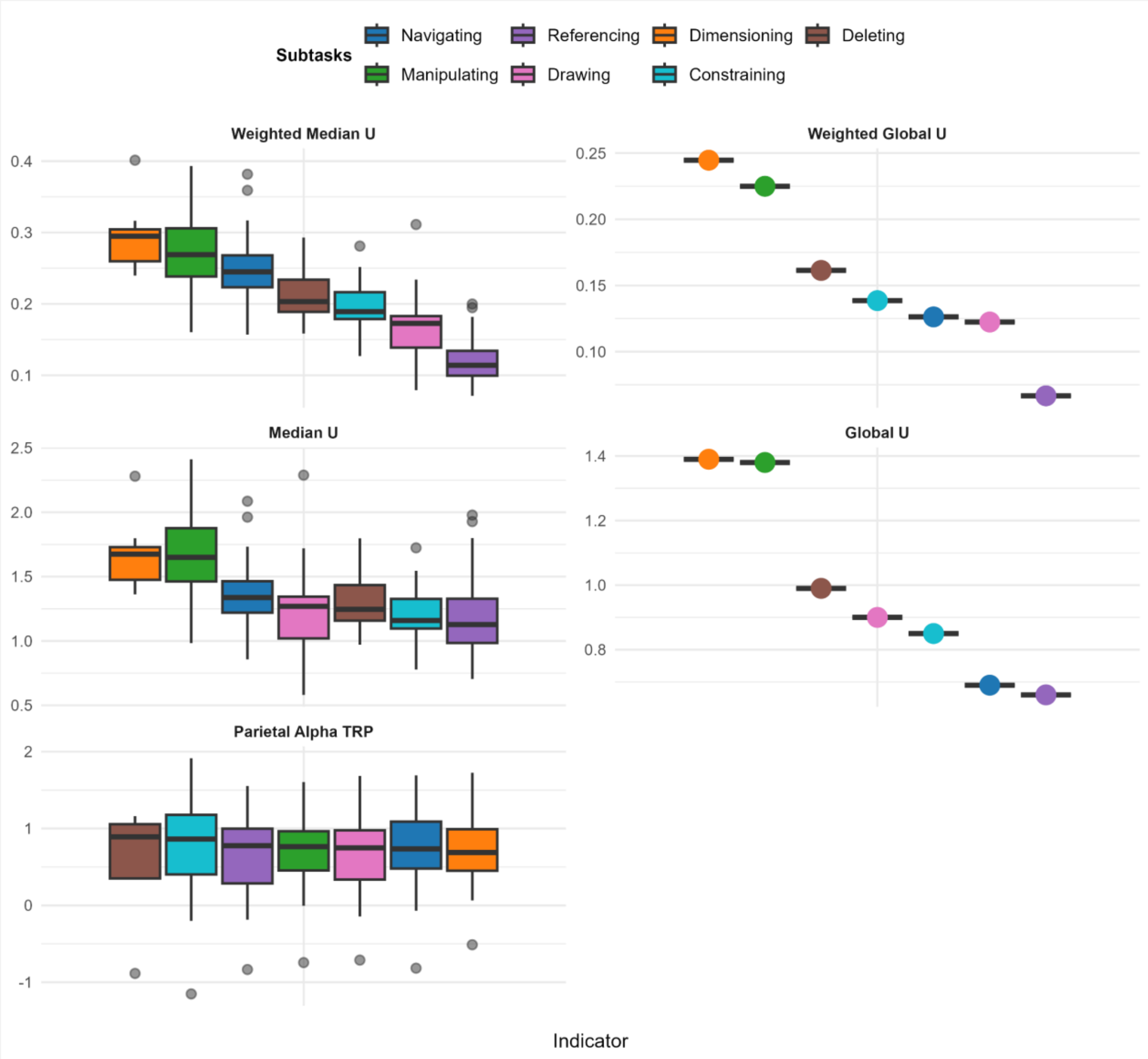

Figure 66 Differences in CL assigned to subtasks when assessed by different indicators

A similar pattern was observed when comparing Weighted Global U to Global U. The two global indices agreed on the ordering of most of the subtasks based on their CL (high values for *dimensioning*/*manipulating*, mid-range for *deleting*, lower values for *navigating*, *drawing*, *constraining,* and *referencing*.

For Parietal Alpha TRP, the subtask distributions were again more compressed, with considerable overlap between boxes. Nevertheless, there was a tendency for *deleting*, *constraining*, and *navigating* to show slightly higher median TRP than *drawing*, *manipulating*, *referencing*, and *dimensioning*, suggesting a somewhat different emphasis of the EEG-based indicator compared to the utilisation metrics. Overall, these plots show that the utilisation-base indicators yielded mostly consistent ranking of both unit task phases and subtasks by CL, which diverges from the ordering indicated by parietal alpha TRP.

6.6.2.2 CLC of subtasks to the overall CL

The results of Friedman test, detailed in Table 79, revealed significant differences between subtasks in their CLCs to the overall CL, regardless of the selected utilisation indicator.

Table 79 Friedman test: Differences in CLC of acquisition, inspection, and unit task phases to the overall CL

Indicator	Statistic	p-value	Effect size
Global U	27.53	$5.72 \cdot 10^{-4}$	0.18
Median U	27.67	$5.42 \cdot 10^{-4}$	0.18
Weighted Global U	27.13	$6.71 \cdot 10^{-4}$	0.18
Weighted Median U	26.7	$7.65 \cdot 10^{-4}$	0.18

In addition, differences remained significant at the level of pairwise comparisons, as detailed in Table 80. *Acquisition* significantly differed with *inspection* in most of the cases; the exception is CLC calculated from Weighted Global U. Moreover, its CLC was significantly differed than CLC of all five included subtasks in most of the cases; the only exception in this case was similarity with *navigating* when indicated by Weighted Med U. Inspection's CLC significantly differed from the contribution of *referencing* and *drawing* in all cases. However, *manipulating* showed significant distinction in CLC from Inspection when indicated by non-weighted indicators, while the significance was omitted for the weighted utilisations. CLC of *navigating* significantly differed from the CLC of at least two of the following subtasks: *referencing*, d*rawing*, *manipulating*. Moreover, CLC of *referencing* was distinguished from the CLC of *navigating*, *drawing*, *dimensioning*, and *manipulating* regardless of the used utilisation indicator. In addition, it significantly differed from *drawing* when CLC was indicated by Global U values (regardless of the weight). Similarly, significant differences in CLC were found between *drawing* and *dimensioning* for all indicators. However, *dimensioning* significantly differed from *manipulating* only when CLC was indicated by Global U values.

Table 80 Significant pairwise differences in CLC of subtasks to the overall CL

Indicator	Pair	Statistic	Adj. p-value	Effect size
Global U	Acquisition-Inspection	187	$8.33 \cdot 10^{-3}$	0.85
	Acquisition-Navigating	190	$5.15 \cdot 10^{-3}$	0.88
	Acquisition-Referencing	190	$5.15 \cdot 10^{-3}$	0.88
	Acquisition-Drawing	190	$5.15 \cdot 10^{-3}$	0.88
	Acquisition-Dimensioning	190	$5.15 \cdot 10^{-3}$	0.88
	Acquisition-Manipulating	190	$5.15 \cdot 10^{-3}$	0.88
	Inspection-Referencing	189	$6.05 \cdot 10^{-3}$	0.87
	Inspection-Drawing	185	$1.14 \cdot 10^{-2}$	0.83
	Inspection-Manipulating	184	$1.33 \cdot 10^{-2}$	0.82
	Navigating-Referencing	190	$5.15 \cdot 10^{-3}$	0.88
	Navigating-Drawing	182	$1.80 \cdot 10^{-2}$	0.80
	Referencing-Drawing	8	$1.80 \cdot 10^{-2}$	0.80
	Referencing-Dimensioning	0	$5.15 \cdot 10^{-3}$	0.88
	Referencing-Manipulating	5	$1.14 \cdot 10^{-2}$	0.83

	Drawing-Dimensioning	0	$5.15 \cdot 10^{-3}$	0.88
	Dimensioning-Manipulating	179	$2.80 \cdot 10^{-2}$	0.78
Median U	Acquisition-Inspection	189	$6.05 \cdot 10^{-3}$	0.87
	Acquisition-Navigating	183	$1.56 \cdot 10^{-2}$	0.81
	Acquisition-Referencing	190	$5.15 \cdot 10^{-3}$	0.88
	Acquisition-Drawing	190	$5.15 \cdot 10^{-3}$	0.88
	Acquisition-Dimensioning	190	$5.15 \cdot 10^{-3}$	0.88
	Acquisition-Manipulating	190	$5.15 \cdot 10^{-3}$	0.88
	Inspection-Referencing	189	$6.05 \cdot 10^{-3}$	0.87
	Inspection-Drawing	177	$3.74 \cdot 10^{-2}$	0.76
	Navigating-Referencing	190	$5.15 \cdot 10^{-3}$	0.88
	Navigating-Drawing	185	$1.14 \cdot 10^{-2}$	0.83
	Navigating-Manipulating	187	$8.33 \cdot 10^{-3}$	0.85
	Referencing-Dimensioning	0	$5.15 \cdot 10^{-3}$	0.88
	Referencing-Manipulating	15	$4.96 \cdot 10^{-2}$	0.74
	Drawing-Dimensioning	0	$5.15 \cdot 10^{-3}$	0.88
Weighted Global U	Acquisition-Navigating	187	$8.33 \cdot 10^{-3}$	0.85
	Acquisition-Referencing	190	$5.15 \cdot 10^{-3}$	0.88
	Acquisition-Drawing	190	$5.15 \cdot 10^{-3}$	0.88
	Acquisition-Dimensioning	190	$5.15 \cdot 10^{-3}$	0.88
	Acquisition-Manipulating	190	$5.15 \cdot 10^{-3}$	0.88
	Inspection-Referencing	190	$5.15 \cdot 10^{-3}$	0.88
	Inspection-Drawing	187	$8.33 \cdot 10^{-3}$	0.85
	Inspection-Manipulating	181	$2.09 \cdot 10^{-2}$	0.79
	Navigating-Drawing	185	$1.14 \cdot 10^{-2}$	0.83
	Navigating-Manipulating	178	$2.80 \cdot 10^{-2}$	0.78
	Referencing-Drawing	2	$7.09 \cdot 10^{-3}$	0.86
	Referencing-Dimensioning	0	$5.15 \cdot 10^{-3}$	0.88
	Referencing-Manipulating	1	$6.05 \cdot 10^{-3}$	0.87
	Drawing-Dimensioning	0	$5.15 \cdot 10^{-3}$	0.88
	Dimensioning-Manipulating	181	$2.09 \cdot 10^{-2}$	0.79
Weighted Median U	Acquisition-Inspection	179	$2.80 \cdot 10^{-2}$	0.78
	Acquisition-Referencing	190	$5.15 \cdot 10^{-3}$	0.88
	Acquisition-Drawing	190	$5.15 \cdot 10^{-3}$	0.88
	Acquisition-Dimensioning	184	$1.33 \cdot 10^{-2}$	0.82
	Acquisition-Manipulating	187	$8.33 \cdot 10^{-3}$	0.85
	Inspection-Referencing	189	$6.05 \cdot 10^{-3}$	0.87
	Inspection-Drawing	178	$3.24 \cdot 10^{-2}$	0.77
	Navigating-Referencing	190	$5.15 \cdot 10^{-3}$	0.88
	Navigating-Drawing	190	$5.15 \cdot 10^{-3}$	0.88
	Navigating-Manipulating	189	$6.05 \cdot 10^{-3}$	0.87
	Referencing-Dimensioning	0	$5.15 \cdot 10^{-3}$	0.88
	Referencing-Manipulating	4	$9.74 \cdot 10^{-3}$	0.84
	Drawing-Dimensioning	0	$5.15 \cdot 10^{-3}$	0.88

Across the utilisation-based indicators, the ordering of subtasks was mostly consistent, as shown in Figure 67. For both Weighted Median U and Median U, three groups can be distinguished:

1) *Dimensioning* and *navigating* had the highest Med values, representing the high CLC of these subtasks to the overall CL;
2) *Manipulating* and *drawing* occupied an intermediate band;

3) *Referencing*, *constraining*, and *deleting* showed the lowest Meds, and thus the lowest CLC to the overall CL.

The global measures (Weighted Global U and Global U) reproduced this hierarchy at the level of aggregated (global) utilisation: *dimensioning* and *navigating* again appeared at the top (although in reversed order), *drawing/manipulating* in the middle, and *constraining/referencing/deleting* at the bottom. Thus, when normalised by time shares, all utilisation metrics ranked subtasks' CLC from highest to lowest in essentially the same way; only the order of *navigating* and *dimensioning* was reversed.

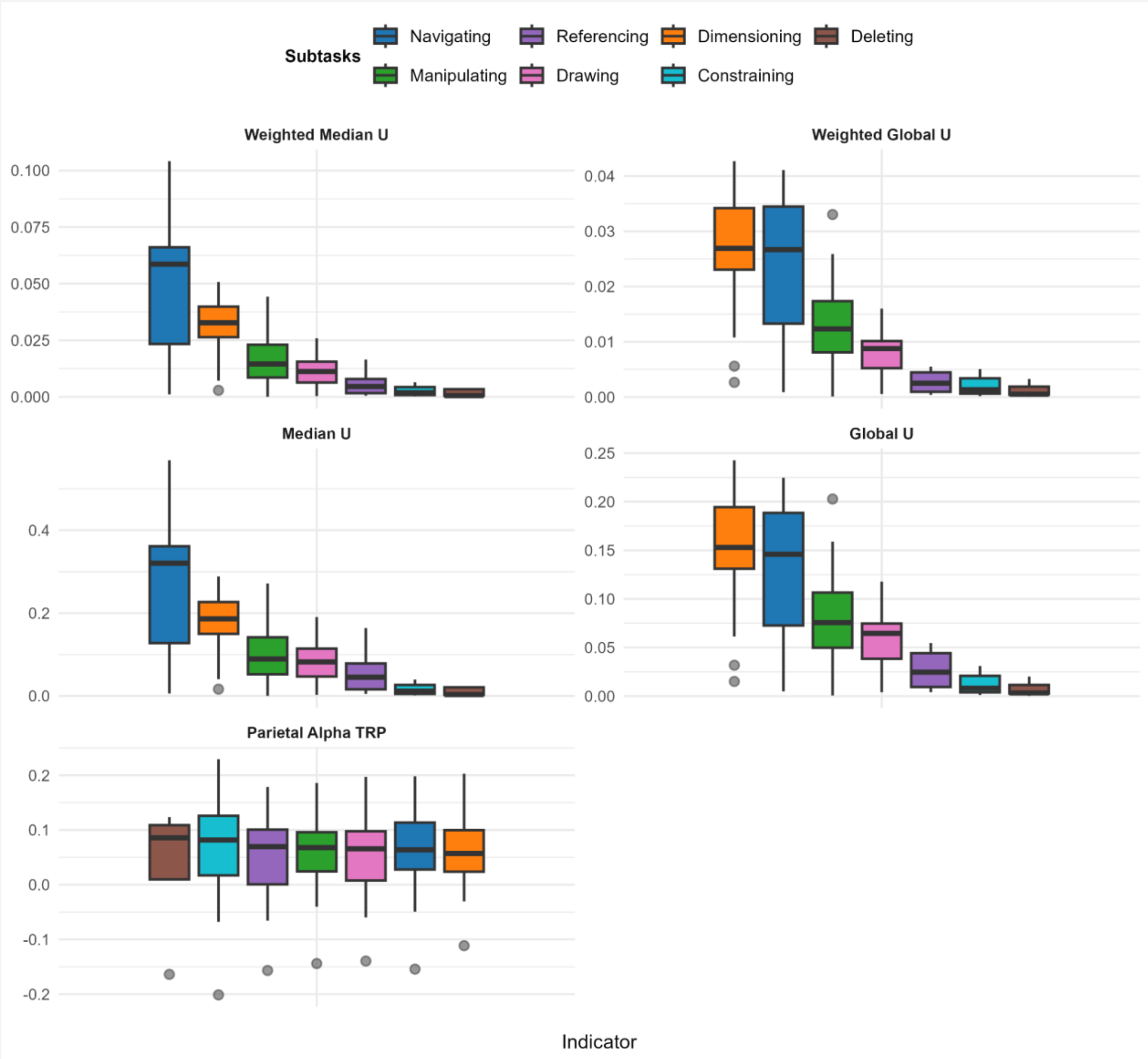

Figure 67 Differences in CL assigned to subtasks when assessed by different indicators and considering time share

Comparing weighted vs. unweighted indicators showed that weighting amplifies contrasts but does not change the rank order. Weighted Median U spread the high-CLC subtasks (*dimensioning*, *navigating*) further away from the low-CLC group than Median U, and Weighted Global U did the same relative to Global U. In other words, weighting emphasised subtasks that combine high operator utilisation with higher information-processing weights, making "dense" subtasks stand out more clearly while preserving the underlying ordering.

In terms of ordering, the utilisation-based indicators and parietal alpha TRP behaved differently. Utilisation metrics (Weighted Median U, Weighted Global U, Median U, Global U) at the subtask level agreed on the same hierarchy, with *dimensioning* and *navigating* having highest, *drawing* and *manipulating* medium, and *constraining*, *referencing*, and *deleting* the lowest information-processing intensity. Weighting (weighted versus non-weighted utilisations) mainly stretched the spacing between groups but did not change the order. Parietal Alpha TRP did not reproduce this order indicated by utilisations.

The stacked bars in Figure 68 highlight how integrating utilisation with time shares changes the relative contribution of subtasks for the overall task CL. Time share alone was dominated by a few subtasks (e.g. *navigating* and *dimensioning*), with rare subtasks such as *deleting* contributing very little. When expressed as shares of cumulative utilisation (for all the four utilisation-based metrics), higher-intensity subtasks like *dimensioning* and *manipulating* gained a larger proportion of the total CL than their raw duration would suggest, whereas time-consuming but lower-intensity subtasks such as *referencing* and *drawing* lost relative contribution.

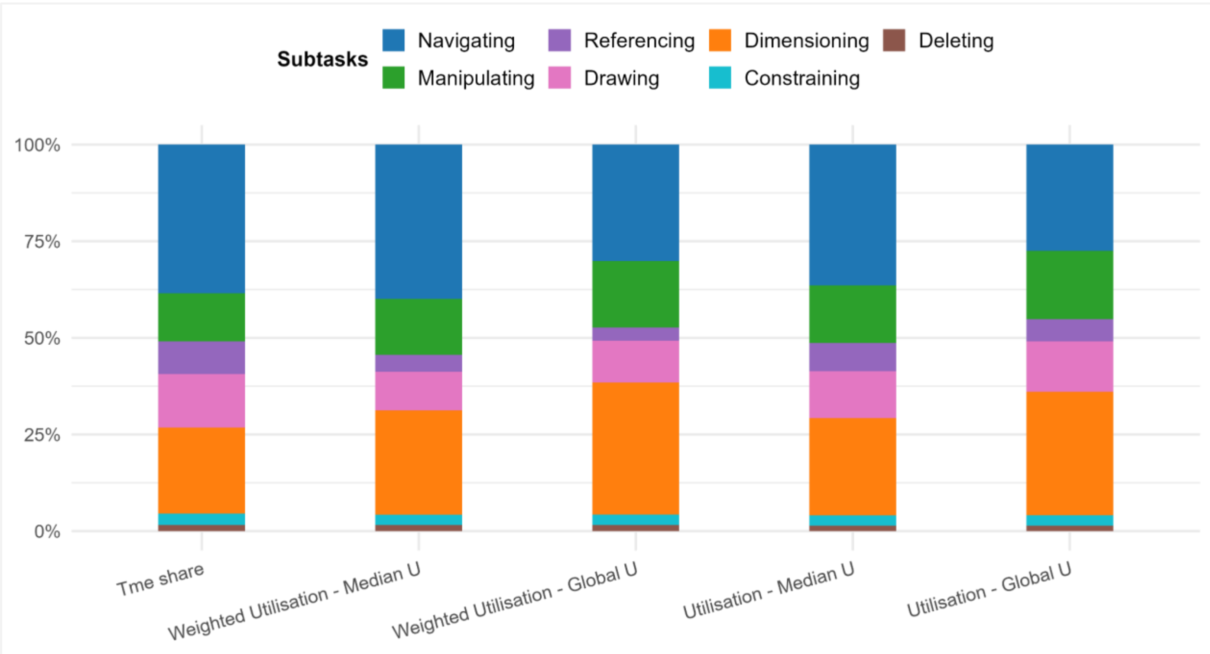

Figure 68 Subtasks' contribution to the overall CL

6.6.3 Discussion

Prior work has shown that higher-level design activities (e.g. understanding, idea generation, and evaluation) can impose different cognitive load (CL) on engineering designers [141]. Extending this view to CAD modelling, the presented results indicate that CL is not uniform across the CAD modelling process, even when examined at finer, CAD-task-specific levels such as unit task phases and intermediate-level subtasks. Thus, segmenting CAD modelling into unit task phases and subtasks is not only descriptively meaningful, but also captures differences in implied CL.

6.6.3.1 Do unit task phases/subtasks differ in CL implied by utilisations?

At the level of analysed unit task phases and intermediate-level subtasks, the presence of many significant pairwise contrasts (with mostly very large effects) indicates that these task segments capture distinct CL demands. The utilisation patterns imply that some subtasks (notably *dimensioning* and *manipulating*) tend to be associated with higher CL, plausibly because they require maintaining and applying precise rules and parameters while coordinating geometry, constraints, and values. In contrast, subtasks such as *referencing* and *drawing* tend to appear at the lower end of implied CL, which may indicate that these CAD operations/subtasks rely more on externally available information or more routine perceptual–motor sequences (even if they remain essential for the CAD modelling progress). Importantly, using weighted variants of utilisation does not fundamentally change the implication that subtasks differ; rather, it tends to sharpen the separations for some contrasts (a steeper gradient), while reducing others (e.g. *acquisition* becoming less distinct from some subtasks), suggesting that weighting changes how strongly differences are expressed rather than whether differences exist.

Overall, these findings support the conclusion that CAD modelling contains multiple task segments that impose measurably different CL, and that utilisation-based indicators are sensitive enough to detect these differences at both hierarchical levels.

6.6.3.2 What is the contribution of each unit task phase/subtask to the overall CAD task CL?

Interpreting contribution to overall CL requires moving beyond how demanding a segment is when it occurs and accounting for how much it occupies the CAD modelling process. The CLC analysis therefore provides a process-level view in which each unit task phase/subtask contributes to the overall CL as a function of both utilisation and exposure (time share in a total CAD task). This distinction is important in CAD, where some segments may be perceptually and cognitively dense (indicated by utilisations) but short when compared to the entire CAD task, whereas others may be less dense yet dominate total time.

At the level of subtasks, CLCs differed systematically, indicating that the overall CL is not only driven by where time is spent but also by which operations lead perceptual and cognitive activity. Across utilisation indicators, a stable hierarchy emerged: *dimensioning* and *navigating* contribute most, *manipulating* and *drawing* form a mid-band, and *referencing*, *constraining*, and *deleting* contribute least. This pattern suggests that subtasks like *dimensioning* and *navigating* combine relatively frequent engagement and/or operator activity with cognitively demanding decisions (e.g. selecting, applying rules, maintaining spatial/parametric context), making them high contributors to overall CL. Conversely, subtasks such as *deleting* tend to remain low impact because they are either infrequent or less demanding in terms of sustained P/C engagement.

6.6.3.3 Is the order of CL associated to each unit task phase/subtask preserved when using different utilisation-based indicators?

The ordering was largely preserved within the family of utilisation-based indicators, but not preserved when compared to the EEG-based indicator (parietal alpha TRP).

At the level of subtasks, utilisation-based indicators produced a mostly consistent hierarchy. Across median- and global-based utilisations, high-CL subtasks remained at the top and low-CL subtasks at the bottom, with only minor local changes (most notably a swap between *navigating* and *dimensioning* depending on whether the indicator was median- or global-based). Again, weighting did not reorder subtasks, but primarily stretched the distances between them, sharpening the gradient from high- to low-CL subtasks.

In contrast, parietal alpha TRP did not reproduce these utilisation-based orderings, neither for unit task phases nor for subtasks. TRP showed more overlap between distributions and a different emphasis in relative ordering, indicating weaker discrimination and/or sensitivity to different cognitive mechanisms than those captured by utilisation of P and C operators. Taken together, the results imply that the preservation of order depends on the indicators being compared; it is robust within utilisation-based modelling, but it is not robust across measurement modalities, which supports the interpretation that utilisation and EEG reflect overlapping yet non-identical aspects of CL in CAD modelling.

6.7 Modelling the dynamic changes in CL throughout CAD modelling

This section extends the analyses presented in Section 6.5 and Section 6.6 from entire-task and task-segment summaries to a time-resolved perspective by modelling how CL evolves throughout the CAD modelling session using a moving-window approach. Dynamic CL was computed for each participant (n=22) using time series of three indicators: 1) parietal alpha TRP, 2) utilisation-based CL derived from participant-specific median operator utilisations ("raw utilisations"), and (3) the proposed integrated measure based on weighted utilisations, where participant-specific median utilisations are multiplied by information-processing weights empirically derived from differences in parietal alpha TRP. For all three indicators, CL time series were calculated using a sliding window (10 s window, 2 s step) to capture within-task fluctuations and enable direct comparison of temporal patterns across indicators.

A 10-s moving window was selected to ensure that each utilisation estimate is computed from a time span that typically contains enough observable CAD activity to be meaningful, given the duration of the modelled task segments. In the screen recordings, many unit task phases and subtasks extend over several seconds to tens of seconds, and within shorter intervals the distribution of P and operators can be sparse (e.g. brief cursor movements or isolated clicks) and therefore unstable. Using a 10-s moving window increases the likelihood that a window includes a representative mix of P and activity (e.g. visual search and interpretation, decision/selection, and execution), so the resulting utilisation values are less sensitive to momentary pauses, micro-actions, or single events. At the same time, 10 s remains short enough to reflect within-task fluctuations rather than collapsing activity over an entire unit task phase or the full task. The step size of 2 s was chosen to preserve temporal resolution while maintaining robustness; with a 10 s moving window this corresponds to an 80% overlap, which provides a smoother utilisation time series and allow changes to be detected without requiring segmentation boundaries to align perfectly with the sampling frames. In combination, a 10-s moving window with a 2-s step (80% overlap) provides a practical compromise between stability of estimates (sufficient activity per window, aligned with typical segment durations) and sensitivity to changes in modelling behaviour over time.

6.7.1 Data analysis

Dynamic CL time series were computed per each participant (n=22) for all three indicators on the same sliding-window grid (window width of 10 seconds, and step 2 seconds, i.e. 80% overlapping windows). The window index is denoted by *n*, and *Tn* is the window length.

For each window *n,* parietal alpha TRP was summarised as the median TRP value within the window, yielding a windowed EEG-based CL estimate (Equation (54)):

$$\mathrm{CL}_n^{\mathrm{EEG}} = Med\,(\mathrm{TRP}_n) \tag{54}$$

TRP_n is a TRP of selected EEG feature in a time window n

For later concordance analyses, TRP was sign-flipped so that higher values consistently indicate higher CL, and then standardised (see below).

For each window *n*, utilisations of perceptual and cognitive operators were computed as the proportion of time the corresponding operator was active within the window (Equation (55)):

$$U_{i,n} = \frac{\int_0^{\mathrm{T}_n} A_i(t) \cdot dt}{\mathrm{T}_n} \tag{55}$$

T_n is duration of a time window n

$$U_{i,n} \in [0,1], \qquad i = P, C$$

Windowed utilisation-based CL was then computed as suggested by Equation (56):

$$\mathrm{CL}_n^{\mathrm{U}} = w_P \cdot U_{P,n} + w_C \cdot U_{C,n} \tag{56}$$

$$w_P = 1, w_P = 2$$

$$U_{P,n} \in [0,1], U_{C,n} \in [0,1]$$

To translate EEG-derived information-processing intensity into the time-resolved utilisation estimate, $\mathrm{CL}_n^{\mathrm{U}}$ was multiplied by a window-specific information-processing weight ω_n (Equation (57)):

$$\mathrm{CL}_n^{\mathrm{U}\omega} = \omega_n \cdot \left(w_P \cdot U_{P,n} + w_c \cdot U_{c,n}\right) \tag{57}$$

ω_n is information − processing weight of a time window n

$$w_P = 1, w_P = 2$$

$$U_{P,n} \in [0,1], U_{C,n} \in [0,1]$$

The weights were derived in Section 6.4 at the level of unit task phases and intermediate-level subtasks. To apply them at the window level, each window *n* was assigned a weight ω_n based on the segment(s) that occurred within that window. Specifically, ω_n was computed as a time-weighted combination of segment weights (Equation (58)):

$$\omega_n = \sum_{s \epsilon S} p_{s,n} \cdot \omega_s \tag{58}$$

$$\sum_s p_{s,n} = 1$$

$p_{s,n}$ is the proportion of window duration occupied by task segment s

For segments lacking an empirically estimated ω_s due to rare occurrence, the Med of available weights (Section 6.4, Table 74) was used as a substitute.

Because absolute scales differ across participants and indicators, each participant's three windowed time series ($\mathrm{CL}_n^{\mathrm{EEG}}, \mathrm{CL}_n^{\mathrm{U}}, \mathrm{CL}_n^{\mathrm{U\omega}}$) was transformed to robust z-scores (Med/MAD standardisation). If minor timing gaps occurred between series, values were aligned on the common utilisation window grid and any missing samples within the shared time span were filled using linear interpolation. The model of CL was visualised for each participant individually as a timeline representing time on x-axis and dynamic changes in CL on y-axis. Figure 69 presents such a model for P#4 who performed the best in the LC task (overall performance score = 0.57), Figure 70 for P#16 who represents an average performer (overall performance score = 0.30, which equals the median value, as shown in Table 65), Figure 71 for P#8 who performed the worst (overall performance score = 0.18). The models for the rest of the participants are available in Appendix B.

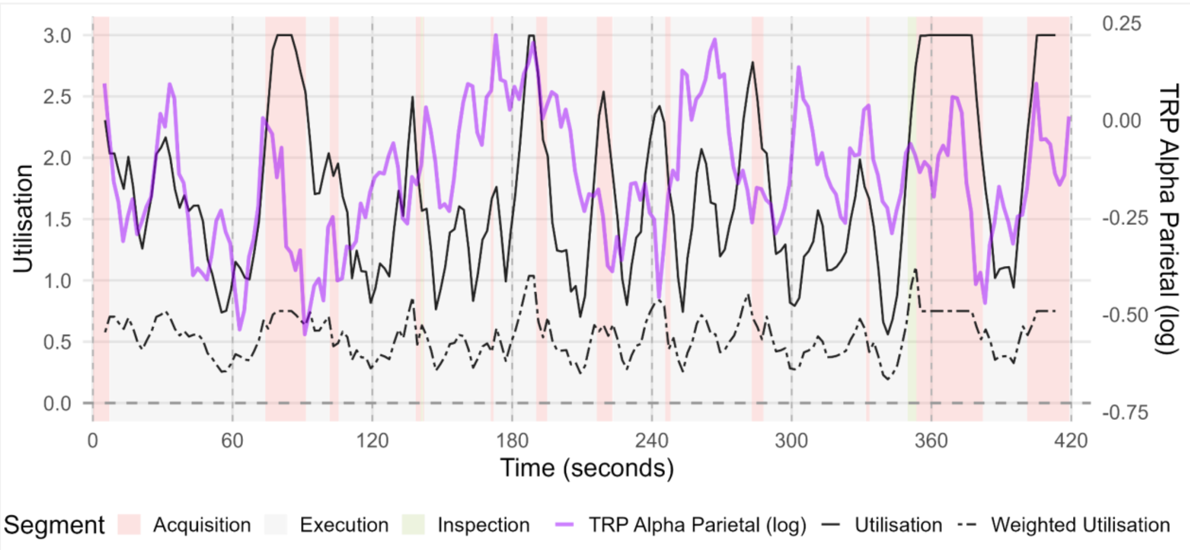

Figure 69 P#4: Dynamic changes in CL across moving windows (10s, 2s) – High Performer

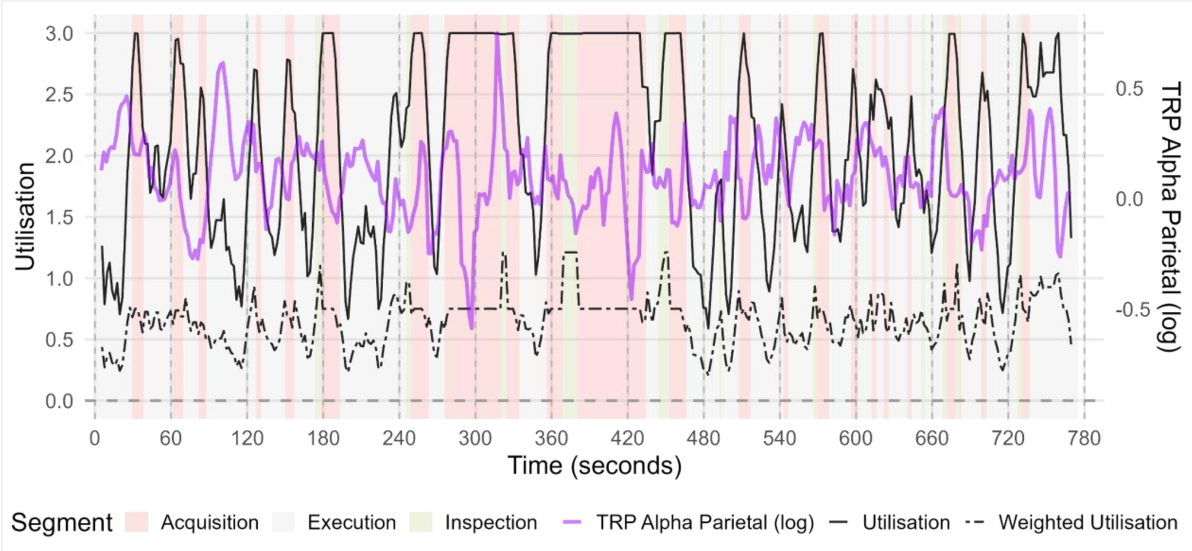

Figure 70 P#16: Dynamic changes in CL across moving windows (10s, 2s) – Average Performer

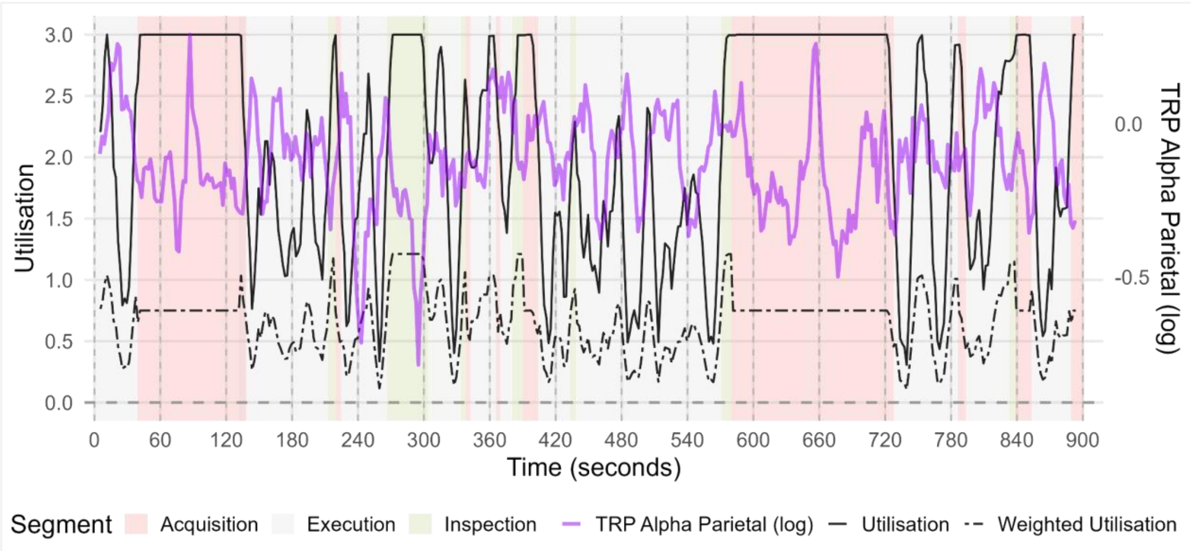

Figure 71 P#8: Dynamic changes in CL across moving windows (10s, 2s) – Low Performer

Concordance between indicators was then quantified at the window level using the following metrics (computed per participant):

1) Monotonic association was calculated to test whether the time-series were ordered similarly at the same time. It was calculated using Spearman ρ and Kendall τ between the robust z-series at synchronous time points. The results are interpreted in a way that values of ρ and τ near 0 indicate weak instantaneous association, while values near 1 indicate strong instantaneous association. Positive values mean the indicators co-vary in the same direction.
2) Temporal coupling and lag were calculated to test whether one indicator precedes the other(s). It was calculated using cross-correlation function (CCF) over lags within ± 10 seconds, and quantified with the peak correlation and the lag at the peak. The positive

peak correlation indicates coupling, while the negative lag at the peak indicates that parietal alpha TRP leads utilisation time-series.

3) Direction of change was calculated to test if the indicators rise/fall together over time. It was calculated by comparing the sign of first differences for each indicator. The agreement rate was used to quantify the fraction of time-windows with identical direction labels; higher values mean more consistent co-directional behaviour of indicators.
4) Overlap of windows indicating high CL was calculated to test if the indicators identify the same high-CL episodes. The overlap was calculated by labelling top 80% of windows on each time-series and computed the F1 score between pairs of indicators. Higher F1 score meant greater overlap of high-CL windows.
5) Shape similarity of time-series were compared by calculating dynamic time warping (DTW) distance on the standardised (robust z-score) indicators. Smaller distance indicates greater shape similarity between the indicators.
6) Segment-wise rank concordance was calculated so see if the indicators ranked segments (subtasks) similarly. This was done by calculating the Med of each time series within each subtask and computing Kendall's W across the three indicators. Higer W indicated more stable, shared ordering of subtask regarding implied CL across the indicators.

All the metrics were first computed per participant (available in Appendix C), and then summarised across participants using Med and interquartile range (IQR).

Finally, to test whether weighted utilisations improve alignment with EEG (parietal alpha TRP) relative to (raw, non-weighted) utilisations, within-participants differences Δ were calculated for each metric (Equation (59)):

$$\Delta = \text{metric}\ (\text{TRP} \leftrightarrow U_{\omega}) - \text{metric}\ (\text{TRP} \leftrightarrow \text{U}) \tag{59}$$

After that, a one-sided paired Wilcoxon signed-rank test was applied to the calculated group differences Δ. These results are interpreted in two ways:

1) For metrics where larger is better (Spearman ρ, Kendall τ, CCF peak, agreement rate, F1) tested hypothesis is that $\Delta > 0$;
2) For metrics where smaller is better (lag magnitude and DTW distance) tested hypothesis is that $\Delta < 0$.

The Hodges-Lehman (HL) estimate and 95% CIs were reported as robust effect-size measures of the Med differences between the groups Δ.

6.7.2 Results

Comparison of alignment between dynamic changes in CL when indicated by parietal alpha TRP, utilisations, and weighted utilisations is summarised in Table 81. The results showed that both utilisation-based indicators well align, but they were weakly aligned with parietal alpha TRP.

For pairs of parietal alpha TRP and utilisation-based indicators, median instantaneous monotonic associations were close to zero. CCF peak correlations were low, indicating limited temporal coupling. Median lags at the CCF peak were small in magnitude, with wide IQRs, suggesting variability in whether and by how much one indicator led the other. The results are inconclusive as in some cases parietal alpha TRP led behaviour (indicated by a negative median lag), while in other cases utilisation indicators led EEG-based signal.

Table 81 Summary of alignment metrics by indicator pairs (Med; IQR)

Metric	TRP vs U	TRP vs weighted U	U vs weighted U
Spearman ρ	$2.05 \cdot 10^{-2}$; 0.25	$2.48 \cdot 10^{-2}$; 0.16	0.92; $9.00 \cdot 10^{-2}$
Kendall τ	$1.42 \cdot 10^{-2}$; 0.17	$1.70 \cdot 10^{-2}$; 0.11	0.82; 0.11
CCF peak	0.13; 0.11	0.11; $6.77 \cdot 10^{-2}$	-
Lag step	1.5; 7.00	6.75; 7.00	-
Lag second	3.00; 14:00	7.00; 13.50	-
Agreement rate	0.34; $4.30 \cdot 10^{-2}$	0.34; $3.65 \cdot 10^{-2}$	0.82; $7.00 \cdot 10^{-2}$
F1	0.22; $8.88 \cdot 10^{-2}$	0.18; $5.40 \cdot 10^{-2}$	0.40; 0.37
DTW distance	139.47; 20.23	150.25; 22.87	91.39; 73.65

Agreement on the direction of change was modelled for both pairs of indicators; the overlap of time windows with indicated high CL was low. Shape similarity metrics were also unfavourable; median DTW distances were large for both pairs of indicators, suggesting differences in time series of EEG-based and utilisation-based indicators. Therefore, these values show weak alignment between parietal alpha TRP and both utilisation indicators, with no obvious advantage of the weighted utilisation.

Across participants, the Med differences for all metrics were small and centred near zero, as detailed in Table 82. Agreement in direction of change and F1 overlap of high-CL windows also showed slightly negative HL estimates, while lag magnitude and DTW distance did not show systematic improvement in the alignment when using Weighted U. Wilcoxon test provided no evidence that Weighted U outperformed U on any metric and only approximately 30% of participants showed changes in the hypothesised direction. Thus, at the group level, weighting the utilisations did not improve the temporal agreement with parietal alpha TRP compared to the simpler unweighted utilisation indicator.

Table 82 Comparison of alignment between utilisation-based indicators and parietal alpha TRP

Metric	Med	IQR	HL	CI low	CI high	p	Improvement
Spearman ρ	$1.65 \cdot 10^{-2}$	$5.76 \cdot 10^{-2}$	$-2.16 \cdot 10^{-2}$	$-4.18 \cdot 10^{-2}$	-	0.97	0.32
Kendall τ	$9.55 \cdot 10^{-3}$	$3.76 \cdot 10^{-2}$	$-1.46 \cdot 10^{-2}$	$-2.90 \cdot 10^{-2}$	-	0.98	0.27
CCF peak	$1.85 \cdot 10^{-2}$	$4.66 \cdot 10^{-2}$	$-2.55 \cdot 10^{-2}$	$-4.75 \cdot 10^{-2}$	-	0.98	0.23
Lag second	0.00	3.00	$4.16 \cdot 10^{-5}$	-	2.00	0.58	0.27
Agreement rate	$9.50 \cdot 10^{-3}$	$2.70 \cdot 10^{-2}$	$-7.48 \cdot 10^{-3}$	$-1.80 \cdot 10^{-2}$	-	0.92	0.36
F1	$5.15 \cdot 10^{-2}$	0.12	$-4.70 \cdot 10^{-2}$	$8.88 \cdot 10^{-2}$	-	0.98	0.30
DTW distance	8.90	15.21	12.16	-	20.62	0.10	0.23

6.7.3 Discussion

Dynamic changes in CL were modelled per participant using time series (moving windows of 10s with 2s step) of three CL indicators: parietal alpha TRP (EEG), Med operator utilisations (U), and weighted utilisations (Weighted U).

As intended by the formulation in Equations (56)-(58), U and Weighted U are computed on the same window grid and are therefore temporally aligned by construction; the weighting step applies a window-specific multiplier to the utilisation estimate rather than redefining the timing of the estimate. Accordingly, the strong agreement between U and Weighted U in instantaneous rank ordering and direction of change mainly indicates that weighting acts primarily as a rescaling of the utilisation signal, while preserving its temporal structure. The informative question is therefore not whether weighting changes when peaks and troughs occur, but whether it changes their relative magnitude in a way that makes utilisation-derived dynamics more consistent with the EEG-based dynamics.

In contrast, both utilisation-based indicators aligned weakly with parietal alpha TRP. For TRP–U and TRP–Weighted U pairs, instantaneous monotonic associations were near zero on average and CCF peak correlations were low, suggesting limited temporal coupling between EEG-derived dynamics and utilisation-derived dynamics under the current windowing and segment-to-weight assignment. Lag analyses further indicated that the relationship is not stable across participants; median lags at the CCF peak were small, but the dispersion was wide, implying substantial between-participant variability in whether changes in TRP tend to precede or follow changes in utilisation. This variability makes it difficult to interpret a single typical temporal offset and suggests that any coupling between neural dynamics and behavioural/operator dynamics is either weak, participant-specific, or dependent on the segment/subtask context.

Agreement-based metrics reinforced this implications. The agreement rate on the direction of change was modest, and the F1 overlap of high-CL windows was low for both TRP–U and

TRP–Weighted U pairs, indicating that the indicators often did not flag the same time windows as high CL. Large DTW (shape similarity metric) likewise suggest that the overall trajectories of parietal alpha TRP and utilisation-based indicators differ in their patterns beyond simple lag shifts or scaling.

Crucially, applying EEG-derived weights did not yield a consistent improvement in alignment with parietal alpha TRP. Across participants, differences between TRP–Weighted U and TRP–U were small and centred near zero, and statistical comparisons provided no robust evidence that Weighted U outperformed U on any metric; only a minority of participants showed improvements in the hypothesised direction (for roughly 23–36%, depending on the metric). Taken together, these findings suggest that dynamic CL inferred from utilisation captures a behavioural/operator-based pattern that is internally consistent, whereas parietal alpha TRP reflects neural dynamics that only weakly track those utilisation-derived changes under the current segmentation and windowing schemes. This supports the interpretation that EEG- and cognitive-modelling-based indicators may be sensitive to partially different components of CL, and that improving alignment may require segment-specific modelling, richer EEG feature sets, or alternative definitions of high-CL episodes that better match the timescales and mechanisms reflected in parietal alpha TRP.

6.8 Relating CAD performance assessments and CL

This part of the analysis investigates relationships between assessed CAD performance of engineering designers (based on completeness of final 3D CAD models and CAD modelling duration) and CL they experienced in CAD tasks, calculated with utilisation-based indicators. By relating these two constructs, this section addresses the research hypothesis stated in Section 1: *method for measuring and analysing CL, based on the monitoring of psychophysiological responses, enables the assessment of the engineering designer's level of CAD activity performance*.

6.8.1 Data analysis

Theil–Sen regression models were developed to relate CL_{Task} values calculated using Weighted Median U and overall performance considering completeness of the resulting CAD model and CAD modelling duration (see (Equation (46)). Theil–Sen regression was selected as a nonparametric, outlier-resistant alternative to ordinary least squares, providing a robust estimate of the monotonic linear trend between utilisation-based indicators and subjective CL without relying on normality assumptions. In line with robust reporting practice, the slope (b) is reported as the Med of all pairwise slopes, and the intercept (a) is reported as a robust location estimate (based on Med, positioning the line through the central data density). Model fit is summarised using a robust error metrics, reported here as the median absolute error (MAE). Uncertainty in the estimated trend is communicated via a bootstrapped 95% confidence interval (CI) for the slope b.

6.8.2 Results

Robust Theil–Sen regressions showed a negative association between overall CAD activity performance and all utilisation-based estimates of CL (see Table 83); higher utilisation-based CL corresponded to lower overall performance (negative slopes across models).

Table 83 Theil-Sen regression models: Utilisation-based indicators and Overall performance

Utilisation-based indicator	Med intercept a	Med slope b	95% CI for b	p-value	MAE
Weighted Median U	0.56	-1.12	-2.14, -0.64	$8.25 \cdot 10^{-5}$	$8.98 \cdot 10^{-2}$
Weighted Global U	0.48	-0.89	-2.23, $8.58 \cdot 10^{-2}$	$6.90 \cdot 10^{-2}$	$9.84 \cdot 10^{-2}$
Median U	0.58	-0.15	-0.24, $8.31 \cdot 10^{-2}$	$1.77 \cdot 10^{-4}$	0.11
Global U	0.63	-0.25	-0.33, $8.97 \cdot 10^{-3}$	$3.21 \cdot 10^{-2}$	$8.70 \cdot 10^{-2}$

Among the tested utilisation-based indicators, Weighted Median U exhibited the strongest relationship with performance, with a steep negative slope (b = −1.12) and a clearly non-zero confidence interval (Figure 72). The non-weighted Median U indicator also showed a significant negative association, although with a substantially smaller slope (b = −0.154). Using global utilisations, Global U remained significantly negatively related to the overall performance. In contrast, Weighted Global U did not reach statistical significance, indicating that the weighted global formulation did not consistently predict performance in this dataset.

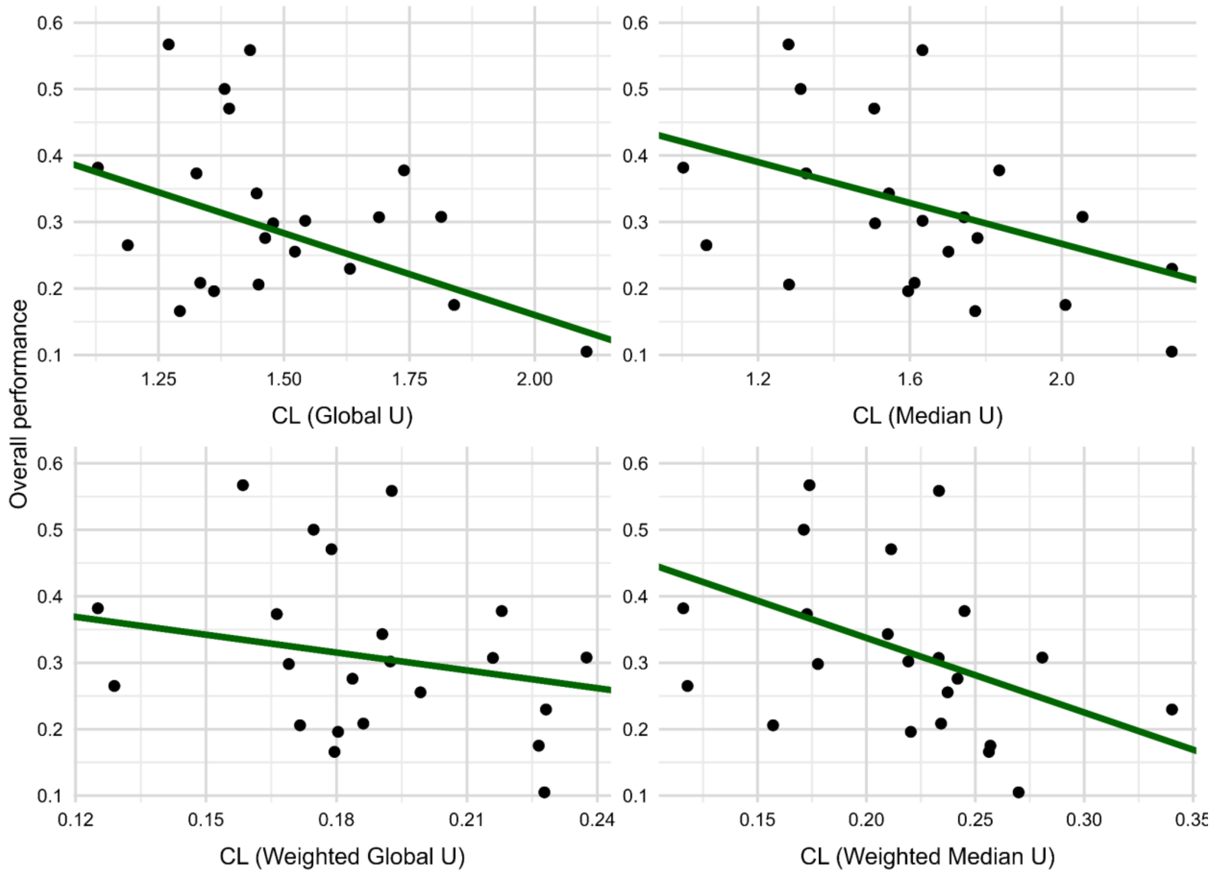

Figure 72 Theil-Sen regression models: Utilisation-based CL and Overall performance

Predictive performance assessed via LOOCV MAE suggested modest differences between indicators. The smallest error was obtained for Global U, followed closely by Weighted Median U. Weighted Global U had a slightly larger error, while Median U showed the largest error among the utilisation-based models (MAE = 0.11). Overall, these results indicate that utilisation-based CL measures, particularly Weighted Median U and Global U, capture meaningful variance in overall performance, with higher inferred CL reliably associated with poorer performance.

6.8.3 Discussion

This analysis examined whether utilisation-based indicators of CL, derived from the proposed CL equation and operator utilisations, relate to overall CAD activity performance. By linking estimated CL to the overall CAD performance, the results provide evidence supporting the research hypothesis that a method for measuring and analysing CL, grounded in monitored psychophysiological responses, enables assessment of engineering designers' CAD performance.

Across all tested utilisation-based indicators, the estimated Theil–Sen slopes were negative, meaning that higher utilisation-based CL was consistently associated with lower overall performance. This pattern aligns with the theoretical expectation that greater cognitive demands during CAD modelling constrain efficient and accurate task execution, ultimately reducing performance when the task requires both sustained attention and precise action sequencing. These findings align with prior research linking CL (or WL) to degraded performance in engineering design tasks (e.g. [135], [139]).

Importantly, the strength and robustness of the association depended on how utilisation was operationalised. Weighted Median U showed the clearest relationship with performance, yielding the steepest negative slope and a confidence interval that excluded zero. This indicates that when operator utilisation is combined with empirically derived information-processing weights, the resulting CL estimate better captures the performance-relevant intensity of CAD activity than utilisation alone. In other words, weighting appears to add explanatory value when it is applied to participant-specific (median) utilisations, likely because it preserves individual differences in how designers distribute perceptual and cognitive resources across segments while simultaneously amplifying segments assumed to be more information-intensive.

The non-weighted Median U and Global U indicators were also significantly related to performance, but with smaller slopes. These findings suggest that even a simpler utilisation-based estimate reflects meaningful variance in engineering designers' performance, supporting the broader premise that operator-based modelling can be used as a proxy for CL-related performance constraints. At the same time, the comparatively weaker slopes for unweighted metrics indicate that utilisation alone may partially blur the distinction between time spent being "busy" with how information-demanding that activity is.

The Weighted Global U indicator did not reach statistical significance. This result is informative because it suggests that weighting does not automatically improve performance prediction when individual variability in utilisation is removed. Since global utilisations are constant across participants by definition, the weighted-global formulation can only differentiate participants through the weighting scheme and segmentation structure, which may be insufficient to explain inter-individual performance differences. Practically, this means that global utilisations are useful for simplified modelling and visualisation, but may be less suited for performance prediction than participant-specific utilisation estimates.

When considering out-of-sample predictive accuracy, LOOCV errors differed only modestly between models; Global U achieved the smallest median absolute error, followed closely by Weighted Median U, whereas Median U produced the largest error. Taken together with the slope results, this pattern implies a trade-off; Weighted Median U appears to capture a stronger performance-related trend (steeper slope and clearer significance), while Global U offers slightly better predictive stability in cross-validation despite a weaker slope. This is consistent with the fact that global measures tend to be less variable and therefore can generalise better in small samples, whereas participant-specific measures capture more variance but can also reflect participant-specific behaviour.

Overall, these results support the central claim that utilisation-based CL modelling can inform assessment of CAD performance; engineering designers who exhibit higher estimated CL tend to show lower overall performance. At the same time, the comparison among indicators clarifies when the proposed weighting scheme is most beneficial; namely, when applied to participant-specific utilisations, while also highlighting the limits of simplified global formulations for predicting individual performance. However, accuracy, generalisability, and sensitivity of the identified relationship still require confirmation on larger samples and additional CAD modelling tasks.

6.9 Experimental study limitations

As outlined in the chapter introduction, the second empirical study was designed to test whether CAD task complexity can be varied in a controlled way and whether the proposed method can quantify corresponding changes in CL using both EEG- and utilisation-based indicators. The aim was therefore to provide CAD-specific evidence for the method's constructs and their relationships under realistic modelling conditions, rather than to claim exhaustive coverage of all determinants of CL in CAD. With that scope in mind, several limitations should be considered when interpreting the findings.

A first limitation concerns how task complexity was defined and operationalised. The two experimental tasks were intentionally designed to differ in complexity in order to impose different levels of CL. Complexity was quantified using proxies for overall geometric and CAD model complexity (e.g. number of surfaces, dimensions, and CAD features), consistent with common approaches in the CAD complexity literature (e.g. [109]). This operationalisation supports transparent task design and comparability with prior work, but it does not capture all possible sources or aspects of complexity. Consequently, the conclusions are primarily limited to CL differences driven by overall geometric and CAD model complexity as operationalised here.

A second limitation concerns the EEG instrumentation and the interpretability of EEG-based results. EEG was recorded using a 14-channel EMOTIV EPOC+ system, selected as a pragmatic trade-off between usability and measurement quality, and consistent with prior engineering design neurocognition studies (e.g. [150], [180]). The device provides adequate temporal resolution for analysing oscillatory dynamics during CAD modelling and supports computation of commonly used EEG features. However, its spatial resolution is limited, and the analyses do not support fine-grained functional localisation or strong anatomical inference. In addition, EEG signals are susceptible to artefacts and methodological variability. Replication using higher-density EEG systems and/or complementary psychophysiological measures (e.g. eye tracking [235] or fNIRS [139]) would strengthen confidence in the robustness and generalisability of the observed EEG-related patterns.

A third limitation concerns the participant sample. Expanding the sample beyond 24 participants would increase statistical power and may strengthen the stability of observed associations between indicators. In addition, the participant pool was gender-imbalanced (2 female, 22 male). While this distribution reflects the representative distribution of the profile of engineering designers, it limits the ability to examine potential gender-related variability and warrants more balanced recruitment in future studies, particularly given the sensitivity of EEG measures to inter-individual differences.

A fourth limitation concerns the behavioural coding required to build the CAD performance models and compute operator utilisations. The segmentation and operator mapping were conducted by a single researcher (the author) following a predefined scheme grounded in the proposed hierarchical task decomposition and operator definitions. Nevertheless, manual annotation inevitably involves judgement calls—especially when segment boundaries are ambiguous, when subtasks overlap, or when the same observable behaviour could plausibly reflect different perceptual/cognitive operations. This introduces a risk of coder-specific bias and prevents quantification of interrater reliability for the present dataset. Future work should therefore include multiple independent coders supported by training, a shared codebook with decision rules and examples, and iterative calibration rounds, with interrater reliability quantified on a representative subset of the data. Establishing such reliability would strengthen confidence that the derived utilisation measures and the resulting CAD performance models are robust to coder subjectivity.

6.10 Conclusions

The second empirical study investigated CL in CAD modelling using the full workflow of the proposed method, combining subjective CL measurement (NASA TLX), psychophysiological measurement (EEG), and process-based cognitive modelling (operator utilisation). Motivated by 1) limited CAD-specific evidence on valid EEG-based CL indicators and 2) the need for a hierarchical CAD process representation that supports interpretable CL localisation, the study pursued two overarching aims. First, to identify EEG features that are sensitive to complexity-driven CL variation in CAD tasks. Second, to operationalise and evaluate the CAD performance and CL models as a basis for hierarchical, process-based CL analysis, including relationships with subjective CL and CAD performance. These aims were addressed through nine research questions (RQ 2.1–RQ 2.9), spanning EEG-based CL indicator identification, process modelling and segmentation, EEG-informed weighting, segment-level CL differences and contributions, dynamic CL alignment, and relating CL indicators and overall CAD performance assessment.

A key contribution of the study is methodological: it is the first to examine CL in CAD by explicitly triangulating 1) NASA TLX, 2) EEG, and 3) indicators derived applying human performance (cognitive) modelling within one coherent framework. Importantly, the subjective NASA TLX analysis provided validation that the designed task manipulation (LC vs. HC CAD task) indeed affected perceived WL dimensions, a step that is essential in experiments including complex CAD tasks where many factors can confound CL. This validation strengthens the interpretation that observed EEG differences are linked to CL rather than to unrelated task characteristics, and it addresses a common limitation in psychophysiological CL studies in CAD where subjective confirmation is often missing.

To answer RQ 2.1, the study tested nine EEG features identified from the literature as promising candidates sensitive to changes in CL. Six of them showed significant differences between HC and LC CAD tasks with moderate to large effect sizes, demonstrating that EEG can detect complexity-driven CL differences in CAD modelling. The features sensitive to CAD task complexity included alpha TRP (global and cortex-region-specific), theta/alpha TRP ratio, and beta TRP. Among them, alpha-band features were consistently the most sensitive, and the convergence of evidence based on effect sizes with subjective WL correlations and Theil-Sen regression models supported parietal and occipital alpha TRP as the most prominent candidates for describing CL variations in the CAD context. This result aligns with earlier findings that posterior alpha (including parietal and occipital cortical areas) is sensitive to CL, but it also reveals that the direction of alpha changes in CAD can differ from patterns often reported in cognitively controlled tasks. Instead of decreases, an increase in alpha TRP with higher CL was observed, consistent with earlier CAD-related evidence and with studies on visuospatial processing. This indicates that EEG-based CL indicators cannot be transferred to CAD contexts without validation and that CAD tasks may involve cognitive mechanisms (particularly visuospatial processing) through which alpha and beta bands behave differently than in heavily controlled tasks and experiments typical for cognitive psychology. This finding motivates further work to generalise these EEG features across other CAD activities (e.g. assembly modelling, simulations, design reviews) and to clarify the cognitive processes that drive the CAD-specific behaviour of frequency bands.

The second aim of the study was to operationalise and validate the proposed CAD performance model and use it for CL analysis across hierarchical levels. RQ 2.2 examined whether unit task phases and intermediate-level subtasks differ in information-processing intensity when measured by the selected EEG-based CL indicator (parietal alpha TRP). To enable segment-wise TRP aggregation and information-processing weight assignment, the CAD performance modelling procedure (Section 5.2) was applied to the LC task for all 22 participants, producing the hierarchical segmentation and operator mapping required by the method. Using these segment boundaries (start and end time), parietal alpha TRP was compared across unit task phases and intermediate-level subtasks. The results did not provide evidence for systematic intensity differences across these segments at the group level. This negative result is informative; it suggests that parietal alpha TRP may be sensitive to overall task-level CL differences (LC vs. HC), while segment-level intensity differences may be more challenging to detect due to within-task variability, participant-specific response directions, and the complexity of CAD behaviour. Indeed, parietal alpha TRP showed both increases and decreases across segments for different participants, implying heterogeneous neural responses during CAD modelling. This heterogeneity raises the possibility that both directions reflect increased processing relative to baseline but through different neural mechanisms or strategies, which would influence any attempt to derive stable weights directly from parietal alpha TRP directionality. Within the available dataset, the analysis did not support reinterpretation of both increase and decrease as high CL in a way that would stabilise segment ordering.

Building on the hierarchical decomposition provided by the CAD performance model, the study proposed EEG-informed information-processing weights and evaluated their usefulness through comparisons with subjective CL (NAS TLX). To answer RQ 2.3 and RQ 2.4, utilisation-based CL indicators were compared in weighted vs. non-weighted forms and in global vs. participant-specific median aggregation forms. At the task level (LC condition), all four utilisation-based indicators (Median U, Global U, and their weighted variants) showed significant positive associations with NASA TLX mental demand using robust Theil–Sen regression. This supports the core assumption that utilisation-based modelling captures psychologically meaningful variance in perceived CL. With respect to weighting, the results indicate that EEG-derived weighting can yield small incremental reductions in prediction error, but the gain is modest, which is consistent with the limited stability of segment-level information-processing weights. With respect to simplification, the global-versus-median comparison provides a clearer practical implication; global utilisations can approximate participant-specific summaries well in this dataset, and in some cases perform better, supporting the method's usability objective when participant-specific utilisation estimation is impractical. Overall, utilisation-based CL appears robust under simplified global aggregation, while weighting should be interpreted as a supportive refinement in the current dataset.

Moving from task-level summaries to a process-level perspective (to answer RQ 2.5-RQ 2.7), utilisation-based indicators revealed that CL is not uniform across CAD modelling; unit task phases and subtasks differed significantly in implied CL, with many pairwise contrasts showing large effects. Contribution analyses further showed that overall CL depends on both how demanding a task segment is when it occurs and how much time it occupies. Across utilisation indicators, a stable hierarchy emerged in intermediate-level subtask contributions, with dimensioning and navigating contributing most, drawing and manipulating occupying a mid-band, and referencing/constraining/deleting contributing least. Importantly, weighting generally amplified contrasts (steeper gradients) but did not change the rank ordering within utilisation-based indicators. In contrast, parietal alpha TRP did not reproduce the utilisation-based ordering, reinforcing that EEG- and utilisation-based measures reflect overlapping but non-identical aspects of CL in CAD modelling.

To address RQ 2.8, dynamic CL time series were modelled per participant using moving windows (10 s, step 2 s) from parietal alpha TRP, utilisation-based CL, and weighted utilisation-based CL. As intended by the formulation, U and weighted U are temporally aligned by construction; weighting modifies magnitude via a window-specific multiplier rather than changing when peaks and troughs occur. The key empirical result is that both utilisation-based dynamics aligned well with each other, but weakly with parietal alpha TRP; instantaneous associations were near zero on average, temporal coupling was low, lag direction varied across participants, overlap of high-CL windows was limited, and shape similarity was weak. Weighting did not improve alignment with TRP at the group level. These findings suggest that, under the current segmentation-to-weight mapping and windowing scheme, utilisation captures a stable behavioural, operator-based dynamic pattern, while parietal alpha TRP reflects neural dynamics that only weakly track those behavioural fluctuations.

Finally, utilisation-based CL estimates were meaningfully related to overall CAD performance assessment (RQ 2.9). In particular, higher utilisation-based CL was associated with lower performance, supporting the thesis hypothesis that CL indicators derived from the method can inform performance assessment. The strength of this relationship depended on the indicator variant: Weighted Median U showed the clearest negative association (steepest slope), while Global U also showed a significant relationship and achieved strong predictive stability in LOOCV. Weighted Global U did not reach significance, suggesting that weighting does not automatically improve performance prediction when individual variability in utilisation is removed. Overall, the results indicate that utilisation-based modelling provides a practical behavioural proxy for performance-relevant CL, and that EEG-based weighting adds explanatory value primarily when applied to participant-specific utilisation estimates.

In summary, the second empirical study advances CL research in CAD by: 1) validating that EEG can detect complexity-driven CL changes in CAD and identifying posterior alpha TRP (parietal/occipital) as the strongest candidate family while revealing CAD-specific band behaviour; 2) operationalising the CAD performance model (22 participants, LC task) to validate reusability of hierarchical task/goal structure, process analysis layers, and operator utilisations; 3) demonstrating that utilisation-based CL indicators relate robustly to subjective mental demand (NASA TLX) and support process-level discrimination across unit task phases and intermediate-level subtasks; 4) showing that EEG-derived information-processing weighting is feasible to integrate and can offer modest gains for subjective CL prediction, but is limited in the present dataset by unstable segment-level TRP differences; 5) detecting which intermediate-level subtasks contribute most to overall CL once time exposure is considered, with stable rankings within utilisation-based indicators; 6) demonstrating weak dynamic alignment between EEG- and utilisation-derived CL under the current time-resolved modelling scheme; and 7) providing evidence that utilisation-based CL, especially in participant-specific form and, in some cases, combined with EEG-informed weighting, is significantly associated with overall output-based CAD performance assessment. Collectively, the findings support the thesis claim that the proposed multimodal CL measurement method in CAD is feasible and informative, while also clarifying where the current approach is strongest (weighted utilisation-based modelling and segment-level interpretation) and where further methodological refinement is most needed (segment-level EEG stability, richer EEG feature combinations, and improved dynamic coupling models).

This page intentionally left blank

7 VALIDATION AND DISCUSSION

This chapter presents the validation and discussion of the research outcomes: 1) the theoretical model of engineering designer's performance in CAD activities (Section 7.1.1), 2) the theoretical model of CL in CAD activities (Section 7.1.2), and 3) the method for measuring and analysing CL in CAD activities (Section 7.2). Building on the validation criteria defined in Section 1.3, formulated as model and method objectives and their main purpose (Table 2 and Table 3), the chapter evaluates the proposed theoretical models and the prescribed method against these criteria through theoretical and empirical structural and performance validity, as prescribed by the Validation Square framework [21], [22]. *It also verifies the research hypothesis and discusses the extent to which research objectives were reached. Finally, the chapter acknowledges key limitations.*

7.1 Validating the theoretical models

Validation of the theoretical models is conducted primarily in terms of theoretical structural validity in the sense of the Validation Square [21], [22], i.e. whether the model constructs, their relationships, and the overall structure form a coherent representation that is appropriate for the intended purpose. To make this evaluation explicit, the validation is organised around the model objectives defined in Section 1.3 (Table 2); mechanisms, usefulness, robustness/generality, and simplicity, which operationalise what "appropriate and coherent" means for the models in this thesis. In addition, where evidence from the empirical studies is relevant (e.g. whether the proposed hierarchy can be instantiated from behavioural data, or whether the CL formulation supports empirical analyses), the model validation also draws on the empirical chapters to demonstrate operational/empirical usability of the theoretical model as implemented on real datasets.

7.1.1 Validating the theoretical model of CAD performance

This section validates the theoretical model of engineering designers' CAD performance as an HCI activity. The purpose of the theoretical models in this thesis is to formalise CAD performance (and CL) by identifying its key elements, their characteristics, and the relationships among them (Section 1.3). Chapter 2 showed that CAD performance is widely discussed in engineering design research, yet comparison across studies remains difficult because definitions, segmentation schemes, and metrics are inconsistent and often constrained by data availability (Section 2.2). In particular, outcome measures of the final CAD model are essential for assessing effectiveness, but they provide limited insight into how performance unfolds over time; process measures are frequently cumulative summaries, while temporally structured descriptions of CAD modelling behaviour are still relatively rare and use inconsistent terminology (commands, events, actions, operations). These conclusions motivate the present domain-specific theoretical model as a synthesis that provides a transferable structure for describing CAD modelling as a process and makes the information-processing perspective explicit so that performance can later be related to CL at the task segment level.

Theoretical structural validation of the model of engineering designers' CAD performance concerns whether its levels, layers, and relations provide a theoretically plausible and empirically usable representation of CAD modelling as an HCI activity. Regarding mechanisms (Table 2), the model is grounded in established human performance modelling theory (Section 2.3) by representing embodied cognition through P, C, and M systems and their corresponding operators, including the possibility of parallel operator activity when different processors are engaged and no output–input dependency exists. In this way, observable CAD events are not treated only as behavioural counts, but can be interpreted in terms of underlying perceptual, cognitive, and motor resource engagement. This focus complements the dominant CAD performance literature (Section 2.2), which has largely operationalised performance through outputs (e.g. available 3D geometry and information from the feature tree) and associated quality rubrics (e.g. completeness, consistency, simplicity, reusability, editability) rather than through explicit mechanisms explaining how those outputs are produced [44], [53], [60], [65], [71]. At the same time, it remains consistent with prior applications of human performance modelling in CAD and related HCI tasks (Section 2.3.3), where GOMS modelling (typically in its simplest, KLM form) has been used to represent user procedures, predict execution time, and guide interface redesign [7], [92]. Unlike these studies, which often examined short, well-defined tasks with predefined execution methods, the present domain-specific theoretical model targets extended, open-ended CAD modelling (in terms of procedure and strategy) and therefore prioritises a descriptive, segmentable process representation that can be aligned with cognitive mechanisms and instantiated from behavioural data.

Regarding usefulness (Table 2), the theoretical model is intended to support investigation of CAD modelling in realistic settings by enabling performance to be examined not only as an outcome, but also as a temporally unfolding process. This is aligned with process-oriented CAD analytics that derive process metrics from CAD logs and timelines (e.g. command/event distributions, transitions, rework/revision patterns, and temporal visualisations of "what engineering designers do and when") to interpret strategies and process evolution (e.g. [12], [30], [43], [58], [59]). The present model contributes to this research direction by providing an explicit and reusable structure that can organise such process descriptors and relate them to mechanisms through operator mapping, thereby improving interpretability and supporting later linkage to CL modelling.

A central conceptual decision enabling both usefulness and empirical usability is the hierarchical decomposition of CAD tasks into three task/goal levels (high-, intermediate-, and low-level subtasks/subgoals), paired with corresponding process layers (CAD actions, CAD operations, and unit task operations) and model information layers. This hierarchy was derived from literature (Section 2.2) and motivated by the need to manage WM limits in cognitively complex CAD modelling tasks [7], [92]. In the CAD performance literature, segmentation is widely used but implemented inconsistently; studies vary substantially in the level of analysis and employ overlapping terminology (actions/events/commands/operations), which limits repeatability and comparability across studies. Some studies segment at higher levels considering design activity in general (e.g. problem definition, planning, evaluation) [33], [50], whereas studies that use CAD logs typically focus on lower-level command/event categories and their distributions or transitions [12], [30], [43], [58], [59]. The proposed three-level structure aligns with prior multi-layer process representations that separate primitive events from higher-level operations and information object states (event/operation levels and product layers) [28], [29], while extending them by explicitly specifying a CAD modelling task/goal hierarchy that can be applied consistently across studies and used to structure empirical analysis. It also resonates with GOMS-based accounts of CAD performance showing that procedure and goal structure (procedural knowledge) can differ fundamentally between users, implying the need for a hierarchy that supports multiple valid goal structures rather than assuming a single "standard" procedure [97]. The proposed structure was supported empirically; in the first empirical study, observed CAD modelling behaviour indicated that engineering designers segment work into goal-driven units and intermediate subgoals (Section 3.4), and in the second empirical study the same decomposition logic was successfully applied to a different participant sample and different CAD tasks (Section 6.4), supporting applicability beyond a single experiment and two CAD modelling tasks.

Importantly, the hierarchy is operational because its elements can be identified from recorded CAD modelling process data. The low-level layer consists of CAD actions observable in screen recordings and behavioural traces (e.g. cursor movement, clicks, typing, rotations, zooming, dragging, system response). Intermediate-level subtasks are reached through clusters of CAD operations that can be identified as functionally and temporally connected groups of CAD actions pursuing a shared subgoal (e.g. navigating, referencing, drawing, constraining, dimensioning, manipulating a view, deleting). High-level subtasks align with unit task goals in feature-based CAD modelling, reached through sketching and applying/editing CAD features as two main operations. This supports process-based analysis at an intermediate abstraction level that is often most informative for distinguishing strategy differences (above commands but below overall task goals) and for making procedures comparable across engineering designers [98]. In the analysed datasets from both empirical studies, the taxonomy was sufficient to code all observed CAD interactions without requiring additional category types, supporting the theoretical model's domain-specific comprehensiveness. At the same time, the proposed theoretical model differs from much of descriptive CAD log analysis (Section 2.2.2.2.2) as it maps observed CAD events to P, C, and M operators, enabling analysis of mechanisms rather than only behavioural summaries.

Regarding robustness and generality (Table 2), the theoretical model is designed to support comparison without overfitting to a specific CAD modelling sequence or geometry. Although CAD modelling processes are highly variable and can lead to visually similar models through many alternative combinations of sketches and features [12], [38], [52], [61], intermediate-level subtasks (clusters of CAD operations) recur across many feature-based workflows and thus provide a stable level for comparison. This directly addresses a central challenge highlighted in the literature; because "best" CAD modelling procedures are rarely universal, process assessment becomes task- and system-dependent and metrics require context-specific interpretation [12], [40], [41]. By combining a stable intermediate layer with a detailed CAD action layer, the theoretical model, once operationalised with the method, supports comparison across participants and tasks while still retaining task- and user-dependent behavioural detail. Moreover, the decision to allow parallel operator activity is consistent with CAD-oriented GOMS work reporting that parts of P and C activity can overlap in well-practiced sequences and that modelling assumptions should reflect such overlap for realistic performance representation [96].

Finally, the structure gained additional empirical support through refinement of unit task phases. While traditional performance models emphasise acquisition and execution, the first empirical study indicated that inspection constitutes an additional important phase in CAD unit tasks. This finding resonates with prior process-oriented studies that emphasise evaluation, checking, and revision behaviours (e.g. creation/revision balance, rework, event transitions) [58], [63], but it formalises inspection as an explicit, recurring component of CAD unit-task structure rather than treating it as an implicit consequence of errors or iteration. Incorporating acquisition–execution–inspection therefore improves descriptive adequacy for CAD modelling, even if inspection is the least predictable in duration and occurrence.

Regarding simplicity (Table 2), the model intentionally uses a compact set of constructs (P/C/M and R operators), three task/goal levels with corresponding process and information layers, and a limited set of reusable CAD action/operation categories, rather than a full cognitive architecture specification. This preserves interpretability and feasibility for realistic CAD modelling sessions while retaining a mechanism-based interpretation compatible with later CL modelling.

Overall, the CAD performance model satisfies the intended objectives by representing relevant mechanisms (P, C, M), providing an operational task/goal hierarchy that can be instantiated from empirical data, and supporting systematic comparison of CAD modelling behaviour across engineering designers and CAD tasks. In that sense, the research fulfilled the first objective of the thesis stated in Section 1.1. By defining CAD performance as a sequence of unit tasks, unit task phases, CAD operations (and their clusters), and CAD actions, and by mapping P, C, and M operators, the domain-specific theoretical model enables multi-level analysis and establishes a common representation that can later be related to CL indicators and CAD performance outcomes.

Implications and directions for further research follow directly from these conceptual choices. First, the task hierarchy and segment taxonomy should be tested and extended to other CAD activities beyond CAD modelling of a single part (e.g. assembly modelling, technical documentation generation, simulation/analysis), while retaining the same abstraction principle. Second, reproducibility should be strengthened by formalising decision rules for borderline segmentation/mapping cases and reporting analyst agreement (interrater reliability) on a representative subset. Third, partial automation of segmentation and operator mapping from CAD logs should be explored to reduce analyst burden and support scaling to larger datasets.

7.1.2 Validating the theoretical model of CL in CAD activities

Theoretical structural validation of the model of CL in CAD activities concerns whether the model's elements and relationships provide a coherent, empirically usable, and theoretically grounded description of how CL emerges during CAD modelling, while meeting the objectives of addressing relevant mechanisms, usefulness, robustness/generality, and simplicity (Table 2; Section 1.3). This rationale follows the conclusions of Chapter 2; while a resource-based view of CL is common, operationalisations differ, CAD-specific evidence remains narrow, and temporally informative measurement aligned with CAD process dynamics is still scarce (Section 2.4). The model proposed here addresses this gap by defining CL in a way that can be quantified and interpreted within the CAD modelling process, using the same transferable segmentation structure used to represent CAD performance.

Consistent with the intended mechanisms objective (Table 2), CL is treated from a resource perspective as the allocation of limited information-processing capacity to meet task demands [87]. The model of CL is explicitly anchored in the theoretical model of CAD performance; CL is defined from the activity (utilisation) of P, C, and M operators representing embodied cognition engaged during CAD performance. This coupling is conceptually important because it enables CL to be quantified within the same hierarchical segmentation used to represent CAD performance (e.g. across the entire task, unit tasks, unit task phases, intermediate-level subtasks, or predefined time windows), thereby preserving interpretability while remaining operational for real-world CAD modelling tasks. In terms of theoretical structural validity (Validation Square [21], [22]), this anchoring also ensures internal coherence; CL is not introduced as an independent variable detached from engineering designer's behaviour, but as a construct that follows directly from the same mechanism-based representation used to formalise CAD performance.

The relationship between operator activity and CL is formalised through an explicit set of equations (detailed in Section 4.3) grounded in related literature (Section 2.4.1). At the level of embodied cognition systems, utilisation is defined as the ratio between the time an operator is active and the available time in an observed period (Equation (33)), and total WL is expressed as the sum of system-specific utilisations (Equation (34)). This reflects WL theory that operationalises load as a ratio of required to available resources [113] and, more specifically, as capacity utilisation (the proportion of resources consumed in a given period) [114], including time-based formulations of information-processing WL [115]. Similar utilisation-based indices have been repeatedly adopted in computational WL modelling using cognitive architectures, where WL is derived from activated time of modules/servers over available time and validated against subjective ratings (typically NASA TLX) [112], [116], [118]–[121]. In that sense, the proposed equations follow an established modelling logic, while adapting it to CAD by grounding "active time" in the P, C, and M operators extracted from the CAD performance model.

To distinguish WL from CL, the model introduces base weights w_i that differentiate the contribution of P, C, and M systems (Equation (35)), operationalising the assumption that C processing contributes more to CL than P processing, while M activity typically does not contribute to CL directly. This weighting principle is consistent with WL modelling in cognitive architectures where different modules are assumed to have different impacts on WL and are weighted accordingly [112], and it aligns with conceptual WL models that distinguish processing stages and emphasise that not all components contribute equally to experienced CL [102]. Conceptually, these base weights provide a simple and transparent mechanism for translating observable operator activity into a cognitive construct while keeping the model interpretable across observation periods of different duration and granularity.

To represent differences in information-processing intensity across CAD task segments, the model extends the base-weighted utilisation formulation with segment-specific weights ω_s (Equation (36)). These weights encode the assumption that unit task phases (acquisition, execution, inspection) and intermediate-level subtasks differ in the type and amount of information (chunks) processed, and thus in CL intensity even when utilisation is similar. This extension is consistent with the common denominator across major conceptual models of CL that distinguish between a task-centred component of load and a component shaped by person, strategy, and context [13], [102], [107]. In CAD terms, utilisation captures how much of the CAD modelling time is spent with perceptual and cognitive resources engaged, whereas ω_s makes explicit that where that engagement occurs in the CAD modelling process matters because CAD task segments differ in their assumed information-processing demands. Importantly, these two components reflect different aspects of CL; utilisation captures duration of engagement, whereas ω_s operationalises relative processing intensity across segments.

This segment logic was not left purely conceptual. Because CL is covert and cannot be measured directly [13], [102] and CAD tasks are lengthy and cognitively complex, temporally sensitive psychophysiological measurement is needed to corroborate assumptions about when CL changes are expected to occur [170]. EEG is frequently selected for CL research due to its high temporal resolution and suitability for computer-based HCI tasks [147]–[149], but CAD-related EEG evidence remains sparse and reported indicators vary across task types and studies (e.g. [86], [122], [126], [130], [136], [137]; Section 2.4.2). Therefore, in this thesis, the proposed CAD task segment ordering and information-processing intensity assumptions were examined empirically in the second empirical study by comparing utilisation-based expectations with psychophysiological evidence once a candidate EEG-based CL indicator for the CAD context was identified (Sections 6.2–6.3). Although this comparison did not yield statistically significant differences in the EEG-based indicator across unit task phases or subtask clusters, it clarified an important boundary condition for robustness/generality (Table 2); utilisation-based rankings and EEG-based rankings may reflect different aspects of CL and may not produce identical segment orders in a given dataset. This result does not undermine the conceptual logic of the model, but it motivates cautious interpretation of ω_s as an assumption that may require stronger empirical corroboration across CAD tasks and samples.

Finally, while the literature shows that CL depends on interacting characteristics of the engineering designer, CAD task, CAD system, and environment [13], comprehensive inclusion of all potential influences is rarely feasible in a transparent way for complex HCI activities [92]. Consistent with the objective of simplicity (Table 2) and the view that dominant factors should be prioritised [102], the present model explicitly accounts for two task-related factors central to this thesis: engineering information format (projection type) and task complexity. These factors can be incorporated as structured adjustments to the utilisation-based formulation (through intensity gains/weight modifications), following the broader logic in WL modelling where task- or user-related influences are introduced via additional factors or weight changes rather than by expanding the model into a large parameter set (e.g. [112]–[115], [118], [121]).

Overall, the combination of 1) operator utilisations defined within the CAD performance model, 2) explicit equations linking utilisation to WL and CL, and 3) segment-specific weights reflecting hierarchical decomposition and content differences provides a coherent descriptive framework for studying CL in CAD modelling. The prescribed domain-specific theoretical model enables CL estimates to be located within the CAD modelling process and later related to CAD performance outcomes, supporting empirical analysis and comparison across tasks and participants. Taken together, the domain-specific theoretical model satisfies the intended objectives by formalising CL in an interpretable and empirically testable way that remains closely coupled to the mechanisms and structure of CAD performance. In that way, it can be stated that the research fulfilled the second objective of the thesis, as defined in Section 1.1.

Implications and directions for further research follow from the same conceptual boundaries. First, stability and transferability of segment-level intensity assumptions (ω_s) should be tested on larger samples and a broader set of CAD modelling tasks, including tasks that vary in representational demands and element interactivity. Second, the model should be examined under different CAD paradigms and interaction modalities (e.g. touch/pen input, VR/AR CAD), where both segmentation taxonomies and operator mapping rules may require adaptation while preserving the core utilisation-based logic. Third, additional structured factors (e.g. expertise or strategy classes) can be incorporated selectively to improve explanatory power, but only if they preserve interpretability and do not compromise the intended model simplicity.

7.2 Validating the method for measuring and analysing CL in CAD activities

This section validates the prescribed method for measuring and analysing CL in CAD modelling against the Validation Square framework [21], [22]. In line with this framework, the section proceeds through theoretical structural validity (the appropriateness and internal coherence of constructs and workflow), empirical structural validity (the suitability and representativeness of the selected example problems and datasets for testing the method), performance validity (usefulness of the produced outputs for the intended purpose), and theoretical performance validity (the plausibility of extend beyond the specific example problems).

Given the purpose of the thesis, the theoretical models provide the conceptual foundation for both prescribing the method and interpreting its outputs. The method is intended to measure and analyse CL in CAD modelling tasks in a way that supports its systematic relation to CAD performance, explicitly considering performance both as a process (how engineering designers interact and progress through task segments) and as outputs (the resulting CAD models and their quality-related characteristics). To operationalise this purpose, the method objectives are defined in Table 3 and organised around functionality, scope, usability, and reliability. In essence, these objectives require that the method: 1) captures dynamic changes in CL and relates them to task segments; 2) prescribes a transparent, minimal, and robust set of metrics and steps for building models of CAD performance and CL; 3) provides process- and output-based performance descriptors without interfering with task execution; and 4) remains applicable to real-world CAD setups while being understandable and reproducible for engineering design researchers. Table 3 therefore provides the baseline evaluation criteria against for the valuation steps that follow.

Accordingly, Sections 7.3.1–7.3.4 evaluate: 1) theoretical structural validity (Section 7.3.1), 2) empirical structural validity (Section 7.3.2), 3) empirical performance validity (Section 7.3.3), and 4) theoretical performance validity (Section 7.3.4). Finally, a consolidated reflection on the method's practical advantages, trade-offs, and the resulting simplified implementation is provided in Section 7.3.5.

7.2.1 Theoretical structural validity

This subsection addresses theoretical structural validity of the prescribed method. In the Validation Square framework [21], [22], theoretical structural validity concerns accepting that 1) individual constructs, methods, tools, and metrics constituting the prescribed method are appropriate for the intended purpose of measuring CL and engineering designers' CAD performance, and that 2) the way these constructs, methods, tools, and metrics were finally put together in the prescribed method for measuring and analysing CL is internally consistent.

The prescribed method (Section 5) integrates analytical (cognitive modelling), and empirical (NASA TLX as the subjective measurement and EEG as the psychophysiological measurement) approaches. Therefore, to prove theoretical structural validity of the method, each element must be supported by prior research as a suitable approach for measuring CL and/or modelling performance in CAD tasks in particular or at least HCI tasks in general. Moreover, the final prescribed method must be logically coherent in terms of information flow; each step must have clearly defined inputs; the outputs produced must be plausible given these inputs; and these outputs must represent adequate inputs for the subsequent steps. The validation argument therefore follows these two criteria in the following paragraphs.

7.2.1.1 Accepting the individual constructs and measures for the intended purpose

The first part of theoretical structural validation addressed whether the elements used in the prescribed method are acceptable choices for its intended purpose (stated in Section 1.3). It therefore asks whether each element integrated into the method is theoretically justified and supported in relevant research streams. This criterion matters because CL is a cover, multidimensional construct with no singe universally accepted operational definition; as a result, research often relies on multiple indirect indicators that capture complementary aspects of CL [13], [101], [102], [105]. In line with this, the prescribed method integrates analytical and empirical approaches, consistent with established classifications of CL assessment methods (Table 19). It also acknowledges the temporal nature of CL (instantaneous, peak, accumulated, average, overall), which motivates the inclusion of measures that can reflect changes over time in complex, lengthy tasks such as CAD [101], [102]. In this thesis, method integration does not mean suggesting a single fused ground-truth value. Instead, it means using complementary measures with defined roles (subjective ratings, behavioural/process indicators, and psychophysiological responses) to check converging or diverging evidence under the same task conditions and aligned task segments (see flowcharts in Chapter 5, and the segmentation and operator mapping in Section 3.4 and Section 4).

Theoretical structural validation was performed through the critical evaluation of literature reviewed in Chapter 2 as a part of Descriptive study I. First, the literature (Section 2.4.3.1) supports EEG as a psychophysiological method for quantifying changes in brain activity associated with CL, particularly in computer-based tasks where continuous visual processing, WM demands, and ongoing interaction with an interface are central [147], [149]. EEG is frequently selected in CL research due to its high temporal resolution and usability in more naturalistic settings, including wireless systems that allow relatively free movement. The latter is an important practical consideration for CAD studies that depend on mouse and keyboard interaction [147], [148]. Moreover, EEG has been applied in design and CAD-relevant HCI tasks (e.g. designing on touchpads [126], [129], [141], complex visuospatial tasks [133], simulation-based HCI tasks [134], and drawing in AutoCAD [137]), with multiple EEG features reported as sensitive to CL changes (see Table 20 in Section 2.4.3 for details). At the same time, prior work also shows that EEG-based CL indicators are not fully consistent across tasks and studies, and that task type can influence which frequency-band features appear sensitive to CL [130]. This directly justifies the method's approach of using EEG as an informative psychophysiological measure, interpreted cautiously and in combination with other method components, rather than treating any single EEG feature as a universal CL indicator [130], [132], [138].

Second, the literature (Section 2.4.3) shows that NASA TLX is widely used as a subjective method for capturing perceived WL/CL across HCI and engineering design contexts, and it remains one of the most commonly adopted self-report tools for CL assessment [169]. In the prescribed method, NASA TLX is therefore included as a complementary perspective to EEG and performance modelling; it supports validation of CL changes as it helps interpret whether experimentally induced differences (e.g. by varying projection format or task complexity) were perceived as more or less demanding at an overall-task level. Importantly, NASA TLX is also commonly used as a validation reference in both utilisation- and EEG-based WL/CL modelling, where model-derived WL estimates are compared or correlated with subjective ratings to evaluate whether the model captures perceived WL/CL [112], [116], [118]–[120]. In that sense, NASA TLX provides a bridge between subjective experience and operator utilisations/EEG features by enabling direct comparison between perceived WL/CL and utilisation- or EEG-based WL/CL estimates. This role is consistent with broader CL research practice, where subjective measures are often used alongside psychophysiological methods to strengthen interpretation and mitigate the limitations of relying on a single indicator for a covert construct [13], [102]. It also aligns with prior studies that explicitly combined EEG and NASA TLX to validate interpretations of CL differences in complex design and HCI tasks [129], [134].

Third, the literature (Section 2.3) supports cognitive modelling (as an analytical approach) as a valid way to represent and quantify human performance in interactive computer tasks, including for evaluation and redesign of HCI systems [101], [105]. This provides a clear rationale for modelling CAD performance as an HCI activity based on observable interaction behaviour. Importantly, cognitive modelling has a history of being used not only to describe task structure but also to support systematic analysis of user procedures and performance mechanisms, which is consistent with the goal of formalising CAD modelling behaviour in a transparent and reusable way [101], [105].

Finally, prior work (Section 2.4.1.3) also supports cognitive modelling as a basis for representing CL-relevant processing by linking WL to information processing and resource utilisation. Studies using cognitive architectures and related modelling approaches commonly quantify WL/CL through module/server activation time relative to available time (i.e. utilisation), and validate these estimates against subjective WL ratings [112], [116], [118]–[121]. This practice is directly aligned with the method's approach of linking CL to the utilisation of P and C operators, and of treating CL as a resource-based construct [113]–[115] [115]. In this sense, the method's structure follows an established modelling logic; utilisation is used as a core parameter for WL, while adapting it to CAD by grounding operator activity in empirically observable CAD behaviour and by combining it with psychophysiological and subjective perspectives.

Together, these arguments support the first validation criterion from the Validation Square [21], [22]; the method is assembled from constructs and measurement approaches that have already been accepted in relevant research streams, and their roles in the method are consistent with how they have been used in prior work.

7.2.1.2 Accepting the internal consistency of the integrated method

The second part of theoretical structural validation addressed criterion concerned with whether the method is internally consistent as an integrated workflow [21], [22]. In this thesis, this is demonstrated through the flowchart representations of the prescribed steps in the workflow (Section 5.2 and Section 5.3) and the summary of required datasets (see Table 57 in Chapter 5). The intention of these flowcharts was to demonstrate that: 1) each step has adequate and realistic inputs available from the required datasets (summarised in Table 57); 2) the step's anticipated outputs are plausible given the defined inputs and processing logic; and 3) the outputs are suitable inputs to the subsequent steps, enabling the method to be executed end-to-end without conceptual gaps.

The method's structure supports this consistency. The steps prescribed for building the model of CAD performance (Figure 45) progress the workflow from analysing observed HCI in a CAD task (HCI script), to analysing the hierarchical structure of CAD tasks (timestamped decomposition into three task/goal levels, three process layers, and three unit task phases), and finally to mapping P, C, and M operators using defined heuristics to produce a model of CAD performance. The outputs of each step are directly used by the next step. Moreover, the built model of CAD performance is used to calculate operator utilisations per task segment, which are then combined with EEG-based CL indicators to assign information-processing weights to P and C utilisations. These quantified values are then used in the CL equation to calculate CL in task segments and to visualise its dynamics over the CAD performance model (see Figure 49). This validates the method based on the second criterion of the theoretical structural validity; the method is coherent as a system, with explicit input–output links between steps and clearly defined required datasets [21], [22].

Finally, the method is not intended to be universal for all design activities. Its domain of application is constrained by its theoretical assumptions (based on the human information-processing perspective) and by the practical objective for capturing and modelling processed-based perspective of HCI in CAD. The application of the proposed method is limited to:

1) CAD modelling tasks (not designing in a broader sense); the method focuses on CAD modelling as interaction with a CAD system and is grounded in MHP/KLM logic for describing HCI and information processing during CAD modelling. It therefore does not represent higher-level design problem solving (e.g. creativity, concept generation, or open-ended synthesis), which may require different operators and mechanisms.
2) Feature-based solid modelling CAD systems with conventional HCI devices; the method assumes a feature-based approach to modelling solids and interaction through a standard monitor, keyboard, and mouse. Other workflows (e.g. VR-based CAD, touch/pen input) may require adjustments to operator definitions, mapping heuristics, and data capture.
3) Description rather than prescription of performance and CL; the method aims to describe CAD performance and its associated CL, not to prescribe their values. This decision is aligned with the high variability in CAD modelling approaches, even among skilled engineering designers, and with the fact that the same solid part can be created through many valid CAD modelling sequences.

These constraints are not weaknesses by themselves; they define where the method's assumptions are expected to hold and where its outputs can be interpreted with confidence.

7.2.2 Empirical structural validity

This section addresses empirical structural validity of the prescribed method, i.e. whether the example problems used in the empirical studies are appropriate for testing the prescribed method [21], [22]. In line with the criteria defined in [21], [22], the example problems should satisfy three requirements: 1) they should be sufficiently similar to tasks for which the adopted constructs and measurement approaches are accepted (so that using human performance modelling and EEG is warranted), 2) they should be representative of the CAD modelling problems for which the method is intended (so that the results generalise to the intended application domain), and 3) they should yield data that are sufficient to execute the method end-to-end and thus judge its performance.

Defining the experimental tasks therefore involved a deliberate trade-off between the first two criteria; tasks had to remain tractable enough to support transparent decomposition and operator mapping, while still reflecting fundamental CAD part-modelling practice. Concretely, the studies used well-specified, screen-mediated part-modelling problems that preserve key characteristics of CAD modelling (sustained visual processing, WM demands, and continuous HCI), but remain bounded enough to allow reproducible segmentation and systematic analysis. The experimental task designs and experimental procedures are described in the study-design sections of Chapter 3 (first empirical study) and Chapter 6 (second empirical study), the procedure of defining the hierarchical task structure and operator mapping rules is described in Section 6.4, while the final version is detailed in Section 4.2 and later tested in the second empirical study (Chapter 6).

The remainder of this section evaluated the example problems against the three criteria above.

First, it was established that the example problems (experimental tasks) are sufficiently similar to those for which the selected constructs and methods are generally accepted. As concluded in Section 2.3, human performance (cognitive) modelling has a strong foundation in HCI research and has been applied to real-world activities, although many published examples are typically based on tasks that are less complex than full CAD modelling sessions. Therefore, human performance modelling is most appropriate when CAD is treated as an HCI process, because it provides explicit constructs for representing how observable interaction unfolds under limited information-processing capacity. In that sense, the Model Human Processor (MHP) and GOMS-style task analysis offer a pragmatic modelling baseline; they support explanation and performance estimation at the interaction level without requiring the strong cognitive specification (e.g. production rules and detailed memory representations), which is difficult to validate for complex and variable CAD workflows. The experimental tasks were therefore intentionally scoped to match the kinds of interactive problems for which these approaches have been most successfully used; they were well-specified, goal-directed, screen-mediated part-modelling activities with identifiable subgoals and repeatable operation and action sequences.

In parallel, EEG has been used to study design cognition, and it has been widely applied for measuring CL in tasks that share key characteristics with CAD (e.g. sustained visual processing, WM demands, and continuous interaction with digital interfaces), even when CAD itself is not the direct context.

Second, the example problems were designed to be representative of the actual problems for which the method is intended. The domain of application was explicitly defined in the research clarification stage and extended in the prescribed method (Section 5.1) through characteristics of suitable CAD modelling tasks. Accordingly, the selected tasks covered essential CAD modelling steps that can be generalised at the level of clusters of CAD operations/intermediate-level subtasks. Moreover, the proposed hierarchical structure of CAD tasks was defined across multiple levels to account for low-level, intermediate-level, and high-level subtasks/subgoals as well as multiple process layers, considering CAD actions, CAD operations, and unit task operations, observed in CAD modelling. This structure was initially suggested based on six CAD modelling sessions with three participants in the first experimental study (Section 3.4) and then validated in the second experimental study using a different participant group and different CAD modelling objects (Section 6.4), supporting its robustness across tasks. At the same time, the example problems were not intended to represent the full range of CAD activities, as they did not include workflows such as assembly modelling or technical documentation.

Third, the empirical studies showed that the example problems provide sufficient data to support conclusions about method performance. In the second experimental study, two CAD tasks with different complexity levels were used to elicit differences in CL; this controlled variation of the condition was grounded theoretically (via element interactivity) and supported empirically through NASA TLX scores as well as output- and process-based CAD performance metrics. In addition, the datasets collected in the studies provided all required inputs specified by the method, satisfied the input requirements, and enabled generation of the anticipated outputs at each step, such that the output requirements were met and the method could be executed end-to-end on the example problems.

7.2.3 Empirical performance validity

This section addresses empirical performance validity, which evaluates whether the outcomes of the prescribed method are useful with respect to the initial purpose (detailed in Section 1.3) for the selected example problems, and whether this usefulness can be attributed to applying the final integrated method rather than to any single construct applied in isolation [21], [22].

Consistent with this definition, the discussion is structured around three criteria: 1) usefulness for the example problems and stated purposes, 2) attribution of usefulness to the integrated method, and 3) practical usability and reproducibility.

7.2.3.1 Usefulness for the example problems and stated purpose

To build confidence in usefulness, the method was implemented in two experimental studies, and the resulting method outputs were evaluated against the degree to which the articulated purposes and objectives (Table 3; Section 1.3) were achieved. In particular, usefulness was assessed with respect to whether the method can:

1) Capture dynamic changes in CL and relate them to task segments (Functionality in Table 3). This was demonstrated by calculating CL estimates at multiple hierarchical levels and process analysis layers: task-level relations to NASA TLX (Section 6.5), segment-level differences and contributions (Section 6.6), and moving window time series with cross-indicator alignment analyses (Section 6.7).
2) Provide explicit, end-to-end steps and required inputs (Functionality and Reliability in Table 3). Chapter 5 specifies the workflow and data requirements, while the second experimental study operationalises them; candidate EEG features are tested and shortlisted for the CAD context (Section 6.2), segment-level information-processing weights are derived from the selected EEG feature (Section 6.3), the CAD performance model is instantiated and produces the required hierarchical segmentation and operator utilisations (Section 6.4), and the resulting indicators are evaluated against subjective CL (Section 6.5-6.7) and performance outcomes (Section 6.8).
3) Generate coherent CAD performance descriptors that support CL interpretation (Functionality in Table 3), by explicitly representing performance as both process (hierarchical segmentation, unit task phases, operator utilisation) and outputs (performance outcomes). This is shown through the model instantiation and utilisation computation (Section 6.4), and through linking CL estimates to CAD performance outcomes (Section 6.8).
4) Operate in realistic CAD modelling without disrupting task execution while remaining implementable by engineering design researchers (Scope and Usability in Table 3). The second empirical study demonstrated feasibility of collecting the required behavioural traces and EEG signals during CAD modelling and processing them into the method outputs under a realistic CAD task (Section 6.1; required datasets summarised in Table 57; workflow described in Chapter 5).

In this thesis, usefulness is primarily linked to generating knowledge that supports further scientific knowledge production; first through theoretical contributions, namely the formalisation of CL and CAD performance via explicit theoretical models (Section 4), and second through practical contributions in the form of an operational how-to procedure that specifies required inputs, processing steps, and outputs for analysing CL in relation to CAD performance (Section 5). From an industrial perspective, usefulness is framed in terms of potential for reducing time and costs and, ultimately, supporting improvements in performance quality. Within this framing, the thesis outcomes support a narrower but defensible claim; the proposed models and method primarily contribute by reducing the time (and related costs) needed to investigate CL in relation to CAD performance, because they offer a structured and replicable pipeline and explicit requirements for data collection and processing. However, the thesis does not directly prescribe how to improve CAD modelling quality. Instead, it provides means for doing so in subsequent work by delivering 1) process-based representations that localise CL to CAD task segments, and 2) performance metrics (process- and output-based) that can be used as targets in future prescriptive studies.

7.2.3.2 Attribution of usefulness to the integrated method

In this thesis, integration does not mean merely using multiple measures in parallel. Instead, the method defines a shared segmentation and information flow so that the method components interact in a structured way:

- The CAD performance model provides a hierarchical, time-stamped segmentation and operator utilisations (Section 5.2; operationalised in Section 6.4)
- EEG is not only recorded as an external correlate; it is used to derive information-processing weights for segments (Section 6.3), which can then parametrise the utilisation-based CL equation (Section 5.3.4, evaluated in Section 6.5).
- Subjective ratings (NASA TLX) provide a task-level validation reference for whether inferred CL relates to the perceived mental demand (Section 6.5).

This workflow makes it possible to test, within the same segment definitions, what is gained by adding or removing specific method components, which directly targets the "attribution" criterion of empirical performance validity.

Attribution of usefulness to the full workflow was supported in the second empirical study by testing method variants rather than presenting a single multimodal outcome. In Section 6.5, overall task-level CL estimates were computed using four utilisation-based indicators that systematically vary two design choices: 1) participant-specific versus global utilisations, and 2) non-weighted versus EEG-weighted utilisations (using segment-level information-processing weights ω derived from parietal alpha TRP). Across all four indicators, Theil–Sen models showed a significant positive association with NASA TLX mental demand, indicating that utilisation-based CL robustly covaries with subjective CL in this dataset. Adding EEG-derived weighting yielded only a (small) reduction in prediction error (weighted indicators had slightly lower LOOCV MAE than their non-weighted counterparts), consistent with the earlier finding that segment-level information-processing weights were not stable in this dataset (no reliable TRP differences between segments and low agreement in segment ranking; Section 6.3). The comparison of participant-specific and global variants further showed that global utilisation summaries can approximate (or exceed) participant-specific median summaries with respect to NASA TLX mental demand prediction (Global U outperformed Median U; Weighted Global U was close to Weighted Median U). Together, these results clarify attribution; the primary explanatory signal is captured by the behavioural/analytical component (segment-aligned utilisation aggregated to the task level), while the EEG-based information-processing weighting acts as a modest, supportive calibration refinement.

Finally, Section 6.8 links CL estimates to CAD performance outcomes. Here, utilisation-derived CL indicators show consistent negative relationships with overall performance, supporting the method's ability to generate outputs that are practically informative for the example problems. Importantly, the strongest performance relationship is observed for the weighted participant-specific indicator (Weighted Median U), whereas the fully global weighted formulation was not consistently significant (Section 6.8). This pattern supports a bounded attribution claim; integration via EEG-derived information-processing weighting appears most informative when individual utilisation patterns are retained, while global simplifications improve usability but can reduce sensitivity for certain inference goals (notably individual-level performance prediction).

7.2.3.3 Practical usability and reproducibility

Beyond demonstrating that the method "works" conceptually and empirically, validation also requires that the workflow is feasible to execute in realistic CAD contexts and that it can be reproduced by other engineering design researchers. In this thesis, usability is addressed by prescribing the workflow in a stepwise manner (Chapter 5) and by explicitly stating the required inputs and outputs (e.g. required datasets in Table 57), such that an engineering design researcher can implement the method without relying on implicit expertise. The method also targets usability through its emphasis on minimising analyst burden; it favours a limited set of core metrics and prescribes clear modelling steps that prioritise transparency and interpretability when linking CL to CAD task segments and performance outcomes (Table 3).

Reproducibility is supported by the fact that the required inputs (CAD interaction traces, hierarchical segmentation outputs, and (where applicable) EEG recordings) were collected successfully under the CAD experimental setup and processed into the method outputs using the prescribed steps (implementation in Sections 6.2–6.4; workflow specification in Chapter 5). At the same time, the full method includes steps that can be time- and expertise-intensive (notably EEG preprocessing and detailed operator mapping). For this reason, it is useful to define a simplified, usability-oriented implementation that preserves the method's core logic of segment-based CL estimation while reducing implementation effort.

A simplified implementation is appropriate when the primary goal is to obtain an interpretable, segment-level proxy of CL dynamics without fully parameterising the method for each individual participant. This includes contexts such as: 1) early-stage studies and pilot deployments where feasibility and rapid iteration are prioritised; 2) resource-constrained settings where EEG acquisition/processing is not available or where detailed operator mapping is impractical; and 3) comparative or descriptive analyses where the emphasis is on visualising and comparing CL patterns across tasks or experimental conditions at the level of task segments rather than modelling fine-grained individual differences.

Because the simplified implementation relies on shared parameter values assigned to CAD task segments, it is most suitable when the task structure and interaction modality are comparable to those used to derive these values (i.e. the same segmentation scheme and similar CAD modelling task). It is less suitable for claims that require high individual specificity (e.g. participant-level diagnosis or performance prediction), because individual variability in segment utilisations is here intentionally reduced.

To maximise usability while preserving the method's core intent, the simplified implementation retains the hierarchical decomposition at the level of unit task phases and intermediate-level subtasks, but omits participant-specific operator mapping and utilisation estimation. Instead, CL is computed using global segment utilisation values (and, where desired, dataset-derived information-processing weights), reducing the required inputs primarily to segment timing information. The following are the proposed steps of the simplified method:

1) Analyse the HCI in a CAD task to extract CAD events based on the list available in Table 54 to create a HCI script
2) Analyse the structure of a CAD task using the HCI script and the proposed levels of the hierarchical structure of CAD tasks (see Section 4.2) to create a timestamped

hierarchical decomposition into unit task phases (*acquisition*, *execution*, *inspection*) and intermediate-level subtasks (*navigating*, *referencing*, *drawing*, *dimensioning*, *constraining*, *manipulating*, *deleting*). The required set of information for each unit task phase and subtask consists of start time, end time, and duration.

3) Compute CL for each intermediate-level subtask using the equation for calculating CL (Equation (44)), global values for P and C utilisations (suggested in Table 74 for this dataset), and information-processing weights (suggested in Table 74 for this dataset).

This simplified implementation still enables modelling and visualising how CL changes across the CAD modelling sequence, but at the temporal granularity of intermediate-level subtasks. Example outputs produced using this approach are shown for representative performers; Figure 73 for P#4 representing a High Performer, Figure 74 for P#16 representing an Average Performer, and Figure 75 for P#8 representing a Low Performer. The main trade-off is reduced temporal and individual precision in exchange for substantially lower analysis effort and easier adoption. The main limitation is generalisability of the shared (global) values beyond the present dataset; extending and validating global segment values across larger samples, CAD systems, and tasks is therefore identified as an important direction for future work.

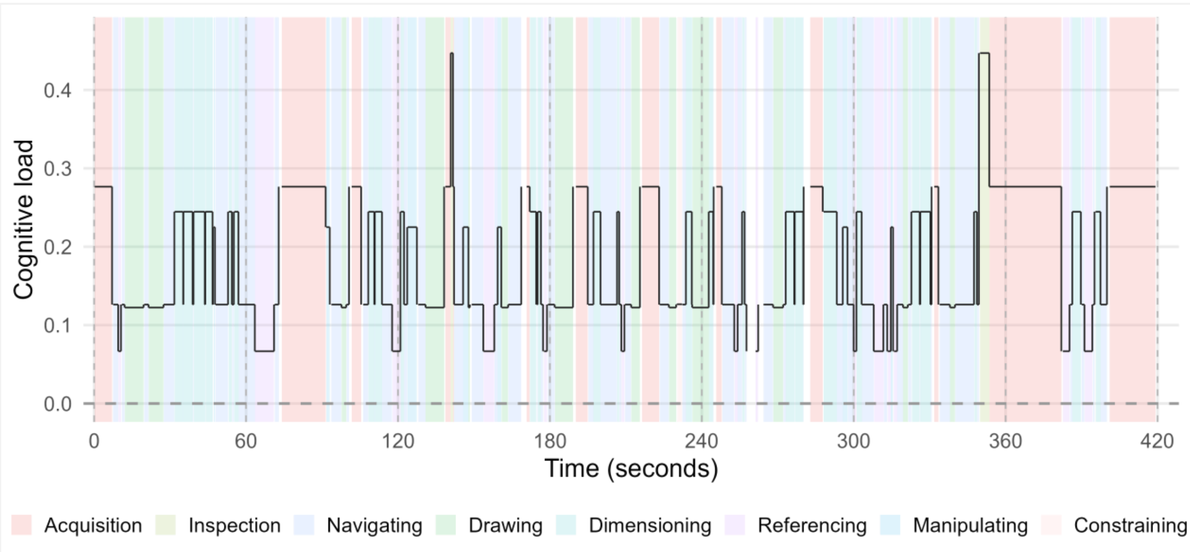

Figure 73 P#4: Dynamic changes in CL indicated by global values of weighted utilisations – High Performer

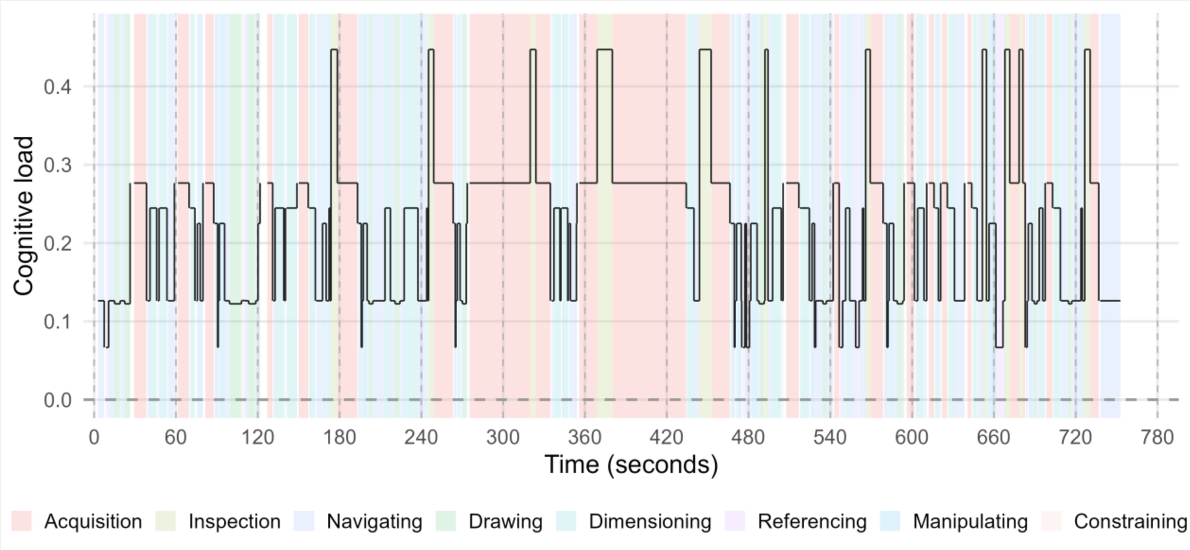

Figure 74 P#16: Dynamic changes in CL indicated by global values of weighted utilisations – Average Performer

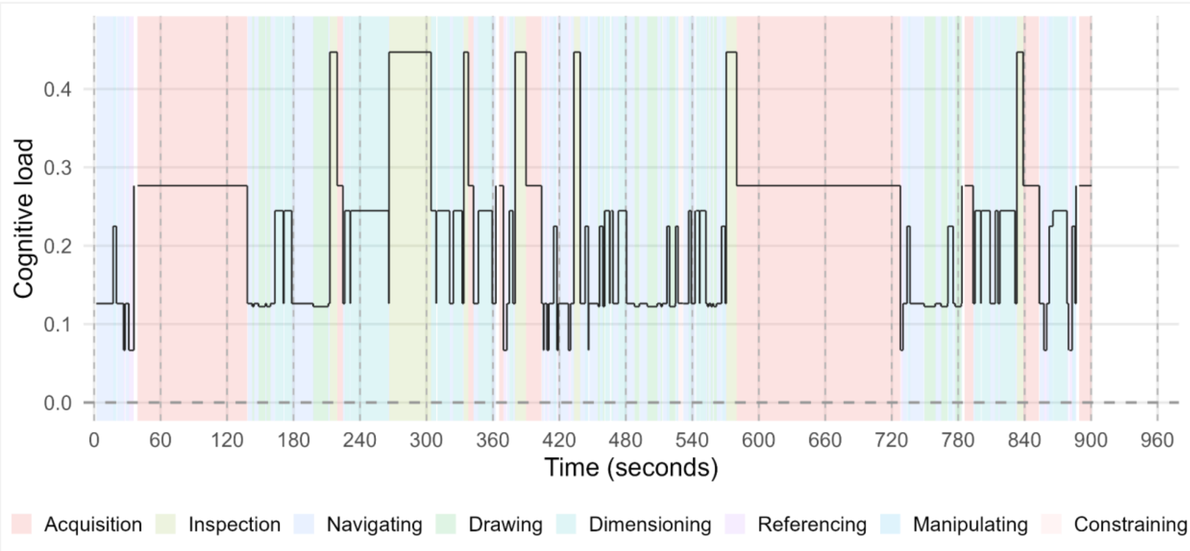

Figure 75 P#8: Dynamic changes in CL indicated by global values of weighted utilisations – Low Performer

7.2.4 Theoretical performance validity

In the Validation Square framework, theoretical performance validity concerns whether the usefulness demonstrated on the selected example problems can be reasonably expected to extend beyond those specific experimental tasks [21], [22]. This final validation step builds on the preceding validation results; what kinds of CAD tasks and study contexts are eligible for applying the method, what evidence supports transfer, and where adaptation is required.

The argument for the extension beyond the example problems rests on two foundations established earlier in the chapter. First, the usefulness of the method components was supported through the empirical studies and the structural validation grounded in previous literature; human performance (cognitive) modelling and operator utilisation provide an interpretable, process-based representation of CAD modelling; EEG provides a psychophysiological perspective on information-processing intensity; and NASA TLX provides a widely used subjective reference for perceived mental demand. Second, usefulness of the integrated method was examined in the second empirical study through end-to-end execution and explicit testing of CL indicator variants. In particular, Section 6.5 showed that task-level utilisation-derived CL estimates robustly covary with NASA TLX mental demand across all tested indicator variants, while EEG-informed weighting yields modest refinement in this dataset and should be interpreted as supportive rather than decisive given the limited stability of segment-level information-processing weights (Section 6.3). The same section also provides evidence relevant to transfer and usability; global utilisation values can approximate (and in some cases exceed) participant-specific summaries with respect to their relationship to subjective CL, which motivates a simplified implementation for contexts where full parameterisation is impractical.

The simplified implementation of the method was not subjected to a separate, full empirical validation cycle. Still, it is explicitly derived from 1) constructs whose usefulness was already supported within the empirical performance validation (Section 7.3.3), and 2) comparative results demonstrating that, for the task-level association with NASA TLX mental demand, global utilisation summaries can preserve the core variations captured by utilisations (Section 6.5). Thus, the simplified implementation is positioned as a bounded transfer of the validated method; it targets the usability objectives in Table 3 while preserving the method's functionality of estimating CL changes with respect to CAD task segments.

The method can be expected to extend to CAD tasks that match its theoretical assumptions (Chapter 5) and data requirements (Table 57), namely:

1) Feature-based part modelling in conventional CAD systems. The method assumes a feature-based solid modelling approach where engineering designers construct geometry through sketches and CAD features and where CAD modelling progress can be represented as a sequence of recurring task segments.
2) Standard HCI interaction modality. The method is designed for screen-mediated interaction with standard devices (monitor, mouse, keyboard). This supports the operator-mapping logic and the utilisation interpretation grounded in human information-processing perspective.
3) Segmentability into the proposed task/goal hierarchy. Applicability is strongest when the CAD modelling activity can be decomposed into unit task phases and intermediate-level subtasks of the type used in the method (e.g. navigating, referencing, drawing,

dimensioning, constraining, manipulating, deleting), so that CL can be interpreted relative to segments that recur across many CAD modelling tasks.

4) Non-intrusive data acquisition. The required behavioural traces (and, when used, EEG) must be obtainable without interfering with task execution, consistent with the functionality and reliability objectives in Table 3.

The method is not intended to be universal for all design activities. Direct application is limited (without adaptation) for several cases:

- Broader design problem solving that is open-ended (e.g. concept generation, creativity-driven synthesis), where the task structure and cognitive mechanisms extend beyond the CAD-modelling-focused HCI process captured by the operator mapping and segment taxonomy.
- CAD activities outside part modelling. such as assembly modelling, technical documentation/drawing production, or simulation/analysis workflows, where the recurring segment set and performance descriptors may differ and would require revising the segmentation scheme, operator mapping rules, and, if used, segment weighting logic.
- Non-standard interaction modalities (e.g. VR/AR CAD, pen/touch-first interfaces), which may require redefinition of operators, mapping heuristics, and potentially different psychophysiological features or validation references.

A key reason extension is plausible within the intended domain is the choice to represent CAD modelling at the level of intermediate-level subtasks (CAD operation clusters) and unit task phases. These subtasks recur across many part-modelling workflows regardless of specific geometry or CAD modelling sequence, enabling the method to remain applicable across a wide range of parts while still localising CL to meaningful segments. This abstraction level supports theoretical generalisation because it reduces reliance on participant-specific command sequences and instead anchors CL interpretation to higher-level, repeatedly observed behavioural segments.

For eligible tasks, theoretical performance validity also depends on using the implementation that matches the inference goal. The full implementation is most appropriate when the objective requires sensitivity to individual differences in how engineering designers allocate perceptual and cognitive resources across CAD task segments, or when linking CL to individual-level performance outcomes is central. The simplified implementation, based on global utilisation values, is appropriate when the goal is a transparent, segment-level proxy of CL dynamics under practical constraints (limited resources, no EEG, or reduced feasibility for participant-specific operator modelling), and where tasks closely match the segmentation structure used to derive the global values in this dataset.

7.2.5 Consolidated method, advantages, and trade-offs

The validation results above support a consolidated interpretation of what the prescribed method contributes, what drives its usefulness, and where simplification is justified. The method contributes a process-based, multimodal approach to studying CL in CAD modelling by combining:

1) Cognitive modelling to describe what the engineering designer does (CAD performance model) and how it does it (operator utilisations);
2) EEG to capture how information-processing intensity changes during CAD modelling (EEG-based CL indicators);
3) NASA TLX to incorporate the engineering designer's perceived WL as an additional, widely used reference.

This integration is advantageous relative to what is available in the literature because CL is a covert, multidimensional construct with no universally accepted operational definition or single gold-standard measure [13], [101], [105]. As a result, any single method is likely to capture only one aspect of the construct (e.g. perceived CL from self-reports, interaction effort and strategy from process measures, or neurophysiological engagement from EEG). Conclusions are typically strengthened when complementary sources converge. The here proposed method therefore triangulates insights from subjective ratings, behavioural/process indicators, and psychophysiological signals as complementary sources of CL evidence, not interchangeable measures of the same quantity [13], [101], [105]. In this thesis, triangulation means using multiple indicators to cross-check CL interpretations; convergence across sources strengthens an inference about CL changes, while divergence helps localise what aspect of CAD task demands is changing (e.g. perceived effort vs. interaction intensity vs. neurophysiological engagement), rather than assuming any single metric is sufficient. The proposed method integrates these sources through a shared process representation; CAD activity is segmented into the same time-aligned task hierarchy, process indicators (e.g. time distributions and operator utilisations) are computed for the same segments, and EEG features are extracted on aligned time windows and then mapped to the same segments for interpretation and validation. This differs from many prior studies that collect multiple measures but report them in parallel; here, the measures are integrated through explicit alignment rules and joint analyses so that subjective, behavioural, and EEG evidence can be interpreted consistently with respect to the same CAD modelling segments [13], [101], [105].

A key advantage is that the method produces a dynamic description across task segments at three suggested hierarchical levels; CL is not only measured as a single score per task, but can be estimated across task segments and visualised relative to the model of CAD performance. This directly addresses a practical limitation of dominant subjective approaches (e.g. NASA TLX), which are typically administered at a level of an overall task and therefore provide limited insight into when CL changes within lengthy, complex CAD activities [101], [102], [181]. It also leverages the high temporal resolution as the main strength of EEG, which is repeatedly highlighted as a reason EEG is commonly selected for studying CL in complex computer-based tasks where cognitive processes rapidly change throughout execution [147], [149]. By anchoring EEG features to a task structure derived from CAD modelling, the method responds to the common challenge that it is often unclear at which stages of CAD tasks CL should increase, motivating segment-based rather than analysis at the level of the overall task [170].

Compared to approaches that rely only on CAD logs or output-based CAD performance metrics, the method's segmentation and operator-based modelling provide an interpretable mechanism for relating CL to how CAD modelling unfolds, not only to final outcomes. Prior CAD-process studies have used diverse segmentations (often at either high-level phases or low-level command categories), making results hard to compare and limiting their ability to explain cognitive demands beyond counts and distributions of these segments (see Section 2.2.1). The prescribed method advances this by tying CL estimation to a theory-grounded process representation (via operator utilisation), aligned with modelling practices that quantify WL/CL through resource activation time relative to available time (i.e. utilisation; see Section 2.4.1.3 for details). At the same time, it avoids over-reliance on any "universal" EEG feature, which is important because prior work shows that EEG-based CL indicators can vary across tasks and studies and should be interpreted cautiously, ideally alongside other measures. This is visible from literature review (Section 2.4.2.1) and confirmed with the results of the second empirical study (Section 6.2), which narrowed the most promising candidates to Alpha Parietal TRP and Alpha Occipital TRP but failed to provide statistical evidence for prescribing one of them as the best EEG-based CL indicator.

The validation results motivate bounded, usability-oriented consolidation. Empirical performance validity (Section 7.3.3) shows that the primary explanatory signal for the task-level association with NASA TLX mental demand is captured robustly by the utilisation-based component, while EEG-derived information-processing weighting provides modest refinement in this dataset (consistent with limited stability of segment-level weights in Section 6.3). This evidence supports presenting the simplified implementation not as a new, independently validated method, but as a constrained derivative of the validated workflow that targets usability objectives (Table 3) while preserving the core functionality of segment-based CL estimation. The trade-off is explicit; reduced analyst burden and easier adoption in exchange for reduced individual specificity and potentially reduced sensitivity for inference goals that depend on individual differences (e.g. participant-level performance prediction).

7.3 Confirming the research hypothesis

In Section 1, the research hypothesised that *the method for measuring and analysing CL, based on monitoring psychophysiological responses, enables assessment of the engineering designer's level of CAD activity performance*. Conceptually, this hypothesis is grounded in the view of CAD modelling as an HCI activity in which performance emerges from the allocation of limited cognitive resources to meet task demands, and in which increases in CL should be reflected in measurable changes in brain activity during CAD modelling. In that sense, psychophysiological monitoring (EEG) provides a continuous, objective perspective on information processing demands during CAD activity, complementing traditional performance assessment based on CAD modelling outcomes and process metrics. Methodologically, the hypothesis implies that CL estimates derived from EEG should be meaningfully related to performance, either by differentiating performance levels (e.g. high vs. low performers), by covarying with performance outcomes (e.g. CAD model completeness, CAD modelling errors, or efficiency), or by explaining performance-relevant process characteristics captured in the CAD performance model (e.g. unit task structure). At the same time, the literature and the present results suggest that this relationship is not expected to be deterministic; CL is multidimensional, EEG-based CL indicators are task- and context-sensitive, and performance depends on additional factors (knowledge, strategy, and task characteristics). Rather, the hypothesis should be interpreted as proposing that psychophysiological CL monitoring can contribute to performance assessment by adding explanatory power about how demanding the CAD modelling process was, and when higher demands occurred, rather than serving as a standalone measure of CAD performance.

Building on the validation results discussed in Section 7.2, particularly the evidence that the method produces coherent, segment-level CL estimates and performance descriptors, this hypothesis was examined empirically in the second experimental study by directly relating task-level CL estimates to overall CAD activity performance (Section 6.8). Overall performance was operationalised as a composite assessment based on completeness of the final CAD model and CAD modelling duration. Robust Theil–Sen regression models were fitted between overall performance and four utilisation-based CL indicators (Weighted Median U, Weighted Global U, Median U, Global U), where the weighted indicators incorporate EEG-derived information-processing weights (ω) estimated from parietal alpha TRP, while the non-weighted indicators reflect utilisation-only estimates.

The results are consistent with the hypothesis, with clear boundary conditions. All estimated slopes were negative, indicating that higher utilisation-based CL is associated with lower overall performance. Statistically reliable associations were observed for Weighted Median U ($b = -1.12$, 95% CI [−2.14, −0.64], $p = 8.25 \cdot 10^{-5}$), Median U ($b = -0.15$, $p = 1.77 \cdot 10^{-4}$), and Global U ($b = -0.25$, $p = 3.21 \cdot 10^{-2}$), whereas Weighted Global U did not reach significance ($p = 6.90 \cdot 10^{-2}$). This pattern clarifies that CL estimates can serve as performance-informative indicators, but their strength depends on operationalisation; the clearest relationship emerged when EEG-derived weighting was applied to participant-specific utilisations (Weighted Median U), while the weighted global formulation was less reliable for capturing performance differences in this dataset.

Taken together, the hypothesis is partially supported with this research. The proposed method demonstrates that CL estimates, especially the EEG-informed, participant-specific formulation, are significantly related to overall CAD activity performance and can therefore contribute to assessing performance level in CAD tasks. At the same time, the results do not justify a broad claim that EEG-based weighting universally improves performance estimation, nor that CL indicators alone provide comprehensive performance prediction. Generalisability and sensitivity should therefore be confirmed on larger samples and additional CAD task types, and future work should further examine when individual-specific parameterisation is necessary versus when simplified global formulations are sufficient for the intended inference goal.

7.4 Limitations

Several limitations and boundaries arise from conceptual choices in the models and from methodological choices in the empirical studies and method implementation. Here, structural choices refer to the deliberate modelling decisions that define the representation and information flow of the approach; most importantly the proposed task/goal hierarchy, process layers, the selected operator set (P, C, M), the heuristic mapping rules that link observable CAD events to operator activity, and the way CL is computed from utilisations.

First, the CAD performance model is tailored to feature-based part CAD modelling performed with conventional HCI devices (monitor, mouse, and keyboard). Other CAD tasks (e.g. assembly modelling, technical documentation generation, simulation workflows) and broader design activity outside CAD would require extending the model with additional goal structures, operation/action classes, and segment definitions. In addition, mapping perceptual and cognitive operators necessarily relies on heuristics grounded in psychological assumptions, which enables operationalisation but makes results sensitive to segmentation decisions and interpretation of behaviour in ambiguous cases (e.g. when "thinking" is not accompanied by overt interaction). The model also does not include factors such as expertise or experience as separate state variables or parameters; their influence is expected to appear indirectly through differences in CAD action or operation distributions, durations, and inferred operator activity rather than being explicitly modelled as an input.

Second, because the CL model is grounded in the CAD performance model, uncertainty in segmentation and operator mapping propagates into utilisation-based CL estimates. This is especially relevant for comparisons that depend on fine temporal localisation (segment-level inference), where small boundary shifts or alternative mapping interpretations may meaningfully change utilisation values. Moreover, segment weights ω_s should not be treated as universal constants; their stability is expected to depend on task characteristics (including complexity), representation formats, and potentially participant characteristics. This boundary is an expected consequence of targeting a minimal, transferable formulation, but it also motivates future work aimed at refining and re-estimating weights across broader task sets and CAD contexts.

Third, the choice between the full and simplified implementation is driven by the inference goal and practical constraints. The full method is most appropriate when the goal requires individual-specific sensitivity (e.g. analysing individual differences in CL dynamics, relating CL to individual performance differences, or interpreting CL relative to participant-specific utilisation patterns), and when the required datasets (including EEG) and analyst effort are feasible. The simplified method is justified when the primary goal is a transparent, segment-level proxy for comparing tasks/conditions or visualising within-task CL dynamics under resource constraints (e.g. no EEG acquisition, limited time for participant-specific operator modelling, or early-stage/pilot studies). The simplified method reduces effort by removing participant-specific parameterisation, but this comes at the cost of reduced sensitivity for claims that depend on individual variability (as supported by the weaker performance relationship for the weighted global formulation in Section 6.8).

Fourth, applying the method beyond conventional feature-based CAD systems with mouse and keyboard input would require adaptation in three places:

- CAD event taxonomy; redefining the observable CAD event set (e.g. gesture-based commands, VR interactions) and ensuring CAD logs can represent them consistently.
- Segmentation scheme; revising intermediate-level subtask types and unit task phase definitions to reflect different interaction structures (e.g. continuous 3D manipulation, bimanual interaction, gesture-driven navigation).
- Operator mapping and parameters; updating mapping heuristics (and potentially operator definitions) to account for different perceptual/motor demands and recalibrating information-processing weights derived from a different paradigm.

Fifth, clear definition and reporting of mapping rules are essential for reproducibility. In the current thesis, the segmentation scheme and mapping logic are specified as a rule-based procedure (with explicit definitions and heuristics in Chapters 4 and 5), and the second empirical study demonstrates end-to-end implementation. However, some judgement is still required in borderline cases (e.g. interpreting pauses, mixed-purpose actions, or overlapping activities), and the thesis does not fully eliminate analyst degrees of freedom. This should be viewed as a limitation of any heuristic mapping approach; strengthening reproducibility in future work would benefit from formal decision trees for ambiguous cases, reporting of exceptions, and, where feasible, inter-rater agreement checks on a representative subset.

Finally, the empirical evidence supporting the method is constrained by the design of empirical studies; the participant sample size and composition, the limited set of CAD tasks, the specific CAD system and experimental setup, and the practical limits of EEG hardware and associated EEG data quality in realistic CAD modelling. These factors bound generalisability and motivate replication across larger and more diverse samples, additional CAD task types (beyond CAD modelling a solid part), and alternative CAD environments.

This page intentionally left blank

8 CONCLUSIONS

The final chapter concludes the research activities conducted within the thesis, highlights the main findings and limitations, and proposes directions for future work.

This thesis addressed the lack of coherent theoretical formalisation of engineering designers' performance in CAD modelling and the limited understanding as well as measurement of cognitive load (CL) during CAD activities. CAD modelling was observed as a cognitively complex human–computer interaction (HCI) activity in which engineering designers allocate perceptual, cognitive, and motor resources to transform input design information into CAD representations. From this perspective, the thesis pursued three objectives: 1) to develop and validate a theoretical model of engineering designers' performance in CAD activities, 2) to develop and validate a theoretical model of CL in CAD activities, and 3) to develop and validate a method for measuring and analysing CL in CAD activities. The work further examined the hypothesis that psychophysiological monitoring enables assessment of CAD activity performance.

These aims were accomplished through the combination of an extensive literature review, two empirical studies, the prescription of theoretical models and a CL measurement method, and their subsequent validation against defined objectives. The first empirical study investigated whether the format of input engineering information influences CAD modelling performance and related cognitive processing. This study was intentionally positioned as an initial empirical step to preliminarily verify prerequisite assumptions needed for later model and method prescription: first, that a practically relevant and theoretically motivated influencing factor (representation format, operationalised through projection type) produces measurable differences in CAD performance outcomes and CAD modelling processes; second, that EEG recording during realistic CAD modelling is feasible and sufficiently sensitive to detect condition-related differences in brain activity before being integrated into a CL measurement and analysis method; and third, that human performance (cognitive) modelling is feasible and informative for representing CAD modelling as a structured process in which observable CAD actions can be mapped to perceptual, cognitive, and motor (P, C, M) operators to enable interpretable process-based analysis. By comparing modelling from isometric and orthographic projections, the study confirmed representation format as an influencing factor reflected in both performance outcomes and processes across multiple hierarchical goal/task levels. In parallel, EEG results indicated that engineering designers' brain activity during CAD modelling is sensitive to representation type, supporting the implication that design information format can modulate CL. The study also demonstrated the feasibility of recording EEG during realistic CAD tasks and provided an initial hierarchical structure of CAD tasks supported by timeline visualisations, thereby establishing a foundation for process-based analysis and subsequent theoretical modelling.

Building on insights from the literature and the first empirical study, the thesis developed two complementary theoretical models; of engineering designers' performance in CAD and CL in CAD activities. The model of engineering designers' performance formalised CAD modelling as a segmentable HCI process that links observable interaction with a CAD system to underlying perceptual, cognitive, and motor (P, C, M) mechanisms and supports systematic comparison across engineering designers and CAD tasks. The CL model is anchored in the model of CAD performance. It quantifies CL through P and C operator utilisations and explicit equations linking these utilisations to WL and CL. In addition, it introduces segment-specific weighting to represent differences in information-processing demands across unit task phases and intermediate subtasks. Together, these models provide an interpretable theoretical framework for monitoring changes in CL within the CAD modelling process and relating it to CAD performance outcomes.

Based on these models, a multimodal method for measuring and analysing CL in CAD was prescribed. The method integrates subjective CL assessment (NASA TLX), psychophysiological measurement (EEG), and utilisation-based human performance modelling. It specifies required data inputs and a coherent information flow to decompose CAD modelling into hierarchical task segments, map observed CAD behaviour to P, C, and M operators, compute utilisation-based indicators and dynamic CL estimates, and relate CL indicators to both process- and outcome-based CAD performance measures.

The second empirical study operationalised and tested the integrated method under varying internal CAD task complexity. It identified parietal and occipital alpha task-related power (TRP) as a prominent EEG-based candidate indicator for CL in the CAD context and confirmed internal task complexity as an additional influencing factor. The study demonstrated that utilisation-based indicators provide process-level discrimination across unit task phases and intermediate subtasks. However, dynamic alignment between EEG-derived and utilisation-derived CL patterns was limited under the tested segmentation and time windowing schemes. This indicates that EEG and utilisation capture overlapping but distinct aspects of CL. Importantly, regression analyses linking utilisation-based CL estimates to overall CAD modelling performance provided evidence supporting the research hypothesis. Across the tested indicators, higher estimated CL was consistently associated with lower performance, and Weighted Median U showed the clearest and most robust relationship, indicating added explanatory value when operator utilisations are combined with EEG-derived information-processing weights. Based on the results of the second empirical study, the thesis also proposed a simplified utilisation-based method variant to reduce barriers for adoption. This simplified method variant preserves the ability to track CL changes across the modelling sequence with reduced experimental data and analysis requirements.

Therefore, the central contribution of the thesis is a method for process-level analysis of CL in CAD modelling, supported by two mutually connected theoretical models. The thesis thus contributes three outcomes: 1) a validated theoretical model of engineering designers' CAD performance that combines an operational hierarchical task structure with explicit links to information-processing mechanisms; 2) a validated theoretical model of CL in CAD activities that formalises CL as a resource-based construct coupled to the CAD performance model; and 3) a validated multimodal method for measuring and analysing CL in CAD that supports dynamic, segment-level interpretation of CL in relation to CAD performance, including a simplified utilisation-based variant to support broader use in CAD research.

Taken together, these outcomes show that the thesis achieved its purpose by developing support for understanding and assessing engineering designers' performance in CAD activities through the modelling, measurement, and analysis of CL. This support is not yet intended as a direct prescriptive tool that automatically improves CAD performance, but as a theoretical and methodological foundation for making engineering designers' performance in CAD modelling more interpretable at the process level. The model of CAD performance defines how CAD modelling can be decomposed into hierarchical task segments and interpreted through perceptual, cognitive, and motor resource use. The CL model anchors CL assessment to this same process logic by linking CL to operator utilisation and segment-specific information-processing demands. The prescribed method then operationalises this theoretical framework by combining subjective assessment, EEG-based psychophysiological indicators, and utilisation-based modelling to examine how CL changes across CAD modelling process and how these changes relate to CAD performance.

The purpose was achieved within the defined scope of research; bounded CAD modelling activities performed in feature-based parametric CAD systems, primarily representative of embodiment and detail design. Within this scope, the thesis provides a coherent way to move beyond outcome-only assessment of CAD performance and towards a temporally structured interpretation of how engineering designers interact with CAD systems, allocate information-processing resources, and experience CL during task execution. At the same time, the achieved support should be interpreted as an evaluated research framework rather than a fully generalised industrial tool. Its transferability to broader engineering design activities, different CAD paradigms, collaborative industrial workflows, and early-stage design activities requires further development and validation.

Several limitations constrain the generalisability of the findings. First, the empirical evidence is strongest for bounded, individual CAD modelling tasks conducted in controlled laboratory settings and within feature-based parametric CAD systems. Therefore, the proposed models and the method should be interpreted as most directly applicable to CAD modelling activities of the kind investigated in this thesis, rather than to the engineering design process as a whole. Broader design activities, such as problem framing, concept generation, early-stage exploration, and iterative co-evolution of problem and solution, were not explicitly modelled. Second, the CAD performance model is closely aligned with explicit feature-based parametric modelling, where task execution can be decomposed into relatively discrete commands, operations, and modelling steps. Its transferability to direct, hybrid, or implicit modelling workflows remains to be demonstrated. In such workflows, interaction may be less command-sequential and more geometry-driven, exploratory, or continuous, which may require adapting the operator taxonomy, segmentation rules, and assumptions about how CAD behaviour is mapped to perceptual, cognitive, and motor operations. Third, the CL indicators and utilisation-based estimates should be interpreted with caution. EEG-based CL indicators showed participant- and context-sensitive behaviour, and segment-level intensity differences were difficult to establish consistently. Moreover, motor demands are partly shaped by the specific CAD system interface, command structure, and interaction devices, while cognitive demands in the empirical tasks may reflect spatial ability and task familiarity rather than broader design reasoning. These factors indicate that richer EEG feature sets, refined segment definitions, stronger empirical corroboration, and more explicit consideration of user expertise, spatial ability, learning effects, CAD familiarity, and design context are needed. Finally, while the proposed method enables CL-informed performance assessment, it does not directly prescribe how to improve CAD performance; rather, it provides the modelling and measurement foundation required for future prescriptive studies addressing training, task design, CAD system interface design, and intelligent CAD system support.

Future research can strengthen and extend the prescribed method in five main directions. First, broader empirical validation is needed across a wider range of CAD tasks, CAD systems, and modelling paradigms. This includes additional feature-based CAD environments, different modelling approaches (e.g. direct, hybrid, or implicit modelling), different types of CAD models (e.g. assemblies), broader range of design activities, and industrially situated tasks. Such studies would clarify which elements of the method are general across CAD activities and which need to be adapted to specific modelling paradigms, interfaces, and design contexts. Second, the mapping heuristics used to assign P, C, and M operators could be refined and standardised through more explicit and reproducible rule sets, evaluation of interrater agreement, and systematic testing of sensitivity to segmentation and mapping choices. This would strengthen the methodological robustness of the CAD performance and CL models and support their use by other researchers. Third, the scalability and practical implementation of the method should be improved. The current multimodal method, particularly when combined with EEG recording and detailed manual task decomposition, is resource intensive. Future work should therefore explore automation of key method inputs by integrating CAD logs, macros, interaction traces, or event-based data extraction. Such automation could reduce manual effort, enable larger samples and more task variant, and make the method more accessible for both academic and industrial use. Fourth, the CL modelling component could be strengthened by calibration and testing of alternative weighting strategies, including participant-specific weighting, robust normalisation procedures, combining multiple EEG-based indicators, and the integration of additional physiological measures such as eye tracking. These extensions could improve the stability of information-processing weights and CL estimates across individuals, and help distinguish between cognitive demands arising from design reasoning, spatial processing, interface interaction, and disruptions related to CAD systems. Fifth, future research should investigate how the method can inform practical interventions in CAD system design, engineering education, and work procedure design. For CAD development, the findings may support user-centred interface improvements, adaptive information prompts, better management of command sequences, reduced cognitive disruption from system response times, and peripheral or multimodal feedback mechanisms. For engineering education, the method may help identify modelling phases, subtasks, or interaction patterns that impose high CL and could therefore benefit from targeted training, scaffolding, or redesigned task inputs. Extending the models and the method towards earlier stages of engineering design, such as conceptual design, problem formulation, and solution space exploration, would require expanding the goal structures, operations, operator set, and action classes, and integrating additional measures to represent higher-level design cognition beyond CAD system interaction.

REFERENCES

[1] D. G. Ullman, "Toward the ideal mechanical engineering design support system," *Res. Eng. Des. - Theory, Appl. Concurr. Eng.*, vol. 13, no. 2, pp. 55–64, 2002, doi: 10.1007/s00163-001-0007-4.

[2] C. McMahon, "Design Informatics: Supporting Engineering Design Processes with Information Technology," *J. Indian Inst. Sci.*, vol. 95, no. 4, pp. 365–377, 2015.

[3] B. F. Robertson and D. F. Radcliffe, "Computer-Aided Design Impact of CAD tools on creative problem solving in engineering design," *Comput. Des.*, vol. 41, no. 3, pp. 136–146, 2009, doi: 10.1016/j.cad.2008.06.007.

[4] D. Robertson and T. J. Allen, "CAD System Use and Engineering Performance," *IEEE Trans. Eng. Manag.*, vol. 40, no. 3, pp. 274–282, 1993, doi: 10.1109/17.233189.

[5] D. Robertson, K. Ulrich, and M. Filerman, "Cognitive complexity and CAD systems: Beyond the drafting board metaphor," Cambridge, Massachusetts, WP #3244-91-MSA, 1991.

[6] D. G. Ullman, T. G. Dietterich, and L. A. Stauffer, "A model of the mechanical design process based on empirical data," *AI EDAM*, vol. 2, no. 1, pp. 33–52, 1988.

[7] S. K. Card, P. Moran, Thomas, and A. Newell, *The psychology of human-computer interaction*. Lawrance Erlbaum Associates, Inc., 1983.

[8] V. Hubka, *Principles of engineering design*, vol. 46, no. 536. Butterworth Scientific, 1980.

[9] S. Lee and J. Yan, "The impact of 3D CAD interfaces on user ideation: A comparative analysis using SketchUp and Silhouette Modeler," *Des. Stud.*, vol. 44, pp. 52–73, 2016, doi: 10.1016/j.destud.2016.02.001.

[10] A. Dillon and M. Sweeney, "The application of cognitive psychology to CAD," *People Comput. IV*, no. January 1988, pp. 477–488, 1988.

[11] J. J. Shah, J. Woodward, and S. M. Smith, "Applied tests of design skills-part II: visual thinking," *J. Mech. Des.*, vol. 135, no. 7, p. 071004, 2013, doi: 10.1115/1.4024228.

[12] P. Rosso, J. Gopsill, S. Burgess, and B. Hicks, "Investigating and characterising variability in CAD modelling and its potential impact on editability: An exploratory study," *Comput. Aided. Des. Appl.*, vol. 18, no. 6, pp. 1306–1326, 2021, doi: 10.14733/cadaps.2021.1306-1326.

[13] J. Sweller, P. Ayres, and S. Kalyuga, *Cognitive Load Theory*. Springer, 2011.

[14] D. Kahneman, *Attention and Effort*. PRENTICE-HALL INC., 1973.

[15] G. Sun and S. Yao, "Investigating the relation between cognitive load and creativity in the conceptual design process," *Proc. Hum. Factors Ergon. Soc.*, pp. 308–312, 2012, doi: 10.1177/1071181312561072.

[16] A. Mohamed-Ahmed, N. Bonnardel, P. Côté, and S. Tremblay, "Cognitive load management and architectural design outcomes," *Int. J. Des. Creat. Innov.*, vol. 1, no. 3, pp. 160–176, 2013, doi: 10.1080/21650349.2013.797013.

[17] C. Zimmerer and S. Matthiesen, "Study on the impact of cognitive load on performance in engineering design," *Proc. Des. Soc.*, vol. 1, no. AUGUST, pp. 2761–2770, 2021, doi:

10.1017/pds.2021.537.

[18] N. Calpin and J. Menold, "The Cognitive Costs of Design Tasks: Examining Cognitive Load Through Verbal and Physical Indicators," *J. Mech. Des.*, vol. 145, no. 10, 2023, doi: 10.1115/1.4062976.

[19] L. Blessing and A. Chakrabarti, *DRM: A Design Reseach Methodology*, no. September. 2009.

[20] P. Cash, T. Stanković, and M. Štorga, *Experimental Design Research: Approaches, Perspectives, Applications*. Springer Nature, 2016.

[21] C. C. Seepersad, K. Pedersen, J. Emblemsvåg, R. Bailey, and J. K. Allen, "The Validation Square: How Does One Verify and Validate a Design Method?," *Decis. Mak. Eng. Des.*, pp. 303–313, 2010, doi: 10.1115/1.802469.ch25.

[22] K. Pedersen, A. Solutions, J. Allen, and F. Mistree, "The''Validation Square''-Validating Design Methods," no. January, 2000.

[23] C. Wu, "The Five Key Questions of Human Performance Modeling," *Int J Ind Ergon.*, vol. 63, pp. 3–6, 2018, doi: 10.1016/j.ergon.2016.05.007.

[24] P. Thorvald, J. Lindblom, and R. Andreasson, "On the development of a method for cognitive load assessment in manufacturing," *Robot. Comput. Integr. Manuf.*, vol. 59, no. May, pp. 252–266, 2019, doi: 10.1016/j.rcim.2019.04.012.

[25] A. Gevins, M. E. Smith, and H. Leong, "Monitoring working memory load during computer-based tasks with EEG pattern recognition methods," *Hum. Factors*, vol. 40, no. 1, pp. 79–91, 1998.

[26] A. Blandford and T. Green, "Methodological Development," in *Research Methods for Human Computer Interaction*, 2008, pp. 158–174.

[27] F. J. O'Donnell and A. H. B. Duffy, *Design Performance*. Springer London, 2005.

[28] Y. Ishino and Y. A. N. Jin, "Acquiring engineering knowledge from design processes," *Artif. Intell. Eng. Des. Anal. Manuf.*, vol. 16, pp. 73–91, 2002.

[29] S. Sadeghi, T. Dargon, L. Rivest, and J.-P. Pernot, "Capturing and analyzing how designers use CAD software," 2016.

[30] S. Bhavnani, J. Garrett Jr, and D. Shaw, "Leading Indicators of CAD Experience," *5th Int. Conf. Comput. Archit. Des. Futur.*, no. 1990, pp. 313–334, 1993.

[31] S. K. Bhavnani, B. E. John, and U. Flemming, "The strategic use of CAD: An empirically inspired, theory-based course," *Conf. Hum. Factors Comput. Syst. - Proc.*, no. May, pp. 183–190, 1999, doi: 10.1145/302979.303036.

[32] I. Chester, "Teaching for CAD expertise," *Int. J. Technol. Des. Educ.*, vol. 17, no. 1, pp. 23–35, 2007, doi: 10.1007/s10798-006-9015-z.

[33] C. J. Atman, K. Deibel, and J. Borgford-Parnell, "The process of engineering design: A comparison of three representations," in *International Conference on Engineering Design, ICED'09*, 2009, pp. 483–494.

[34] R. F. Hamade and H. A. Artail, "A study of the influence of technical attributes of beginner CAD users on their performance," *CAD Comput. Aided Des.*, vol. 40, no. 2, pp. 262–272, 2008, doi: 10.1016/j.cad.2007.11.001.

[35] A. Rynne and W. Gaughran, “Cognitive modeling strategies for optimum design intent in parametric modeling (PM),” *Comput. Educ. J.*, vol. 18, no. 3, pp. 55–68, 2008, doi: 10.18260/1-2--2651.

[36] R. F. Hamade and H. A. Artail, “A study of the influence of learning style of beginner computer-aided design users on their performance,” *J. Eng. Des.*, vol. 21, no. 5, pp. 561–577, 2010, doi: 10.1080/09544820802409289.

[37] S. T. Pektas, “Effects of cognitive styles on 2D drafting and design performance in digital media,” pp. 63–76, 2010, doi: 10.1007/s10798-008-9060-x.

[38] A. Rynne, W. F. Gaughran, and N. Seery, “Defining the variables that contribute to developing 3D CAD modelling expertise,” 2010.

[39] T. J. Branoff and M. Dobelis, “The relationship between spatial visualization ability and students’ ability to model 3D objects from engineering assembly drawings,” *Eng. Des. Graph. J.*, vol. 76, no. 3, pp. 106–111, 2012.

[40] R. P. Diwakaran and M. D. Johnson, “Analyzing the effect of alternative goals and model attributes on CAD model creation and alteration,” *Comput. Des.*, vol. 44, no. 4, pp. 343–353, Apr. 2012, doi: 10.1016/J.CAD.2011.11.003.

[41] X. Peng, P. Mcgary, M. Johnson, B. Yalvac, and E. Ozturk, “Assessing Novice CAD Model Creation and Alteration,” *Comput. aided Des. Appl.*, no. 2, pp. 9–19, 2012, doi: 10.3722/cadaps.2012.PACE.9-19.

[42] H. M. Steinhauer, “Correlation between a student’s performance on the mental cutting test and their 3D parametric modeling ability,” *Eng. Des. Graph. J.*, vol. 76, no. 3, pp. 44–48, 2012.

[43] C. Xie, Z. Zhang, S. Nourian, A. Pallant, and E. Hazzard, “Time series analysis method for assessing engineering design processes using a CAD tool,” *Int. J. Eng. Educ.*, vol. 30, no. 1, pp. 218–230, 2014.

[44] P. Company, M. Contero, J. Otey, and R. Plumed, “Approach for developing coordinated rubrics to convey quality criteria in MCAD training,” *CAD Comput. Aided Des.*, vol. 63, pp. 101–117, 2015, doi: 10.1016/j.cad.2014.10.001.

[45] J. Gopsill, C. Snider, L. Shi, and B. Hicks, “Computer aided design user interaction as a sensor for monitoring engineers and the engineering design process,” *Proc. Int. Des. Conf. Des.*, vol. DS 84, pp. 1707–1718, 2016.

[46] J. Sweany, P. Goodrum, and J. Miller, “Analysis of empirical data on the effects of the format of engineering deliverables on craft performance,” *Autom. Constr.*, vol. 69, pp. 59–67, 2016, doi: 10.1016/j.autcon.2016.05.017.

[47] B. Stone *et al.*, “A Multi-User Computer-Aided Design Competition : Experimental Findings and Analysis of Team-Member Dynamics,” vol. 17, no. September, pp. 1–10, 2017, doi: 10.1115/1.4035674.

[48] C. McComb, J. Cagan, and K. Kotovsky, “Mining Process Heuristics from Designer Action Data Via Hidden Markov Models,” *J. Mech. Des. Trans. ASME*, vol. 139, no. 11, pp. 1–12, 2017, doi: 10.1115/1.4037308.

[49] C. McComb, J. Cagan, and K. Kotovsky, “Capturing Human Sequence-Learning Abilities in Configuration Design Tasks Through Markov Chains,” *J. Mech. Des. Trans. ASME*, vol. 139, no. 9, pp. 1–12, 2017, doi: 10.1115/1.4037185.

[50] J. Buckley, N. Seery, and D. Canty, “Heuristics and CAD modelling: an examination of student behaviour during problem solving episodes within CAD modelling activities,” *Int. J. Technol. Des. Educ.*, vol. 28, no. 4, pp. 939–956, 2018, doi: 10.1007/s10798-017-9423-2.

[51] K. Eves, J. Salmon, J. Olsen, and F. Fagergren, “A comparative analysis of computer-aided design team performance with collaboration software,” *Comput. Aided. Des. Appl.*, vol. 15, no. 4, pp. 476–487, 2018, doi: 10.1080/16864360.2017.1419649.

[52] X. Garikano, M. Garmendia, A. P. Manso, and E. Solaberrieta, “Strategic knowledge-based approach for CAD modelling learning,” *Int. J. Technol. Des. Educ.*, vol. 29, no. 4, pp. 947–959, 2019, doi: 10.1007/s10798-018-9472-1.

[53] J. Otey, P. Company, and M. Contero, “Assessment of parametric assembly models based on CAD quality dimensions,” *Comput. Des. Appl.*, vol. 16, pp. 628–653, 2019.

[54] H. Arshad, V. Phadnis, and A. Olechowski, “Paired computer-aided design: The effect of collaboration mode on differences in model quality,” in *Proceedings of the ASME 2020 DETC2020-22730*, 2020, pp. 1–9.

[55] T. Delahunty, N. Seery, R. Dunbar, and M. Ryan, “An exploration of the variables contributing to graphical education students’ CAD modelling capability,” *Int. J. Technol. Des. Educ.*, vol. 30, no. 2, pp. 389–411, 2020, doi: 10.1007/s10798-019-09503-x.

[56] K. Kitajima, K. Oikawa, and T. Murakami, “Study on Feature Analysis of CAD Operation Log and Its Application,” in *The 5th International Conference od Design Engineering and Science*, 2020, no. November 2020.

[57] H. E. Otto and F. Mandorli, “Parametric feature-based solid model deficiency identification to support learning outcomes assessment in CAD education,” *Comput. Aided. Des. Appl.*, vol. 18, no. 2, pp. 411–442, 2020, doi: 10.14733/cadaps.2021.411-442.

[58] J. Y.-H. Chen, “Development of a Novel Computer-Aided Design Experiment Protocol for Studying Designer Behaviours,” University of Toronto, 2021.

[59] K. Leonardo and A. Olechowski, “Identifying computer-aided design action types from professional user analytics data,” *Proc. ASME Des. Eng. Tech. Conf.*, vol. 3B-2021, pp. 1–10, 2021, doi: 10.1115/DETC2021-72102.

[60] V. Phadnis, H. Arshad, D. Wallace, and A. Olechowski, “Are two heads better than one for computer-aided design?,” *J. Mech. Des. Trans. ASME*, vol. 143, no. 7, 2021, doi: 10.1115/1.4050734.

[61] A. Aranburu, J. Cotillas, D. Justel, M. Contero, and J. D. Camba, “How Does the Modeling Strategy Influence Design Optimization and the Automatic Generation of Parametric Geometry Variations?,” *Comput. Des.*, vol. 151, p. 103364, Oct. 2022, doi: 10.1016/J.CAD.2022.103364.

[62] Y. Deng, M. Mueller, C. Rogers, and A. Olechowski, “The multi-user computer-aided design collaborative learning framework,” *Adv. Eng. Informatics*, vol. 51, p. 101446, 2022, doi: 10.1016/j.aei.2021.101446.

[63] R. Celjak, N. Horvat, and S. Škec, “Comparing Collaborative Cad Modelling Patterns of High-Performing and Low-Performing Teams,” *Proc. Des. Soc.*, vol. 3, no. JULY, pp. 1007–1016, 2023, doi: 10.1017/pds.2023.101.

[64] N. Horvat, J. Šklebar, M. Štorga, and S. Škec, “Create or revise ? A comparative study on CAD rework after team-based engineering design review in virtual reality and desktop

interface," *Adv. Eng. Informatics*, vol. 65, 2025, doi: https://doi.org/10.1016/j.aei.2025.103177.

[65] J. Šklebar, T. Martinec, and M. Štorga, "Evaluating ChatGPT ' s role in collaborative CAD task completeness," vol. 5, pp. 1585–1594, 2025.

[66] R. Guay, "Purdue Spatial Visualization Test." 1976.

[67] P. Rosso, J. Gopsil, B. Hicks, and S. Burgess, "Investigating and characterising variability in CAD modelling: An overview," in *Proceedings of CAD'20*, 2020, pp. 226–230, doi: 10.14733/cadconfp.2020.226-230.

[68] M. M. Andreasen, "Modelling—The Language of the Designer," *J. Eng. Des.*, vol. 5, no. 2, pp. 103–115, 1994, doi: 10.1080/09544829408907876.

[69] C. L. Dym and D. C. Brown, *Engineering design: Representation and reasoning*, Second Edi. Cambridge University Press, 2011.

[70] J. J. Shah and M. T. Rogers, "Assembly modeling as an extension of feature-based design," *Res. Eng. Des.*, vol. 5, no. 3–4, pp. 218–237, 1993, doi: 10.1007/BF01608364.

[71] T. J. Branoff, K. L. Devine, and J. Brown, "Evaluating a Rubric for Assessing Constraint-Based Solid Models Evaluating a Rubric for Assessing Constraint-Based," 2016.

[72] A. P. Garland and S. J. Grigg, "Evaluation of humans and software for grading in an engineering 3D CAD course," *ASEE Annu. Conf. Expo. Conf. Proc.*, 2019, doi: 10.18260/1-2--32764.

[73] S. J. Kirstukas, "Development and evaluation of a computer program to assess student CAD models," *ASEE Annu. Conf. Expo. Conf. Proc.*, vol. 2016-June, 2016, doi: 10.18260/p.26781.

[74] K. Jaakma and P. Kiviluoma, "Auto-assessment tools for mechanical computer aided design education," *Heliyon*, vol. 5, no. October, p. e02622, 2019, doi: 10.1016/j.heliyon.2019.e02622.

[75] P. Company, F. Naya, M. Contero, and J. D. Camba, "On the role of geometric constraints to support design intent communication and model reusability," *Comput. Aided. Des. Appl.*, vol. 17, no. 1, pp. 61–76, 2020, doi: 10.14733/cadaps.2020.61-76.

[76] M. H. Rahman, C. Schimpf, C. Xie, and Z. Sha, "A computer-aided design based research platform for design thinking studies," *J. Mech. Des.*, vol. 141, no. 12, pp. 1–12, 2019, doi: 10.1115/1.4044395.

[77] V. S. Phadnis, D. R. Wallace, and A. Olechowski, "A multimodal experimental approach to study CAD collaboration," *Comput. Aided. Des. Appl.*, vol. 18, no. 2, pp. 328–342, 2021, doi: 10.14733/cadaps.2021.328-342.

[78] A. Vrolijk and A. Olechowski, "Computer-Aided Design as a Research Instrument for Engineering Design," *J. Eng. Educ.*, 2024.

[79] M. D. Byrne and R. W. Pew, "A History and Primer of Human Performance Modeling," *Rev. Hum. Factors Ergon.*, vol. 5, no. 1, pp. 225–263, 2009, doi: 10.1518/155723409x448071.

[80] J. R. Anderson and C. Lebiere, *The atomic components of thought*. Psychology Press Taylor & Francis Group, 1998.

[81] D. G. Ullman, S. Wood, and D. Craig, "The importance of drawing in the mechanical design process," *Comput. Graph.*, vol. 14, no. 2, pp. 263–274, Jan. 1990, doi: 10.1016/0097-

8493(90)90037-X.

[82] M. D. Byrne, “Applying a Cognitive Architecture to HCI,” 2001.

[83] G. Salvendy and W. Karwowski, *Handbook of human factors and ergonomics*, vol. 6, no. August. 2021.

[84] B. E. John and D. E. Kieras, “The GOMS family of analysis techniques: Tools for design and evaluation,” 1994. doi: 10.1201/9781482276275-15.

[85] G. A. Miller, “The magical number seven, plus or minus two: Some limits on our capacity for processing information,” *Psychol. Rev.*, vol. 63, no. 2, pp. 81–97, 1956, doi: 10.1177/001088049003100202.

[86] P. Chandler and J. Sweller, “Cognitive load while learning to use a computer program,” *Appl. Cogn. Psychol.*, vol. 10, no. 2, pp. 151–170, 1996, doi: 10.1002/(SICI)1099-0720(199604)10:2<151::AID-ACP380>3.0.CO;2-U.

[87] J. Sweller, “Element interactivity and intrinsic, extraneous, and germane cognitive load,” *Educ. Psychol. Rev.*, vol. 22, no. 2, pp. 123–138, 2010, doi: 10.1007/s10648-010-9128-5.

[88] A. Baddeley, “Working memory: Looking back and looking forward,” *Nat. Rev. Neurosci.*, vol. 4, no. 10, pp. 829–839, 2003, doi: 10.1038/nrn1201.

[89] Y. Liu, R. Feyen, and O. Tsimhoni, “Queueing Network-Model Human Processor (QN-MHP): A computational architecture for multitask performance in human-machine systems,” *ACM Trans. Comput. Interact.*, vol. 13, no. 1, pp. 37–70, 2006, doi: 10.1145/1143518.1143520.

[90] D. E. Kieras, “Towards a Practical GOMS Model Methodology for User Interface Design,” *Handb. Human-Computer Interact.*, no. January 1988, pp. 135–157, 1988, doi: 10.1016/b978-0-444-70536-5.50012-9.

[91] B. E. John, “Contributions to engineering models of human-computer interaction: Volume I,” 1988.

[92] S. K. Card, T. P. Moran, and A. Newell, “The Keystroke-Level Model for User Performance Time with Interactive Systems,” *Commun. ACM*, vol. 23, no. 7, pp. 396–410, 1980, doi: 10.1145/358886.358895.

[93] W. D. Gray, B. E. John, and M. E. Atwood, “Precis of project Ernestine or an overview of a validation of GOMS,” *Conf. Hum. Factors Comput. Syst. - Proc.*, pp. 307–312, 1992, doi: 10.1145/142750.142821.

[94] B. John, A. Vera, M. Matessa, M. Freed, and R. Remington, “Automating CPM-GOMS,” no. February, pp. 147–154, 2002, doi: 10.1145/503376.503404.

[95] J. D. Baskin and B. E. John, “Comparison of GOMS Analysis Methods,” *Conf. Hum. Factors Comput. Syst. - Proc.*, no. April, pp. 261–262, 1998, doi: 10.1145/286498.286742.

[96] R. Gong and D. Kieras, “Validation of the GOMS model methodology in the development of a specialized, commercial software application,” *Conf. Hum. Factors Comput. Syst. - Proc.*, pp. 351–357, 1994, doi: 10.1145/191666.191782.

[97] G. T. Lang, R. E. Eberts, M. M. Barash, and M. G. Gabel, “Extracting and Using Procedural Knowledge in a CAD Task,” *IEEE Trans. Eng. Manag.*, vol. 38, no. 3, pp. 257–268, 1991, doi: 10.1109/17.83758.

[98] S. K. Bhavnani and B. E. John, “The Strategic Use of Complex Computer Systems,” vol. 15,

no. 2000, pp. 107–137, 2000, doi: 10.1207/S15327051HCI1523.

[99] S. K. Bhavnani and B. E. John, "Delegation and circumvention: Two Faces of Efficiency," in *CHI 98 Los Angeles USA*, 1998, no. April, pp. 273–280, doi: 10.1145/274644.274683.

[100] S. K. Bhavnani and B. E. John, "The Efficient Use of Complex Computer Systems," vol. 15, no. 2, pp. 97–124, 1996.

[101] F. Paas, J. E. Tuovinen, H. Tabbers, and P. W. M. Van Gerven, "Cognitive load measurement as a means to advance cognitive load theory," *Educ. Psychol.*, vol. 38, no. 1, pp. 63–71, 2003, doi: 10.1207/S15326985EP3801_8.

[102] B. Xie and G. Salvendy, "Prediction of Mental Workload in Single and Multiple Tasks Environments," *Int. J. Cogn. Ergon.*, no. May 2015, pp. 37–41, 2010, doi: 10.1207/S15327566IJCE0403.

[103] B. Raufi and L. Longo, "An Evaluation of the EEG Alpha-to-Theta and Theta-to-Alpha Band Ratios as Indexes of Mental Workload," *Front. Neuroinform.*, vol. 16, no. March, 2022, doi: 10.3389/fninf.2022.861967.

[104] G. Johannsen, "Workload and workload measurement," in *Mental Workload - Its Theory and Measurement*, N. E. Moray, Ed. New York: Plenum Press, 1979.

[105] P. M. Linton, B. D. Plamondon, A. O. Dick, A. C. Bittner, and R. E. Christ, "Operator workload for military system acquisition," in *Application of human performanc emodels to system design*, vol. 2, G. R. McMillan, D. Beevis, E. Salas, M. H. Strub, R. Sutton, and L. van Breda, Eds. New York: Plenum Press, 1989.

[106] P. A. Hancock, G. Matthews, and C. Florida, "Workload and Performance : Associations , Insensitivities , and Dissociations," *Hum. Factors*, no. January 2019, 2018, doi: 10.1177/0018720818809590.

[107] F. G. W. C. Paas and J. J. G. Van Merriënboer, "Instructional control of cognitive load in the training of complex cognitive tasks," *Educ. Psychol. Rev.*, vol. 6, no. 4, pp. 351–371, 1994, doi: 10.1007/BF02213420.

[108] J. Elkind, S. Card, J. Hochberg, and B. Huey, *Human Performance Models for Computer-Aided Engineering*. 1990.

[109] M. D. Johnson, L. M. Valverde, and W. D. Thomison, "An investigation and evaluation of computer-aided design model complexity metrics," *Comput. Aided. Des. Appl.*, vol. 15, no. 1, pp. 61–75, 2018, doi: 10.1080/16864360.2017.1353729.

[110] J. C. Lavrsen and J. Daalhuizen, "Balancing cognitive load in design work: A conceptual and narrative review," in *DRS2024*, 2024, pp. 23–28.

[111] M. Zhao, W. Jia, D. Yang, P. Nguyen, T. A. Nguyen, and Y. Zeng, "A tEEG Framework for Studying Designer ' s Cognitive and Affective States," *Des. Sci.*, vol. 6, pp. 1–47, 2020, doi: 10.1017/dsj.2020.28.

[112] S. Jo, R. Myung, and D. Yoon, "Quantitative prediction of mental workload with the ACT-R cognitive architecture," *Int. J. Ind. Ergon.*, vol. 42, no. 4, pp. 359–370, 2012, doi: 10.1016/j.ergon.2012.03.004.

[113] C. D. Wickens, "Processing Resources in Attention, Dual Task Performance, and Workload Assessment," 1991.

[114] M. A. Just, P. A. Carpenter, and A. Miyake, "Neuroindices of cognitive workload:

Neuroimaging, pupillometric and event-related potential studies of brain work," *Theor. Issues Ergon. Sci.*, vol. 4, no. 1–2, pp. 59–88, 2003, doi: 10.1080/14639220210159735.

[115] K. C. Hendy, J. Liao, and P. Milgram, "Combining time and intensity effects in assessing operator information- processing load," *Hum. Factors*, vol. 39, no. 1, pp. 30–47, 1997, doi: 10.1518/001872097778940597.

[116] C. Wu and Y. Liu, "Queuing network modeling of driver workload and performance," *IEEE Trans. Intell. Transp. Syst.*, vol. 8, no. 3, pp. 528–537, 2007, doi: 10.1109/TITS.2007.903443.

[117] H. Oh, Y. Yun, and R. Myung, "Driver behavior and mental workload for takeover safety in automated driving: ACT-R prediction modeling approach," *Traffic Inj. Prev.*, vol. 25, no. 3, pp. 381–389, 2024, doi: 10.1080/15389588.2023.2300640.

[118] S. Park, S. Jeong, and R. Myung, "Modeling of multiple sources of workload and time pressure effect with ACT-R," *Int. J. Ind. Ergon.*, vol. 63, pp. 37–48, 2018, doi: 10.1016/j.ergon.2017.07.003.

[119] S. Park and R. Myung, "Predicting task-related properties of mental workload with ACT-R cognitive architecture," *Proc. Hum. Factors Ergon. Soc.*, pp. 773–777, 2013, doi: 10.1177/1541931213571169.

[120] S. Cao and Y. Liu, "Mental workload modeling in an integrated cognitive architecture," *Proc. Hum. Factors Ergon. Soc.*, pp. 2083–2087, 2011, doi: 10.1177/1071181311551434.

[121] S. Cao and Y. Liu, "Modelling workload in cognitive and concurrent tasks with time stress using an integrated cognitive architecture," *Int. J. Hum. Factors Model. Simul.*, vol. 5, no. 2, p. 113, 2015, doi: 10.1504/ijhfms.2015.075360.

[122] S. Martin-Michiellot and P. Mendelsohn, "Cognitive load while learning with a graphical computer interface," *J. Comput. Assist. Learn.*, vol. 16, no. 4, pp. 284–293, 2000, doi: 10.1046/j.1365-2729.2000.00141.x.

[123] Z. Bilda, J. S. Gero, and T. Purcell, "To sketch or not to sketch? That is the question," *Des. Stud.*, vol. 27, no. 5, pp. 587–613, 2006, doi: 10.1016/j.destud.2006.02.002.

[124] G. B. Dadi, P. M. Goodrum, T. R. B. Taylor, and C. M. Carswell, "Cognitive Workload Demands Using 2D and 3D Spatial Engineering Information Formats," *J. Constr. Eng. Manag.*, vol. 140, no. 5, pp. 1–8, 2014, doi: 10.1061/(asce)co.1943-7862.0000827.

[125] A. M. Maier, N. Baltsen, H. Christoffersen, and H. Störrle, "Towards Diagram Understanding: A Pilot-Study Measuring Cognitive Workload Through Eye-Tracking," *Proc. Intl. Conf. Hum. Behav. Des.*, no. October, pp. 1–6, 2014.

[126] T. A. Nguyen and Y. Zeng, "A physiological study of relationship between designer's mental effort and mental stress during conceptual design," *CAD Comput. Aided Des.*, vol. 54, pp. 3–18, 2014, doi: 10.1016/j.cad.2013.10.002.

[127] T. Chandrasekera and S. Y. Yoon, "The Effect of Tangible User Interfaces on Cognitive Load in the Creative Design Process," *Proc. 2015 IEEE Int. Symp. Mix. Augment. Real. - Media, Art, Soc. Sci. Humanit. Des. ISMAR-MASH'D 2015*, no. February 2020, pp. 6–8, 2015, doi: 10.1109/ISMAR-MASHD.2015.18.

[128] B. G. Young, A. Wodehouse, and M. Sheridan, "Physical Interaction Mappings : Utilizing Cognitive Load Theory in Order To Enhance Physical Product Interaction," *Proc. 20th Int. Conf. Eng. Des. (ICED 15), Vol. 1 Des. Life*, no. July, pp. 1–10, 2015.

[129] T. A. Nguyen and Y. Zeng, "Effects of stress and effort on self-rated reports in experimental

study of design activities," *J. Intell. Manuf.*, vol. 28, no. 7, pp. 1609–1622, 2016, doi: 10.1007/s10845-016-1196-z.

[130] M. Schultze-Kraft, S. Dahne, M. Gugler, G. Curio, and B. Blankertz, "Unsupervised classification of operator workload from brain signals," *J. Neural Eng.*, vol. 13, no. 3, pp. 1–15, 2016, doi: 10.1088/1741-2560/13/3/036008.

[131] R. Yu and J. Gero, "Exploring designers' cognitive load when viewing different digital representations of spaces: A pilot study," *Smart Innov. Syst. Technol.*, vol. 65, pp. 457–467, 2017, doi: 10.1007/978-981-10-3518-0_40.

[132] B. Majdic, C. Cowan, J. Girdner, W. Opoku, O. Pierrakos, and E. Barrella, "Monitoring Brain Waves in an Effort to Investigate Student ' s Cognitive Load During a Variety of Problem Solving Scenarios," in *Systems and Information Engineering Design Symposium (SIEDS)*, 2017, pp. 186–191.

[133] A. Dan and M. Reiner, "EEG-based cognitive load of processing events in 3D virtual worlds is lower than processing events in 2D displays," *Int. J. Psychophysiol.*, vol. 122, pp. 75–84, 2017, doi: 10.1016/j.ijpsycho.2016.08.013.

[134] M. K. Choi, S. M. Lee, J. S. Ha, and P. H. Seong, "Development of an EEG-based workload measurement method in nuclear power plants," *Ann. Nucl. Energy*, vol. 111, pp. 595–607, 2018, doi: 10.1016/j.anucene.2017.08.032.

[135] P. Nguyen, T. A. Nguyen, and Y. Zeng, "Empirical approaches to quantifying effort, fatigue and concentration in the conceptual design process: An EEG study," *Res. Eng. Des.*, vol. 29, no. 3, pp. 393–409, 2018, doi: 10.1007/s00163-017-0273-4.

[136] T. Kosch, M. Funk, A. Schmidt, and L. L. Chuang, "Identifying cognitive assistance with mobile electroencephalography: A case study with in-situ projections for manual assembly," *Proc. ACM Human-Computer Interact.*, vol. 2, no. EICS, 2018, doi: 10.1145/3229093.

[137] M. Z. Baig and M. Kavakli, "Analyzing novice and expert user's cognitive load in using a multi-modal interface system," *26th Int. Conf. Syst. Eng. ICSEng 2018 - Proc.*, 2019, doi: 10.1109/ICSENG.2018.8638206.

[138] E. M. Barrella, C. Cowan, J. Girdner, M. K. Watson, and R. Anderson, "Measuring connections: Engineering students' cognitive activities and performance on complex tasks," in *Proceedings - Frontiers in Education Conference, FIE*, 2019, vol. 2019-Octob, doi: 10.1109/FIE43999.2019.9028595.

[139] T. Shealy, J. Gero, M. Hu, and J. Milovanovic, "Concept generation techniques change patterns of brain activation during engineering design," *Des. Sci.*, pp. 1–27, 2020, doi: 10.1017/dsj.2020.30.

[140] H. Nolte and C. McComb, "The cognitive experience of engineering design: an examination of first-year student stress across principal activities of the engineering design process," *Des. Sci.*, vol. 7, no. May, 2021, doi: 10.1017/dsj.2020.32.

[141] W. Jia, F. von Wegner, M. Zhao, and Y. Zeng, "Network oscillations imply the highest cognitive workload and lowest cognitive control during idea generation in open-ended creation tasks," *Sci. Rep.*, vol. 11, no. 1, pp. 1–24, 2021, doi: 10.1038/s41598-021-03577-1.

[142] J. Mathur, S. R. Miller, T. W. Simpson, and N. A. Meisel, "Analysis of the knowledge gain and cognitive load experienced due to the computer-aided instruction of additive

manufacturing processes," 2021.

[143] T. Dissaux and S. Jancart, "Impact of cognitive load associated with learning and using parametric tools in architectural design," 2022, doi: 10.31428/10317/11415.

[144] C. Zimmerer, T. Nelius, and S. Matthiesen, "Using eye tracking to measure cognitive load of designers in situ," *Des. Comput. Cogn.*, pp. 481–495, 2023, doi: 10.1007/978-3-031-20418-0_29.

[145] M. Cass and R. Prabhu, "Looking Beyond Self-Reported Cognitive Load: Investigating the Use of Eye Tracking in the Study of Design Representations in Engineering Design," *Proc. Des. Soc.*, vol. 3, no. JULY, pp. 2475–2484, 2023, doi: 10.1017/pds.2023.248.

[146] D. G. Ullman, *The Mechanical Design Process*, Fourth Edi. 2010.

[147] J. Gero and J. Milovanovic, "A framework for studying design thinking through measuring designers' minds, bodies and brains," *Des. Sci.*, vol. 6, pp. 1–40, 2020, doi: 10.1017/dsj.2020.15.

[148] P. Antonenko, F. Paas, R. Grabner, and T. van Gog, "Using Electroencephalography to Measure Cognitive Load," *Educ. Psychol. Rev.*, vol. 22, no. 4, pp. 425–438, 2010, doi: 10.1007/s10648-010-9130-y.

[149] Y. Borgianni and L. Maccioni, "Review of the use of neurophysiological and biometric measures in experimental design research," *Artif. Intell. Eng. Des. Anal. Manuf.*, vol. 34, no. 2, pp. 248–285, 2020.

[150] S. Vieira, M. Benedek, J. Gero, S. Li, and G. Cascini, "Design spaces and EEG frequency band power in constrained and open design," *Int. J. Des. Creat. Innov.*, vol. 00, no. 00, pp. 1–28, 2022, doi: 10.1080/21650349.2022.2048697.

[151] W. Jia and Y. Zeng, "EEG signals respond differently to idea generation , idea evolution and evaluation in a loosely controlled creativity experiment," *Sci. Rep.*, vol. 11, pp. 1–20, 2021.

[152] S. Vieira *et al.*, "The Neurophysiological Activations of Novice and Experienced Professionals When Designing and Problem-Solving," *Proc. Des. Soc. Des. Conf.*, vol. 1, no. June, pp. 1569–1578, 2020, doi: 10.1017/dsd.2020.121.

[153] S. Vieira *et al.*, "The neurophysiological activations of mechanical engineers and industrial designers while designing and problem-solving," *Des. Sci.*, vol. 6, no. September 2018, pp. 1–35, 2020, doi: 10.1017/dsj.2020.26.

[154] T. A. Nguyen and Y. Zeng, "Analysis of design activities using EEG signals," in *International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, 2010, vol. 44137, pp. 277–286, doi: 10.1115/DETC2010-28477.

[155] P. Promsorn, V. Boonyahotra, and P. Sittiprapaporn, "Spatial Abilities Improve Brain-Computer Interface Performance Indexed by Electroencephalography," in *14th International Conference on Electrical Engineering/Electronics, Computer, Telecomunications and Information Technology*, 2017, pp. 34–37.

[156] T. A. Nguyen and Y. Zeng, "Clustering Designers ' Mental Activities Based on Eeg Power," *Tools Methods Compet. Eng.*, pp. 1–7, 2012.

[157] B. J. Call, W. Goodridge, I. Villanueva, N. Wan, and K. Jordan, "Utilizing electroencephalography measurements for comparison of task- specific neural efficiencies: Spatial intelligence tasks," *J. Vis. Exp.*, vol. 2016, no. 114, pp. 1–13, 2016, doi: 10.3791/53327.

[158] C. J. Liu, C. F. Huang, C. Y. Chou, M. C. Lu, Y. Y. Chang, and M. C. Ho, “Applying frequency bands to explore the identification of two dimensional figures,” *Appl. Mech. Mater.*, vol. 311, pp. 196–201, 2013, doi: 10.4028/www.scientific.net/AMM.311.196.

[159] I. Riečanský and S. Katina, “Induced EEG alpha oscillations are related to mental rotation ability: The evidence for neural efficiency and serial processing,” *Neurosci. Lett.*, vol. 482, no. 2, pp. 133–136, 2010, doi: 10.1016/j.neulet.2010.07.017.

[160] A. Gevins, M. E. Smith, L. McEvoy, and D. Yu, “High-resolution EEG mapping of cortical activation related to working memory: Effects of task difficulty, type of processing, and practice,” *Cereb. Cortex*, vol. 7, no. 4, pp. 374–385, 1997, doi: 10.1093/cercor/7.4.374.

[161] M. Tovey, “Thinking styles and modelling systems,” *Des. Stud.*, vol. 7, no. 1, pp. 20–30, 1986, doi: 10.1016/0142-694X(86)90004-9.

[162] D. K. Lieu and S. Sorby, *Visualization, Modeling, and Graphics for Engineering Design*, vol. 148. Cengage Learning, 2016.

[163] W. Klimesch, “EEG alpha and theta oscillations reflect cognitive and memory performance: A review and analysis,” *Brain Res. Rev.*, vol. 29, no. 2–3, pp. 169–195, 1999, doi: 10.1016/S0165-0173(98)00056-3.

[164] S. Chikhi, N. Matton, and S. Blanchet, “EEG power spectral measures of cognitive workload: A meta-analysis,” *Psychophysiology*, vol. 59, no. 6, pp. 1–24, 2022, doi: 10.1111/psyp.14009.

[165] E. Basar, *Brain Functions and Oscillations, Volume II: Integrative Brain Functions. Neurophysiology and Cognitive Processes*. Springer, 1999.

[166] C. Pernet *et al.*, “Issues and recommendations from the OHBM COBIDAS MEEG committee for reproducible EEG and MEG research,” *Nat. Neurosci.*, vol. 23, no. 12, pp. 1473–1483, 2020, doi: 10.1038/s41593-020-00709-0.

[167] U. Maurer, S. Brem, M. Liechti, S. Maurizio, L. Michels, and D. Brandeis, “Frontal Midline Theta Reflects Individual Task Performance in a Working Memory Task,” *Brain Topogr.*, vol. 28, no. 1, pp. 127–134, 2015, doi: 10.1007/s10548-014-0361-y.

[168] Ş. Harputlu Aksu, E. Çakıt, and M. Dağdeviren, “Investigating the relationship between EEG features and N-back task difficulty levels with NASA-TLX scores among undergraduate students,” *Proc. 6th Int. Conf. Intell. Hum. Syst. Integr. (IHSI 2023) Integr. People Intell. Syst. Febr. 22–24, 2023, Venice, Italy*, vol. 69, no. February, 2023, doi: 10.54941/ahfe1002828.

[169] J. Morton *et al.*, “Danger, high voltage! Using EEG and EOG measurements for cognitive overload detection in a simulated industrial context,” *Appl. Ergon.*, vol. 102, no. March, p. 103763, 2022, doi: 10.1016/j.apergo.2022.103763.

[170] A. Gevins and M. E. Smith, “Neurophysiological measures of cognitive workload during human-computer interaction,” *Theor. Issues Ergon. Sci.*, vol. 4, no. 1–2, pp. 113–131, Jan. 2003, doi: 10.1080/14639220210159717.

[171] D. Grimes, D. S. Tan, S. E. Hudson, P. Shenoy, and R. P. N. Rao, “Feasibility and pragmatics of classifying working memory load with an electroencephalograph,” in *CHI 2008 proceedings: Cognition, Perception, and Memory*, 2008, pp. 835–844.

[172] A. M. Brouwer, M. A. Hogervorst, J. B. F. Van Erp, T. Heffelaar, P. H. Zimmerman, and R. Oostenveld, “Estimating workload using EEG spectral power and ERPs in the n-back task,”

J. Neural Eng., vol. 9, no. 4, 2012, doi: 10.1088/1741-2560/9/4/045008.

[173] M. A. Hogervorst, A. M. Brouwer, and J. B. F. van Erp, "Combining and comparing EEG, peripheral physiology and eye-related measures for the assessment of mental workload," *Front. Neurosci.*, vol. 8, no. OCT, pp. 1–14, 2014, doi: 10.3389/fnins.2014.00322.

[174] B. S. Cheema, S. Samima, M. Sarma, and D. Samanta, *Mental workload estimation from EEG signals using machine learning algorithms*, vol. 10906 LNAI. Springer International Publishing, 2018.

[175] K. Guan, X. Chai, Z. Zhang, Q. Li, and H. Niu, "Evaluation of Mental Workload in Working Memory Tasks with Different Information Types Based on EEG," *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. EMBS*, vol. 100191, no. 37, pp. 5682–5685, 2021, doi: 10.1109/EMBC46164.2021.9630575.

[176] A. Gundel and G. F. Wilson, "Topographical changes in the ongoing EEG related to the difficulty of mental tasks," *Brain Topogr.*, vol. 5, no. 1, pp. 17–25, 1992, doi: 10.1007/BF01129966.

[177] A. Mastropietro, I. Pirovano, A. Marciano, S. Porcelli, and G. Rizzo, "Reliability of mental workload indeks assessed by EEG with different electrode configurations and signal pre-processing pipelinse," *Sensors*, vol. 23, p. 1367, 2023.

[178] F. Lukačević, N. Becattini, M. M. Perišić, and S. Škec, "Differences in engineers' brain activity when CAD modelling from isometric and orthographic projections," *Sci. Rep.*, vol. 13, no. 1, pp. 1–16, 2023, doi: 10.1038/s41598-023-36823-9.

[179] P. Nguyen, T. A. Nguyen, and Y. Zeng, "Segmentation of design protocol using EEG," *Artif. Intell. Eng. Des. Anal. Manuf. AIEDAM*, vol. 33, no. 1, pp. 11–23, 2018, doi: 10.1017/S0890060417000622.

[180] S. Li, N. Becattini, and G. Cascini, "Neuro-cognitive Insights into Engineering Design: Exploring EEG Predictive Associations with Task Performance," *J. Mech. Des.*, 2024, doi: 10.1115/1.4066681.

[181] S. G. Hart and L. E. Staveland, "Development of NASA-TLX (Task Load Index): results of empirical and theoretical research," in *Human Mental Workload*, P. A. Hancock and N. Meshkati, Eds. Amsterdam,: Elsevier Science, 1988, pp. 139–183.

[182] G. Goldschmidt, "To see eye to eye: the role of visual representations in building shared mental models in design teams," *CoDesign*, vol. 3, no. 1, pp. 43–50, 2007, doi: 10.1080/15710880601170826.

[183] D. Boa and B. Hicks, "Information operations : A model for characterising information interaction of engineers," in *Analyzing Cognitive Processes during Design : Proceedings of the HBiD 2014*, 2014.

[184] Y. Shi, J. Du, and Q. Zhu, "The impact of engineering information format on task performance: Gaze scanning pattern analysis," *Adv. Eng. Informatics*, vol. 46, no. July, p. 101167, 2020, doi: 10.1016/j.aei.2020.101167.

[185] J. D. Summers and J. J. Shah, "Representation in engineering design: A framework for classification," in *Proceedings of the ASME Design Engineering Technical Conference*, 2004, vol. 3, no. January 2004, pp. 439–448, doi: 10.1115/detc2004-57514.

[186] A. Oti and N. Crilly, "Immersive 3D sketching tools: Implications for visual thinking and communication," *Comput. Graph.*, vol. 94, no. 2021, pp. 111–123, 2021, doi:

10.1016/j.cag.2020.10.007.

[187] K. L. Norman, “Spatial Visualization-A Gateway to Computer-Based Technology,” *J. Spec. Educ. Technol.*, vol. XII, no. 3, 1994.

[188] A. K. Goel, S. Vattam, B. Wiltgen, and M. Helms, “Cognitive, collaborative, conceptual and creative - Four characteristics of the next generation of knowledge-based CAD systems: A study in biologically inspired design,” *CAD Comput. Aided Des.*, vol. 44, no. 10, pp. 879–900, 2012, doi: 10.1016/j.cad.2011.03.010.

[189] G. Goldschmidt, “On visual design thinking: the vis kids of architecture,” *Des. Stud.*, vol. 15, no. 2, pp. 158–174, 1994, doi: 10.1016/0142-694X(94)90022-1.

[190] M. Suwa and B. Tversky, “What do architects and students perceive in their design sketches? A protocol analysis,” *Des. Stud.*, vol. 18, pp. 385–403, 1997.

[191] L. Hay, P. Cash, and S. McKilligan, “The future of design cognition analysis,” *Des. Sci.*, vol. 6, no. 20, pp. 1–26, 2020, doi: 10.1017/dsj.2020.20.

[192] J. W. Peirce, J. R. Gray, S. Simpson, R. MacAskill, M. R., Höchenberger, H. Sogo, and J. Kastman, E., Lindeløv, “PsychoPy2: Experiments in behavior made easy,” *Behav. Res. Methods*, 2019, doi: 10.3758/s13428-018-01193-y.

[193] M. X. Cohen, “Where Does EEG Come From and What Does It Mean?,” *Trends Neurosci.*, vol. 40, no. 4, pp. 208–218, 2017, doi: 10.1016/j.tins.2017.02.004.

[194] R. Renu and G. Mocko, “Retrieval of solid models based on assembly similarity,” *Comput. Des. Appl.*, vol. 13, no. 5, pp. 628–636, 2016, doi: 10.1080/16864360.2016.1150708.

[195] A. Cardone, S. K. Gupta, and M. Karnik, “A survey of shape similarity assessment algorithms for product design and manufacturing applications,” *J. Comput. Inf. Sci. Eng.*, vol. 3, no. 2, pp. 109–118, 2003, doi: 10.1115/1.1577356.

[196] P. Wittenburg, H. Brugman, A. Russel, A. Klassmann, and H. Sloetjes, “ELAN : a Professional Framework for Multimodality Research,” 2006.

[197] D. R. Krathwohl, “A Revision of Bloom’s Taxonomy: An Overview,” *Theory Pract.*, vol. 41, no. 4, pp. 212–218, 2002, doi: 10.1207/s15430421tip4104.

[198] R. Ornstein, J. Johnstone, J. Herron, and C. Swencionis, “Differential right hemisphere engagement in visuospatial tasks,” *Neuropsychologia*, vol. 18, no. 1, pp. 49–64, 1980, doi: 10.1016/0028-3932(80)90083-4.

[199] J. E. Roberts and M. Ann Bell, “Two- and three-dimensional mental rotation tasks lead to different parietal laterality for men and women,” *Int. J. Psychophysiol.*, vol. 50, no. 3, pp. 235–246, 2003, doi: 10.1016/S0167-8760(03)00195-8.

[200] I. Gerlič and N. Jaušovec, “Multimedia: Differences in cognitive processes observed with EEG,” *Educ. Technol. Res. Dev.*, vol. 47, no. 3, pp. 5–14, 1999, doi: 10.1007/bf02299630.

[201] B. R. Fajen and F. Phillips, “Spatial perception and action,” in *Handbook of Spatial Cognition*, First Edit., D. Waller and L. Nadel, Eds. Washington DC: American Psychological Association, 2013, pp. 67–80.

[202] A. Delorme and S. Makeig, “EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis,” *J. Neurosci. Methods*, vol. 134, no. 1, pp. 9–21, Mar. 2004, doi: 10.1016/j.jneumeth.2003.10.009.

[203] S. Li, N. Becattini, and G. Cascini, “Correlating design performance to EEG activation: Early

evidence from experiental data," in *Proceedings of the Design Society*, 2021, pp. 771–780, doi: 10.1017/pds.2021.77.

[204] W. De Clercq, A. Vergult, B. Vanrumste, W. Van Paesschen, and S. Van Huffel, "Canonical Correlation Analysis Applied to Remove Muscle Artifacts From the Electroencephalogram," *IEEE Trans. Biomed. Eng.*, vol. 53, no. 12, pp. 2583–2587, Nov. 2006, doi: 10.1109/TBME.2006.879459.

[205] G. Pfurtscheller and F. H. Lopes da Silva, "Event-related EEG/MEG synchronization and desynchronization: basic principles," *Clin. Neurophysiol.*, vol. 110, no. 11, pp. 1842–1857, 1999.

[206] "ARTool: Aligned Rank Transform." https://cran.r-project.org/web/packages/ARTool/index.html.

[207] J. O. Wobbrock, L. Findlater, D. Gergle, and J. J. Higgins, "The Aligned Rank Transform for nonparametric factorial analyses using only ANOVA procedures," *Conf. Hum. Factors Comput. Syst. - Proc.*, pp. 143–146, 2011, doi: 10.1145/1978942.1978963.

[208] S. G. Willis, G. H. Wheatley, and O. R. Mitchell, "Cerebral processing of spatial and verbal-analytic tasks: An EEG study," *Neuropsychologia*, vol. 17, no. 5, pp. 473–484, 1979, doi: 10.1016/0028-3932(79)90054-X.

[209] W. Klimesch, H. Schimke, and G. Pfurtscheller, "Alpha Frequency , Cognitive Load and Memory Performance," *Brain Topogr.*, vol. 5, no. 3, pp. 241–251, 1993.

[210] B. Xie and G. Salvendy, "Review and reappraisal of modelling and predicting mental workload in single- and multi-task environments," *Work Stress An Int. J. Work , Heal. Organ.*, no. January 2013, pp. 37–41, 2000.

[211] M. Teplan, "Fundamentals Of EEG Measurment," *Meas. Sci. Rev.*, vol. 2, no. 2, pp. 1–11, 2003.

[212] W. Visser, "Design: one, but in different forms," *Des. Stud.*, vol. 30, no. 3, pp. 187–223, 2009, doi: 10.1016/j.destud.2008.11.004.

[213] M. Štorga, M. M. Andreasen, and D. Marjanović, "The design ontology: foundation for the design knowledge exchange and management," *J. Eng. Des.*, vol. 21, no. 4, pp. 427–454, 2010, doi: 10.1080/09544820802322557.

[214] C. Eckert and J. Clarkson, "The reality of design," in *Design Process Improvement*, C. Eckert and J. Clarkson, Eds. 2005, pp. 1–29.

[215] J. Stempfle and P. Badke-Schaub, "Thinking in design teams - An analysis of team communication," *Des. Stud.*, vol. 23, no. 5, pp. 473–496, 2002, doi: 10.1016/S0142-694X(02)00004-2.

[216] T. Martinec, S. Škec, N. Horvat, and M. Štorga, "A state-transition model of team conceptual design activity," *Res. Eng. Des.*, vol. 30, no. 1, pp. 103–132, 2019, doi: 10.1007/s00163-018-00305-1.

[217] J. S. Gero, "Design prototypes. A knowledge representation schema for design," *AI Mag.*, vol. 11, no. 4, pp. 26–36, 1990.

[218] O. W. Klaproth, C. Vernaleken, L. R. Krol, M. Halbruegge, T. O. Zander, and N. Russwinkel, "Tracing Pilots ' Situation Assessment by Neuroadaptive Cognitive Modeling," *Front. Neurosci.*, vol. 14, no. August, pp. 1–12, 2020, doi: 10.3389/fnins.2020.00795.

[219] C. Wu, Y. Liu, and C. M. Quinn-walsh, "Queuing Network Modeling of a Real-Time Psychophysiological Index of Mental Workload — P300 in Event-Related Potential (ERP)," in *IEEE Transations on Systems, MAN, and Cybernetics - Part A: Systems and Humans*, 2008, vol. 38, no. 5, pp. 1068–1084.

[220] N. Y. Kim, R. House, M. H. Yun, C. S. Nam, and A. R. Harrivel, "Neural Correlates of Workload Transition in Multitasking : An ACT-R Model of Hysteresis Effect," *Front. Hum. Neurosci.*, vol. 12, no. January, pp. 1–14, 2019, doi: 10.3389/fnhum.2018.00535.

[221] S. Prezenski and N. Russwinkel, "A Proposed Method of Matching ACT-R and EEG-Data," in *Proceedings of the 14th International Conference on Cognitive Modeling (ICCM 2016)*, 2016, pp. 249–251.

[222] M. Suwa, J. Gero, and T. Purcell, "Analysis of cognitive processes of a designer as the foundation for support tools," *Artif. Intell. Des. '98*, no. November, 1998, doi: 10.1007/978-94-011-5121-4.

[223] V. Anand, S. R. Sreeja, and D. Samanta, "An automated approach for task evaluation using EEG signals," *ArXiv*, pp. 1–19, 2019.

[224] J. M. Stern, *Atlas of EEG Patterns*. Wolters Kluwer, Lippincott Williams & Wilkins.

[225] C. Gerloff, J. Richard, J. Hadley, A. E. Schulman, M. Honda, and M. Hallett, "Functional coupling and regional activation of human cortical motor areas during simple, internally paced and externally paced finger movements," *Brain*, vol. 121, no. 8, pp. 1513–1531, 1998, doi: 10.1093/brain/121.8.1513.

[226] J. Z. Bakdash, L. R. Marusich, and J. H. Bolin, "Repeated Measures Correlation," *Front. Psychol.*, vol. 8, no. April, pp. 1–13, 2017, doi: 10.3389/fpsyg.2017.00456.

[227] J. S. Bridle, "Probabilistic Interpretation of Feedforward Classification Network Outputs , with Relationships to Statistical Pattern Recognition," in *NASA ASI Series Neurocomputing*, 1989, no. F 68.

[228] H. H. Jasper, "The ten-twenty electrode system of the International Federation," *Electroencephalogr. Clin. Neurophysiol.*, vol. 10, pp. 370–375, 1958.

[229] S. G. Hart, "NASA-task load index (NASA-TLX); 20 years later," *Proc. Hum. Factors Ergon. Soc.*, pp. 904–908, 2006, doi: 10.1177/154193120605000909.

[230] R. R. Wilcox, "A Note on the Theil-Sen Regression Estimator When the Regressor Is Random and the Error Term Is Heteroscedastic," *Biometrical J.*, vol. 40, no. 3, pp. 261–268, 1998.

[231] O. Jensen, "Distractor inhibition by alpha oscillations is controlled by an indirect mechanism governed by goal-relevant information," *Commun. Psychol.*, vol. 2, no. 36, 2024, doi: 10.1038/s44271-024-00081-w.

[232] A. M. Tuladhar, N. Huurne, J. Schoffelen, E. Maris, R. Oostenveld, and O. Jensen, "Parieto-Occipital Sources Account for the Increase in Alpha Activity with Working Memory Load," *Hum. Brain Mapp.*, vol. 28, pp. 785–792, 2007, doi: 10.1002/hbm.20306.

[233] E. Wianda and B. Ross, "The roles of alpha oscillation in working memory retention," *Brain Behav.*, vol. 9, p. e01263, 2019, doi: 10.1002/brb3.1263.

[234] M. Benedek, R. J. Schickel, E. Jauk, A. Fink, and A. C. Neubauer, "Alpha power increases in right parietal cortex reflects focused internal attention," *Neuropsychologia*, vol. 56, pp. 393–400, 2014, doi: 10.1016/j.neuropsychologia.2014.02.010.

[235] M. Cass and R. Prabhu, "Looking Beyond Self-Reported Cognitive Load: Comparing Pupil Diameter against Self-Reported Cognitive Load in Design Tasks," *J. Mech. Des.*, 2024, doi: 10.1115/1.4067343.

APPENDIX

Appendix A

High performer

The high performer's modelling process in the isometric projection condition was divided into 14 unit tasks, presented in Figure 76 in the order of their appearance (from the left to the right). Unit tasks represent either sketching or applying a CAD feature; it can be noticed from Figure 76 that they were performed in strict Sketch → CAD feature order. After the first, the longest sketching period, approximately in every consecutive minute a new CAD feature was applied to the created sketch. Moreover, two occurrences of the unit task *Extrude Cut E* can be noticed from Figure 76; the participant made an error while solving this subgoal and therefore had to edit it.

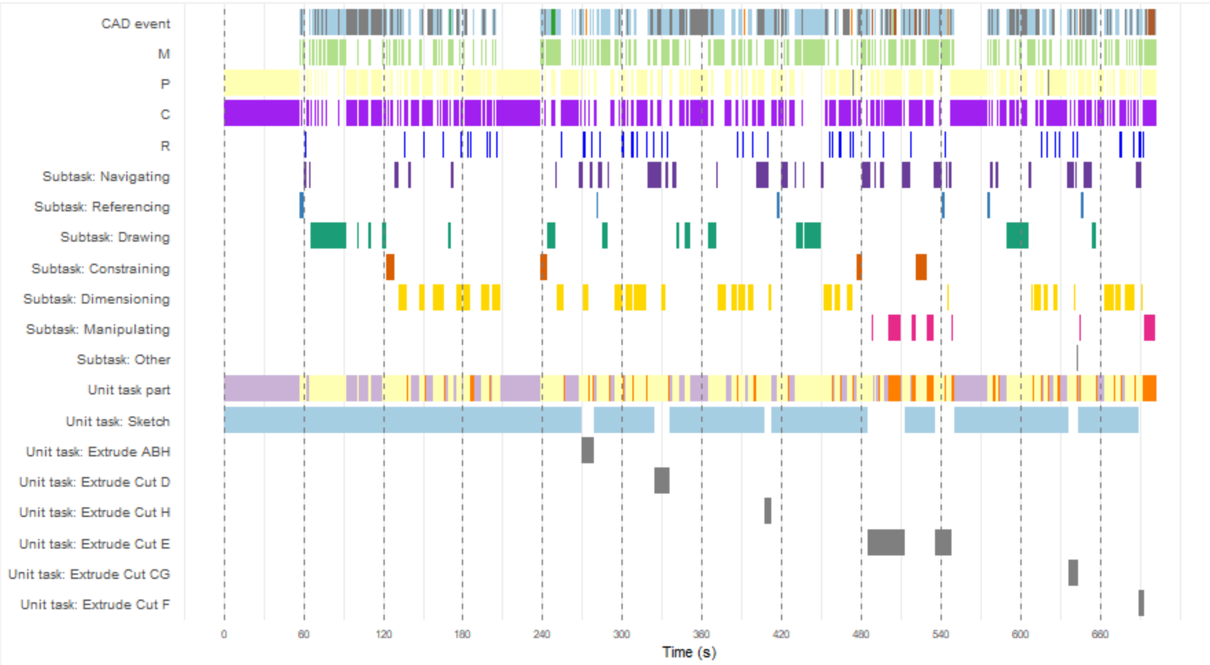

Figure 76 High performer: Descriptive model of CAD performance - Isometric condition

Similar organisation of unit tasks was observed in the orthographic condition as well; descriptive model is visualised in Figure 77. The sequence of the unit tasks was similar to the isometric condition, only the overall number of the unit tasks through which the goal was reached was lower (see Table 84 for details). In this condition, the participant did not make any errors that would require editing the sketch and the associated CAD feature.

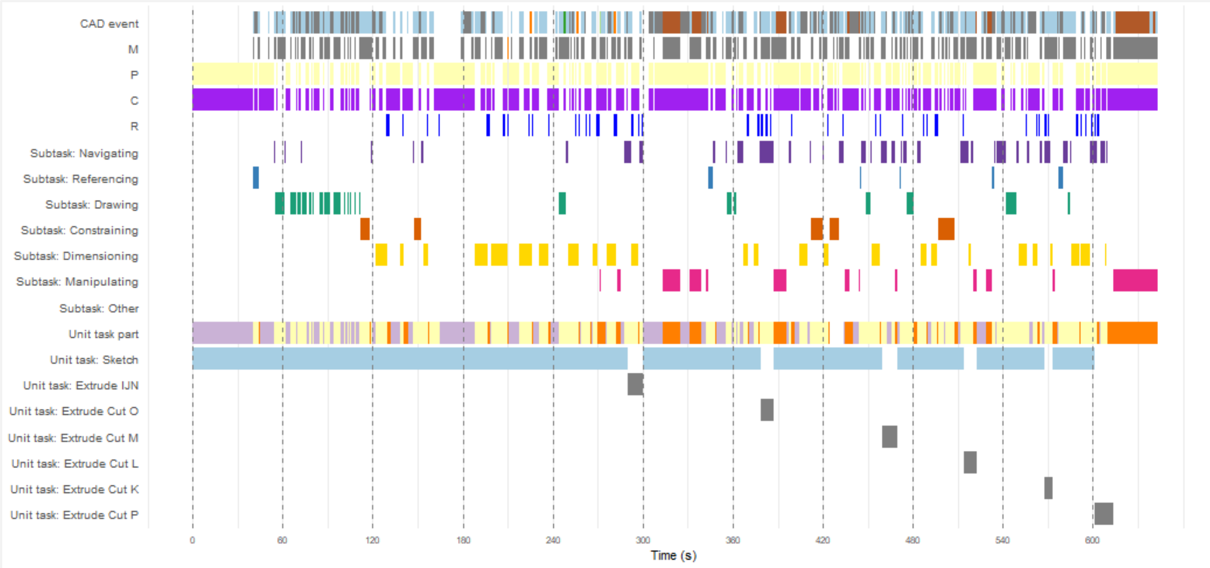

Figure 77 High performer: Descriptive model of CAD performance - Orthographic condition

Decomposition of unit tasks into specific types, unit task phases, and subtask clusters is detailed in the following sections.

Unit tasks

High performer in both conditions used the same two types of CAD features: *Extrude Cut* and *Extrude*. An average duration of each unit task is available in Table 84.

Table 84 High performer: Types, number of occurrences, and average duration of unit tasks

Unit task	N of occurrences		Duration (M) in [s]		Duration (Med) in [s]	
	Iso	Ortho	Iso	Ortho	Iso	Ortho
Sketch	7	6	87.6	79.9	83.2	45
Extrude Cut	6	5	11.5	9	8.55	9
Extrude	1	1	9	10	9	10

There were many more *acquisition*, *execution*, and *inspection* task parts than unit tasks in both conditions, implying that participants divided unit tasks into smaller subtasks, for which subgoals were defined during their *acquisition* parts. The number of "closed" *acquisition-execution-inspection* sequences was similar in both conditions; 29 in the isometric (see Figure 78) and 28 in the orthographic (see Figure 79) condition.

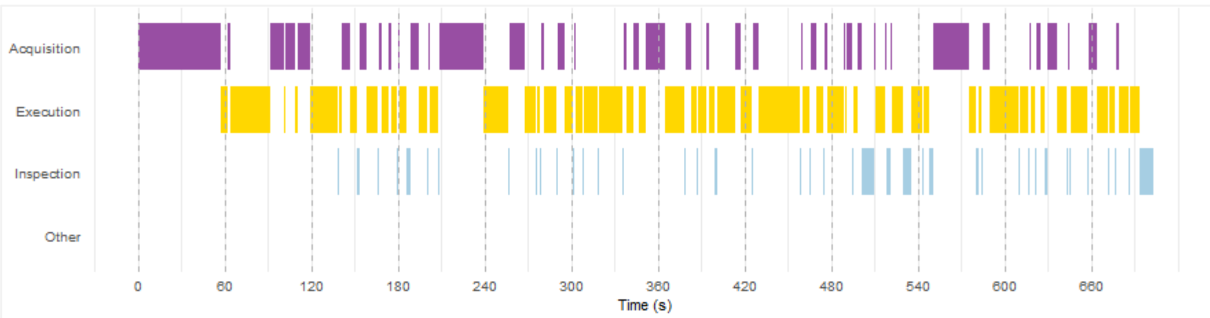

Figure 78 High performer: Visualisation of unit task phases - Isometric condition

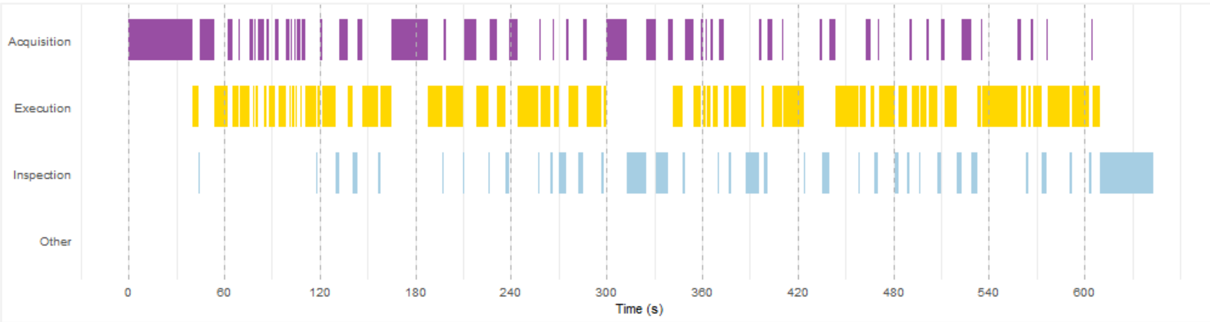

Figure 79 Visualisation of unit task phases - Orthographic condition

The number of *acquisition* and *execution* occurrences was higher in the orthographic condition, while the number of *inspection* occurrences was lower. However, when the time distribution of these three unit task phases are compared, the results are the opposite - *execution* and *acquisition* took larger portion of the CAD modelling task when conducted from the isometric projection, while *inspection* took larger percentage when modelling from the orthographic projection, despite the lower number of occurrences in these conditions. Details are available in Table 85.

Table 85 High performer: Number of occurrences, average duration, and distribution of unit task phases

Unit task phase	N of occurrences		Duration (M; SD) in [s]		Duration (Med; MAD) in [s]		Distribution [%]	
	Iso	Ortho	Iso	Ortho	Iso	Ortho	Iso	Ortho
Execution	56	61	6.46; 5.80	5.21	5.00; 2.97	4	54.20	50.20
Acquisition	47	59	5.34; 9.57	3.32	2.00; 2.97	2	35.70	30.90
Inspection	46	39	1.54; 1.82	3.08	1.00; 0.00	1	10.10	18.90
Other	0	0	-	-	-	-	-	-

Distribution of the subtask (CAD operation) clusters across unit tasks is presented in Figure 80 for the isometric condition, and in Figure 81 for the orthographic condition. These distributions imply that *acquisition* and *inspection* took larger percentage of time when sketching than when applying CAD features. Moreover, the High performer spent the largest proportion of sketching time on *dimensioning* in almost all sketching unit tasks, while *navigating* dominated application of CAD features (see Figure 80 and Figure 81).

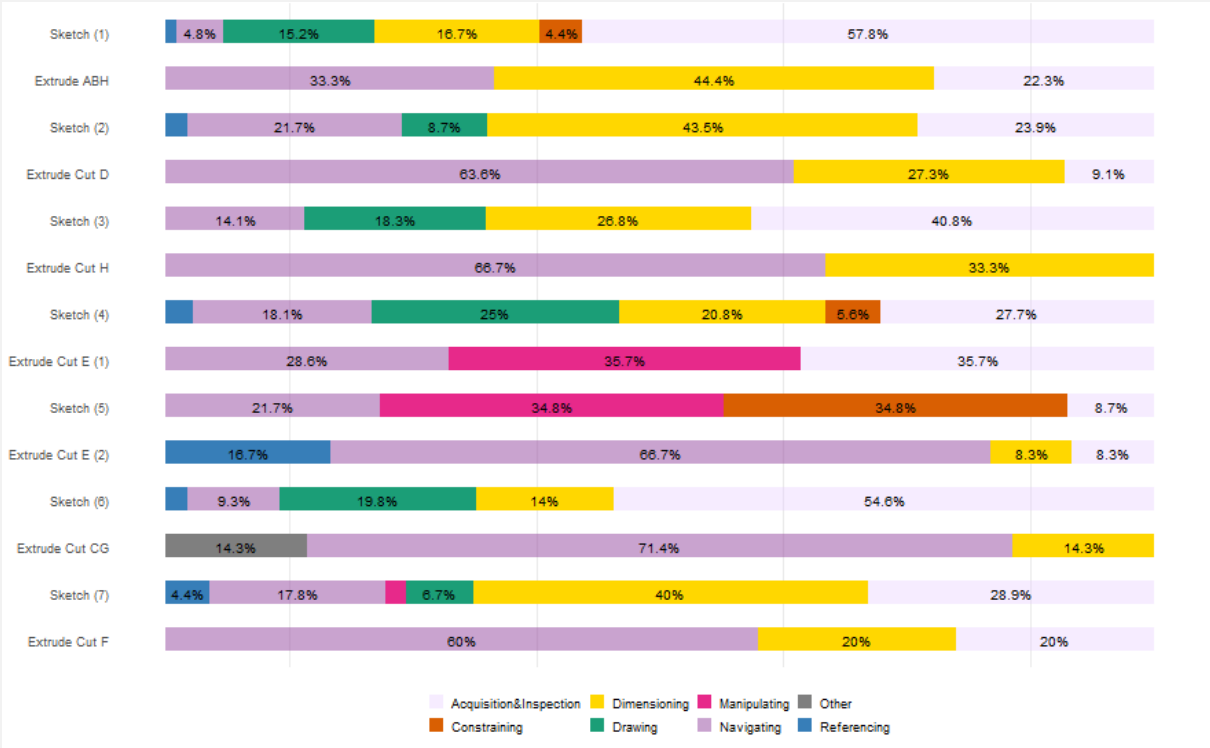

Figure 80 High performer: Distribution of subtask clusters across unit tasks - Isometric condition

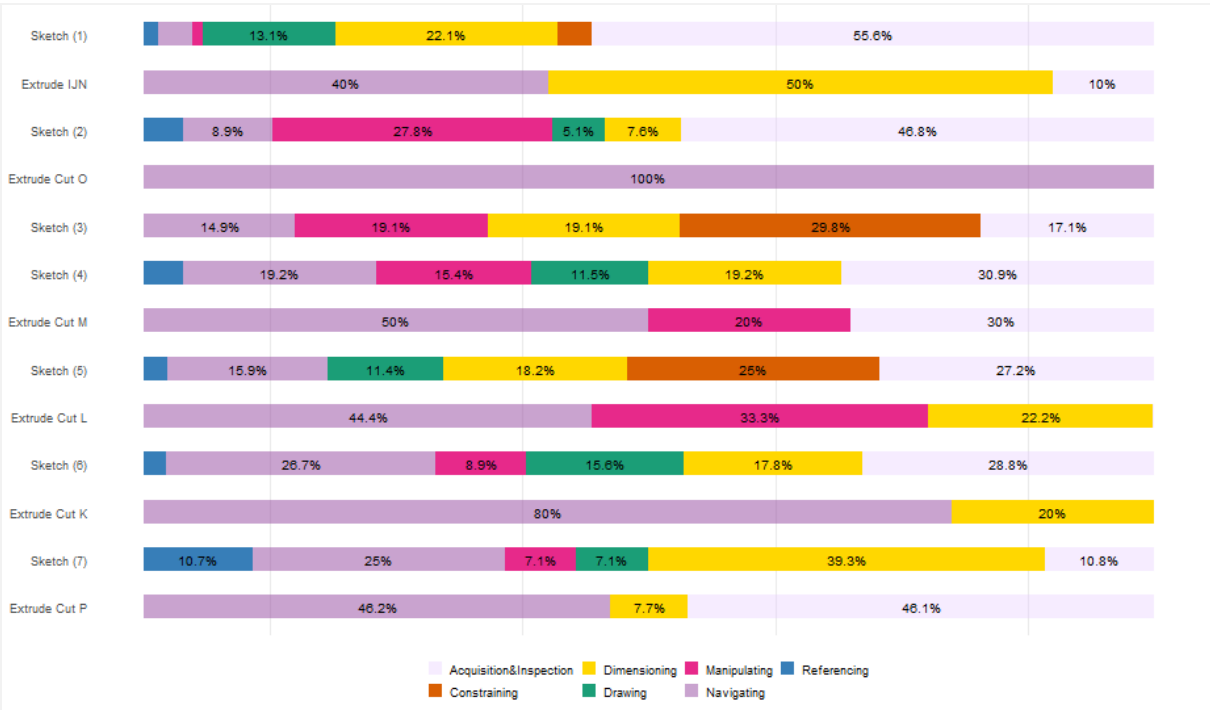

Figure 81 High performer: Distribution of subtask clusters across unit tasks - Orthographic condition

Subtask (CAD operation) clusters

The number of occurrences and mean as well as Med duration of the subtask clusters is shown in Table 86. *Other* stands for transition periods, such as mouse movements captured at the end of the *execution* for which no context indicated intention of this action.

Table 86 High performer: Number of occurrences and average duration of subtasks at the level of the entire task

Subtask cluster	N of occurrences		Duration (M) in [s]		Duration (Med) in [s]	
	Iso	Ortho	Iso	Ortho	Iso	Ortho
Navigating	54	51	1.94	1.69	2	1
Dimensioning	33	26	4.27	4.62	5	4.5
Drawing	14	24	6.86	2.46	4.5	2
Manipulating	8	16	3.62	4.88	2.5	2
Constraining	6	5	4	7	4.5	6
Referencing	6	6	2	2.17	2	2
Deleting	0	0	-	-	-	-
Other	1	0	1	0.4	1	-

The rank of the subtask clusters regarding their distribution is shown in Figure 82; *dimensioning, navigating*, and *drawing* took the largest portion of task performance time in both conditions. The rest of the task (percentage of the distribution that is not covered in Figure 82) was spent on *Acquisition*, *Inspection*, and transitions between the subtasks during *Execution*.

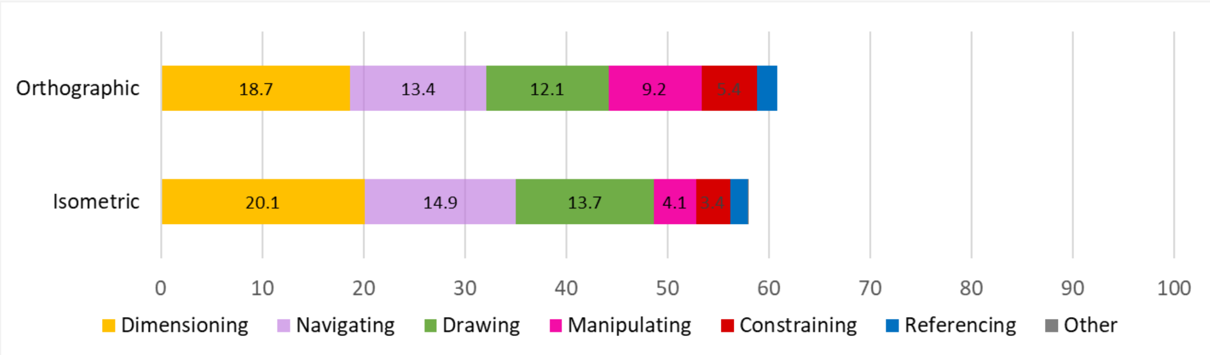

Figure 82 High performer: Distribution of subtasks

CAD actions

The number of occurrences of CAD actions was similar in both conditions, as shown in Table 87. The largest difference was in the number of CAD actions *hovering over an icon/feature/point*.

Table 87 High performer: Number of occurrences and average duration of CAD actions at the level of the entire task

CAD action	N of occurrences		Duration (M) in [s]		Duration (Med) in [s]	
	Iso	Ortho	Iso	Ortho	Iso	Ortho
Moving a cursor	273	274	0.952	0.909	1	1
Clicking on an icon/feature/textbox	181	173	0.001	0.001	0	0
Hovering over an icon/feature/point	148	185	1.01	0.714	1	1
Typing a value	27	21	0.185	0.143	0	0
Rotating view	13	15	1	3.87	1	2
Zooming in/out	13	10	0.0769	0.1	0	0
Dragging a geometric entity	2	5	2	2	2	2
Entering (striking) a command	2	1	0.001	0.001	0	0

Operators’ utilisation throughout the task performance

Utilisations of operators are first visualised and discussed across the unit tasks and the subtasks (CAD operation) clusters.

Across the unit tasks

Figure 83 (isometric) and Figure 84 (orthographic) show P, C, and M operator utilisations across the High performer’s execution of unit tasks. In both conditions, operator utilisations seem to depend on the unit task type; *sketch* unit tasks showed higher utilisation of C and P systems, whereas several *extrude*/*extrude cut* steps coincided with M peaks. Both sequences ended with an increase in the utilisation of all three operators, indicating convergent processing demands and simultaneous M activity in the final step.

Despite these shared characteristics, the dynamics of changes in utilisations differed between the projections. Frequent crossovers between operators and sharper local drops can be noticed in the isometric condition, particularly for P and C during feature-application unit tasks. In contrast, the orthographic condition showed smoother trajectories for P and C utilisations, including a noticeable plateau of relatively high P and C across the sequence *Sketch_3* – *Sketch_5*. Here, changes were more clearly localized to specific *extrude* unit tasks, where M spiked while P and C temporarily dipped, before returning to a “stable” level in the next *sketch* unit task.

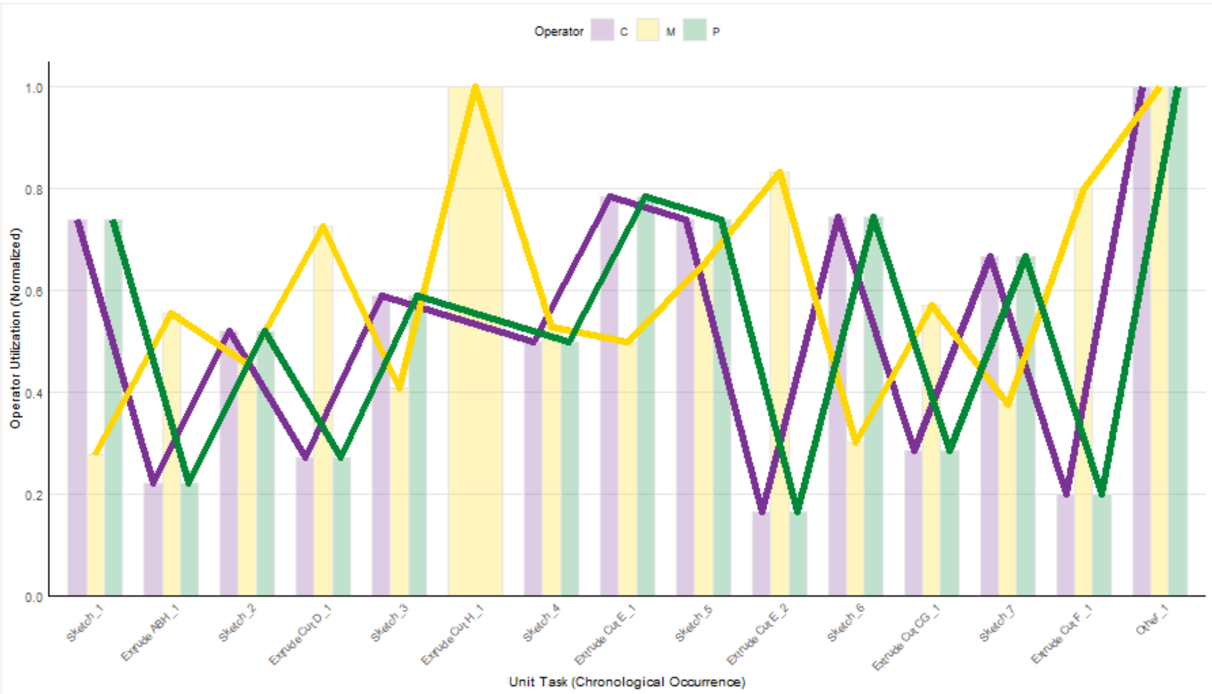

Figure 83 High performer: Average utilisation of operators across unit tasks – Isometric condition

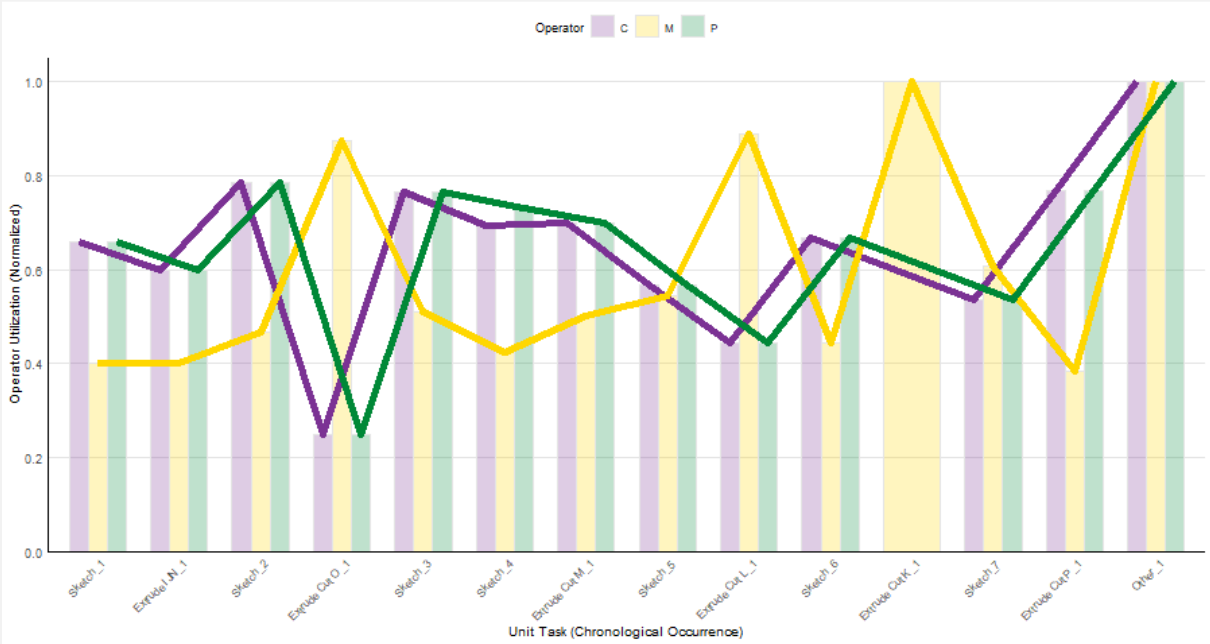

Figure 84 High performer: Average utilisation of operators across unit tasks – Orthographic condition

Finally, in the orthographic condition, C and P rose more gradually and in parallel toward the final unit tasks, whereas in the isometric condition low-to-high transition preceded the final convergence at the final unit task. Overall, for this participant, the orthographic projection is associated with a more stable P and C utilisation patterns across unit tasks, while the isometric projection showed stronger reallocations of operator demands from unit task to unit task.

Across the subtask (CAD operator) clusters

Across subtask clusters, operator utilisations indicated the effect of projection (isometric vs. orthographic), which depended on the subtask type (Table 88).

Table 88 High performer's average operator utilisation across subtasks

	Navigating		**Referencing**		**Drawing**		**Constraining**		**Dimensioning**		**Manipulating**	
Operator	**Iso**	**Ortho**	**Iso**	**Ortho**	**Iso**	**Ortho**	**Iso**	**Ortho**	**Iso**	**Ortho**	**Iso**	**Ortho**
P	0.56	0.52	0.56	0.64	0.37	0.32	0.50	0.52	0.50	0.50	1.00	1.00
C	0.56	0.52	0.56	0.64	0.37	0.32	0.50	0.52	0.50	0.50	1.00	1.00
M	0.67	0.76	0.67	0.50	0.75	0.83	0.63	0.57	0.53	0.54	0.90	1.00

The highest information-processing demand in both projections was found when *manipulating* (both P and C were maximal in both conditions, while M increased from 0.90 in isometric to 1.00 in the orthographic condition), indicating a stronger M system activity when modelling from orthographic projections. *Drawing* was consistently M-dominant, with M exceeding P and

C in both conditions but becoming even more motor-loaded in orthographic (M: 0.83 vs. 0.75; P and C: 0.32 vs. 0.37). A similar shift appeared in *navigating*, where modelling from the orthographic projection was associated with higher M (0.76 vs. 0.67) alongside slightly lower P and C utilisation (0.52 vs. 0.56).

In contrast, *referencing* showed the opposite pattern; modelling from the orthographic projection increased P and C utilisation (0.64 vs. 0.56) while reducing M utilisation (0.50 vs. 0.67), suggesting more interpretive and visual processing when extracting information from the orthographic projections. *Constraining* differences were smaller but pointed in the same direction (slightly higher P and C, and lower M in the orthographic projection condition. Utilisations were similar across projections when *dimensioning*, the only difference is 0.53 vs. 0.54 for the M operator.

Overall, the operator utilisation averages indicate that moving from isometric to orthographic projections was associated with more motor-intensive *navigating*, *drawing*, and *manipulating*, but more perceptually and cognitively demanding *referencing*, while *dimensioning* remained unchanged.

Low performer

In both timeline plots (isometric condition in Figure 85; orthographic condition in Figure 86), the Low participant's performance was organised around prolonged *Sketch* unit tasks, with *feature-application* unit tasks (*Extrude / Extrude Cut / Chamfer*) occurring as discrete blocks that interrupt *sketching*. Despite this shared structure, the organisation into unit tasks differed notably between projections.

Modelling from the isometric projection was relatively linear; one dominant *sketch* unit task spanned most of the early and mid-execution, with only brief interruptions. *Feature application* appeared as a short sequence of isolated unit tasks in the second half of the process. Importantly, there were no explicit rework unit tasks (e.g. no *delete* operations), and the unit-task sequence suggests a single forward pass from *sketching* into *feature* completion.

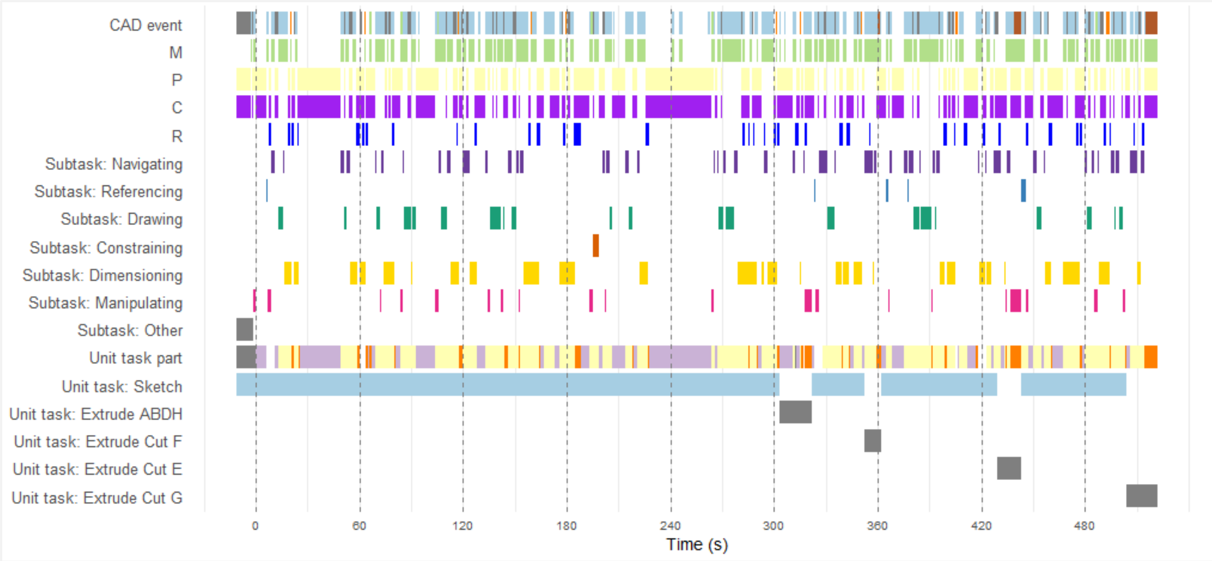

Figure 85 Low performer: Descriptive model of CAD performance - Isometric condition

Modelling from the orthographic projection was longer and more segmented, with greater unit-task variety and a clearer two-phase structure. First, there was an extended *sketching* phase (roughly up to 660 s), but unlike the isometric case it was punctuated by additional *feature* unit tasks earlier in the timeline (e.g. *Extrude JN* and *Extrude O* appear before the main late-stage refinement). Second, the final third of the session (after 660 s) became a dense *feature* application and refinement window, where multiple *feature-application* unit tasks occurred in close succession. Crucially, orthographic condition included explicit rework: *Delete Extrude Cut PK* and *Delete sketch* appear as unit tasks, indicating correction and iteration before finalising the model. *Sketching* also became more fragmented late in the session, returning in short bursts between *feature* unit tasks rather than remaining continuous.

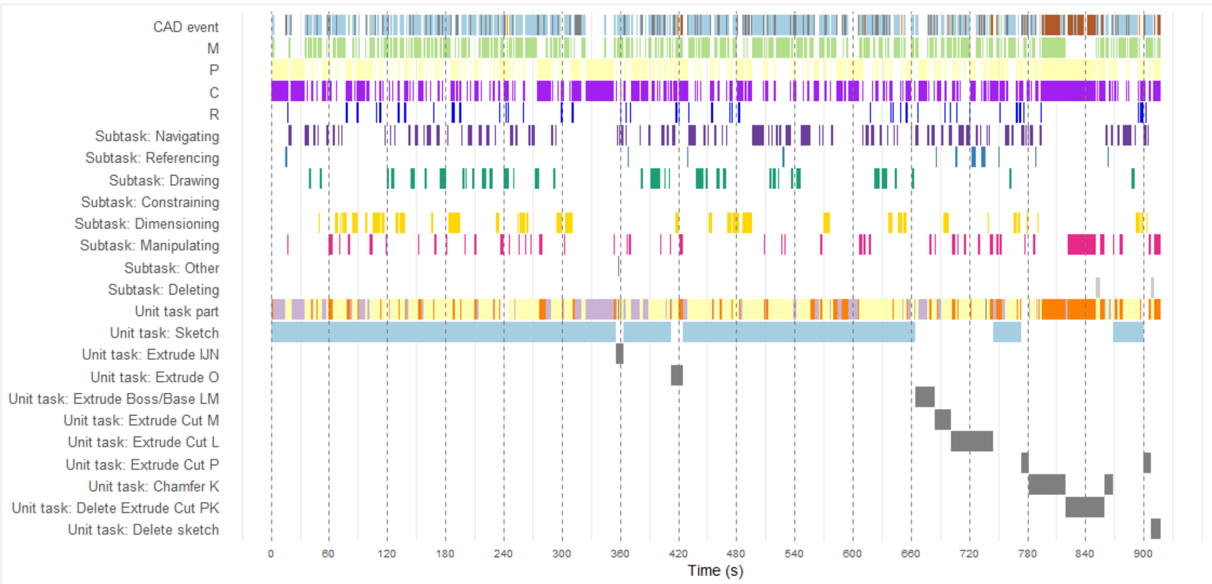

Figure 86 low performer: Descriptive model of CAD - Orthographic condition

Overall, the isometric condition showed a more streamlined unit-task organisation (long *sketching* followed by a short, one-way chain of *feature* applications), whereas the orthographic condition showed a more iterative unit-task organisation with more unit-task types, more switching, and visible rework (delete operations) concentrated in the late refinement stage.

Unit tasks

Compared to the isometric condition, CAD modelling from the orthographic condition by the Low performer was organised into more unit tasks and included additional unit task types (see Table 89), indicating a more iterative and corrective workflow.

Table 89 Low performer: Types, number of occurrences, and average duration of unit tasks

Unit task	N of occurrences		Duration (M) in [s]		Duration (Med) in [s]	
	Iso	Ortho	Iso	Ortho	Iso	Ortho
Sketch	4	6	118	118	64	75
Extrude Cut	3	4	14	19	14	12
Extrude	1	3	19	13.3	19	12
Chamfer	-	2	-	23.5	-	23.5
Delete Extrude Cut	-	1	-	40	-	40
Delete Sketch	-	1	-	9	-	9

Modelling from the orthographic projection by the Low performer included more overall occurrences of unit task phases (215 vs. 128) and showed a clear shift in the time distribution toward *execution* and *inspection* (see Table 90, Figure 87 and Figure 88). In the isometric condition, time was primarily allocated to *execution* (56.10%) and *acquisition* (31.40%), with *inspection* comprising only 10.20%. In contrast, in the orthographic condition, the share of time devoted to *execution* increased (60.50%), *acquisition* decreased substantially (19.40%), and *inspection* nearly doubled (19.90%). Thus, orthographic modelling for the Low performer was characterised by less time spent in *acquisition* and substantially more time spent *inspecting*, alongside slightly more *execution*.

Table 90 Low performer: Number of occurrences, average duration, and distribution of unit task phases

Unit task phase	N of occurrences		Duration (M) in [s]		Duration (Med) in [s]		Distribution [%]	
	Iso	Ortho	Iso	Ortho	Iso	Ortho	Iso	Ortho
Execution	48	82	5.98	6.61	5	5	56.10	60.5
Acquisition	39	70	4.13	2.49	2	1	31.40	19.40
Inspection	39	61	1.33	2.92	1	1	10.20	19.90
Other	2	2	6.00	1	6	1	2.34	0.22

Across unit task phases, the number of occurrences increased in the orthographic condition for all unit task phases: *execution* increased from 48 to 82, *acquisition* from 39 to 70, and *inspection* from 39 to 61. Regarding duration, *execution* was slightly longer per occurrence in the orthographic condition (mean 6.61 s vs. 5.98 s; median 5 s in both), while *acquisition* became notably shorter (mean 2.49 s vs. 4.13 s; median 1 s vs. 2 s), indicating many brief *acquisition* episodes distributed throughout the modelling session. *Inspection* was clearly longer in the orthographic condition (mean 2.92 s vs. 1.33 s; median 2 s vs. 1 s), which is also visible in Figure 88 as denser and more sustained *inspection* windows than in Figure 87.

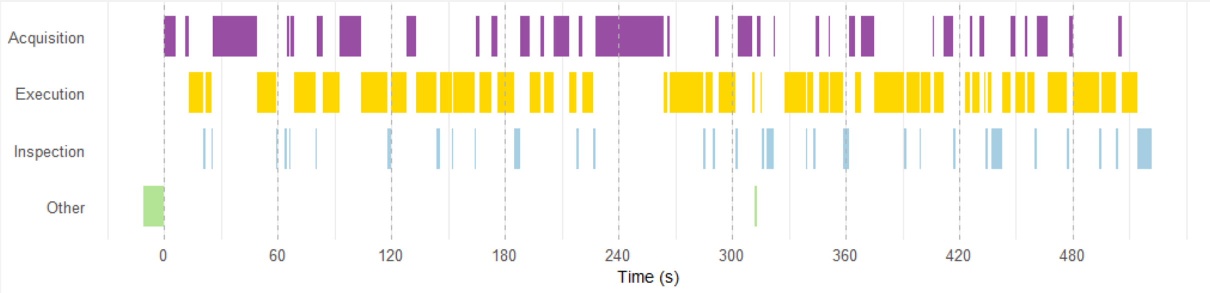

Figure 87 Low performer: Visualisation of unit task phases - Isometric condition

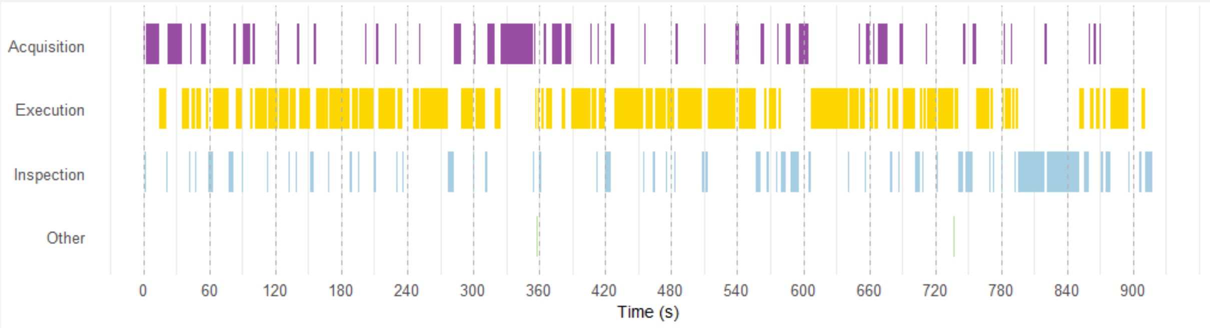

Figure 88 Low performer: Visualisation of unit task phases - Orthographic condition

Finally, the number of *acquisition–execution–inspection* cycles differed strongly between conditions; the Low performer completed many more cycles in the orthographic condition (29 vs. 17). Combined with the higher number of unit task phase occurrences, this suggests that modelling from the orthographic projection involved more frequent transitions through unit task phases, and that the process was not only longer but also more repeatedly re-initiated, aligning with the more segmented unit-task structure and the increased inspection share visible in Figure 88.

In the isometric condition, unit tasks showed a relatively structured distribution of subtask clusters (see Figure 89). *Sketch* unit tasks were consistently dominated by *acquisition* and *inspection*, which took the largest share in *Sketch(1)* (43.9%) and remained substantial across subsequent sketches (20–37%). The remaining sketch time was mainly split between *dimensioning* (e.g. 22.9–36.7%) and *navigating* as well as *drawing* (typically 12–21%), with only small percentages of *manipulating*. In contrast, the *feature-application* unit tasks were less diverse and tended to be driven by a small number of subtask clusters: *Extrude ABDH* and *Extrude Cut G* were dominated by *acquisition* and *inspection* (63.1% and 55.6%), while *Extrude Cut F* primarily consisted of *navigating* (60%) with smaller contributions of *acquisition* and *inspection* (30%) and *dimensioning* (10%). The most distinct deviation is *Extrude Cut E*, which was dominated by *manipulating* (50%), indicating that even in isometric modelling the Low performer occasionally relied heavily on *manipulation* during *feature application*.

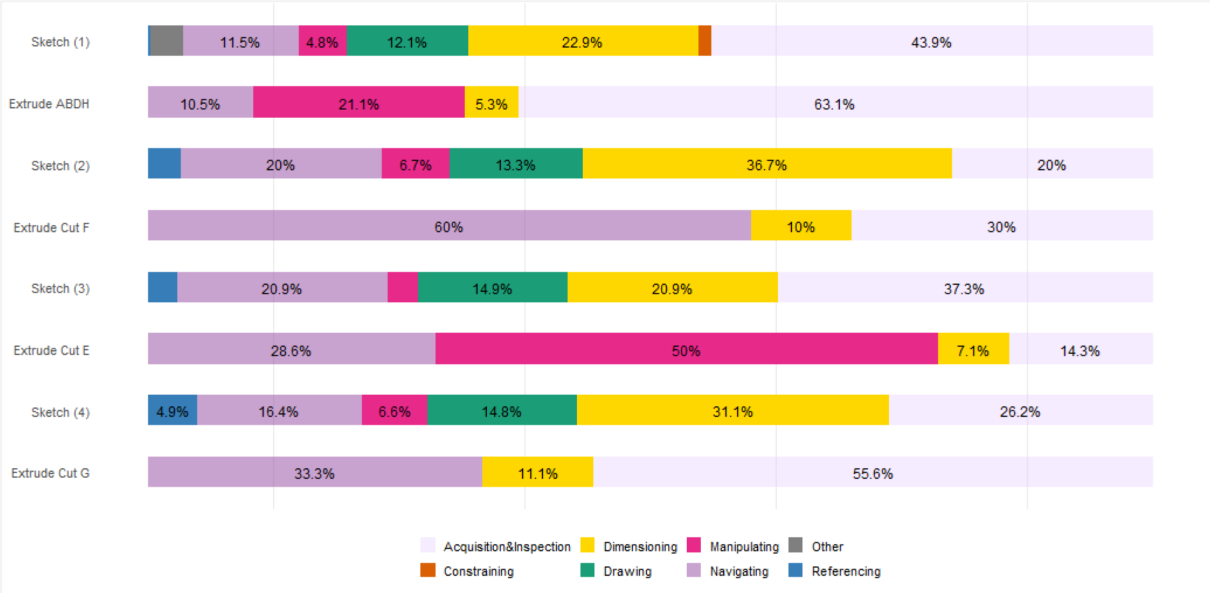

Figure 89 Low performer: Distribution of subtask clusters across unit tasks - Isometric condition

The orthographic condition showed greater diversity within unit tasks and more pronounced evidence of iterative refinement and rework (see Figure 90). *Sketch* unit tasks varied more obviously in their dominant subtasks: *Sketch(1)* remained dominated by *acquisition* and *inspection* (35%), but *Sketch(2)* was dominated by *drawing* (39.7%) with substantial *dimensioning* (32.7%), while later *sketches* (e.g. *Sketch(6)* and *Sketch(8)*) became *navigating*-dominant (62.6% and 38.7%). *Feature-application* unit tasks were also more heterogeneous than in the isometric condition, with multiple unit tasks showing large *manipulating* shares and explicit rework tasks that were driven by *manipulating*, alongside a non-negligible *deleting* CAD operations.

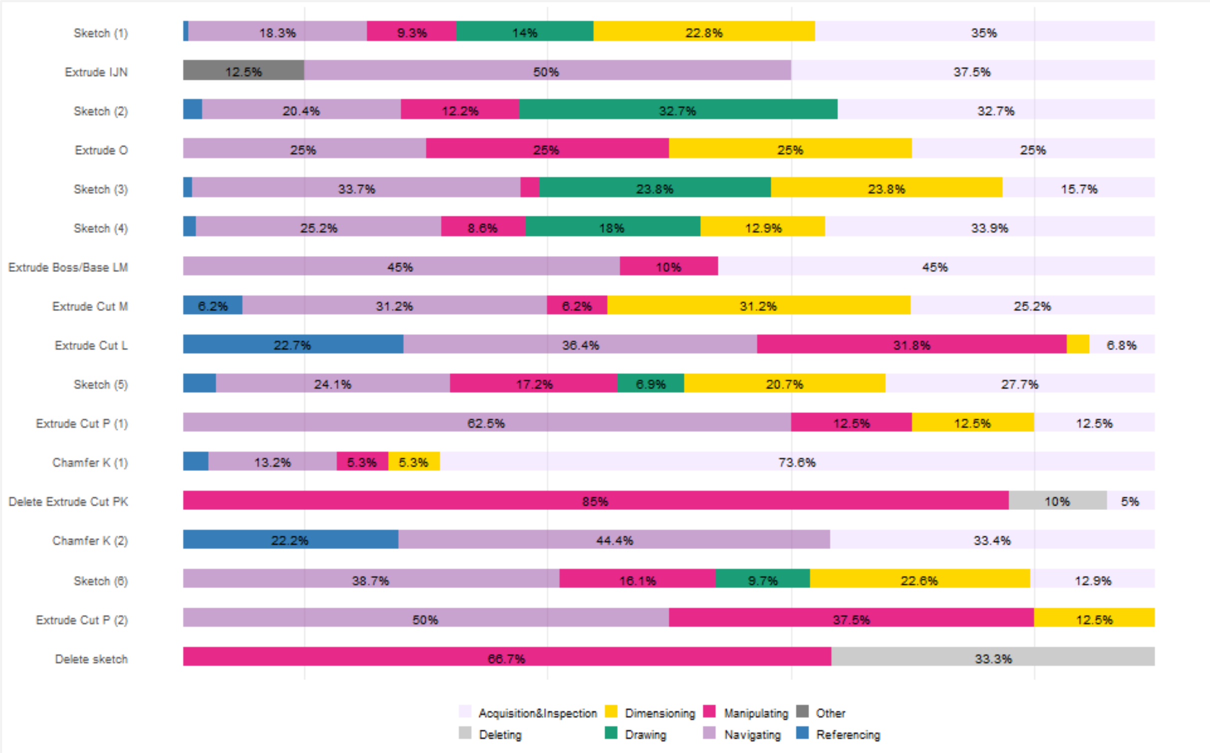

Figure 90 Low performer: Distribution of subtask clusters across unit tasks - Orthographic condition

Overall, for the Low performer, isometric condition exhibited more repeatable subtask distributions (*sketch* dominated by *acquisition* and *inspection* with *dimensioning* and *navigating*, and *feature-application* unit tasks often driven by *acquisition* and *inspection* or *navigating*), whereas orthographic condition produced more mixed within-unit-task compositions, with higher *manipulating* involvement and clear distributions related to rework (*delete* operations), consistent with a more iterative modelling process.

Subtask (CAD operation) clusters

Analysis of the cumulative distribution of subtask clusters (considering the entire CAD modelling task, not individual unit tasks) for the Low performer revealed some differences in what the modelling time was spent on between the projections (see Table 91 and Figure 91). Compared to isometric, the orthographic condition contained more *navigating* occurrences (130 vs. 60) and they were slightly longer on average. *Dimensioning* was also more frequent in the orthographic condition (40 vs. 31), with similar durations, indicating a comparable "typical" length but more frequent *dimensioning* events overall. *Drawing* increased in the orthographic condition as well (39 vs. 22), with a modest increase in mean duration. The largest increase was again observed in *manipulating*, which rose from 21 occurrences in isometric to 63 in orthographic condition, with a higher mean duration (2.05 s vs. 1.62 s) while the median remained 1 s in both conditions, suggesting that the orthographic difference was driven primarily by many more *manipulation* occurrences, rather than longer typical execution. Finally, *referencing* was more frequent in the orthographic condition (13 vs. 5) with similar typical duration. Two subtask clusters appeared only in one projection: *constraining* was observed only in the isometric, whereas *deleting* appeared only in the orthographic condition.

Table 91 Low performer: Number of occurrences and average duration of subtasks at the level of the entire task

Subtask	N of occurrences		Duration (mean) in [s]		Duration (Med) in [s]	
	Iso	Ortho	Iso	Ortho	Iso	Ortho
Navigating	60	130	1.4	1.68	1	1
Dimensioning	31	40	3.9	3.72	4	3.5
Other	2	1	5	1	5	1
Drawing	22	39	2.77	3.08	3	3
Manipulating	21	63	1.62	2.05	1	1
Constraining	1	-	4	-	4	-
Referencing	5	13	1.4	1.62	1	1
Deleting	-	3	-	2.33	-	2

When looking at how total time was distributed across subtask clusters, Figure 91 suggests that modelling from the isometric projection was dominated by *dimensioning* (22.7%), followed by *navigating* (15.8%) and *drawing* (11.4%), while *manipulating* contributed relatively little (6.4%). In the orthographic condition, *navigating* became the largest component (23.8%), *dimensioning* dropped noticeably (16.2%), *drawing* increased slightly (13.1%), and *manipulating* strongly increased (14.1%). In addition, modelling from the orthographic projection included a visible *deleting* contribution, which was absent in the isometric condition, consistent with the rework observed in the unit-task-level analyses.

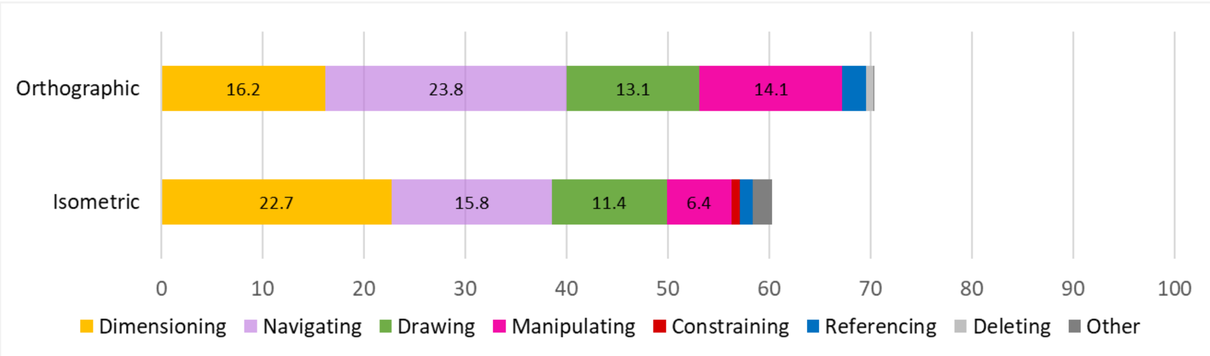

Figure 91 Low performer: Distribution of subtasks

CAD actions

The number of occurrences of almost all CAD actions differed between the conditions (see Table 92). Similar counts were observed only for *typing a value* (25 vs. 26) and *dragging a geometric entity* (absent in isometric and only one occurrence in orthographic). Overall, the Low performer executed substantially more CAD actions when modelling from the orthographic projection, while the typical (median) duration of these actions stayed largely comparable to the isometric condition.

Table 92 Low performer: Number of occurrences and average duration of CAD actions at the level of the entire task

CAD action	N of occurrences		Duration (M) in [s]		Duration (Med) in [s]	
	Iso	Ortho	Iso	Ortho	Iso	Ortho
Moving a cursor	275	543	0.88	0.86	1	1
Clicking on an icon/feature/textbox	180	312	0.001	0.001	0	0
Hovering over an icon/feature/point	157	299	0.40	0.46	0	0
Typing a value	25	26	0.58	0.19	0	0
Rotating view	6	27	2	2.30	0.5	1
Zooming in/out	30	91	0.2	0.12	0	0
Dragging a geometric entity	-	1	-	1	-	1
Entering (striking) a command	12	26	0.001	0.001	0	0

Operator utilisation throughout the task performance

Utilisations of operators are first visualised and discussed across the unit tasks and the subtasks (CAD operation) clusters.

Across the unit tasks

Figure 92 (isometric) and Figure 93 (orthographic) show P, C, and M operator utilisations across the Low performer's execution of unit tasks. In both conditions, operator utilisations depended on the unit task type: *sketch* unit tasks generally showed higher utilisation of the P and C systems, whereas several *feature-application* unit tasks (*Extrude* / *Extrude Cut* / *Chamfer*, including *delete*-related steps in orthographic) coincided with increases in M. In both conditions, the final unit tasks were performed with elevated utilisations across all three operators, indicating convergent processing demands near task completion.

Despite these shared characteristics, the dynamics of changes in utilisations differed between the projections. In the isometric condition (Figure 92), P and C remained relatively smooth and mid-to-high early in the sequence (*Sketch_1* and *Extrude ABDH_1*), after which a clear reallocation occurred around *Extrude Cut F_1*; M peaked while both P and C dropped to their lowest levels (0.3–0.4), creating a pronounced M-dominant event. Following this drop, P and C recovered gradually across *Sketch_3* → *Extrude Cut E_1*, where both operators rose to high values (0.85–0.9), before the sequence ended with moderately high, converging utilisations during *Extrude Cut G_1* as the final unit task.

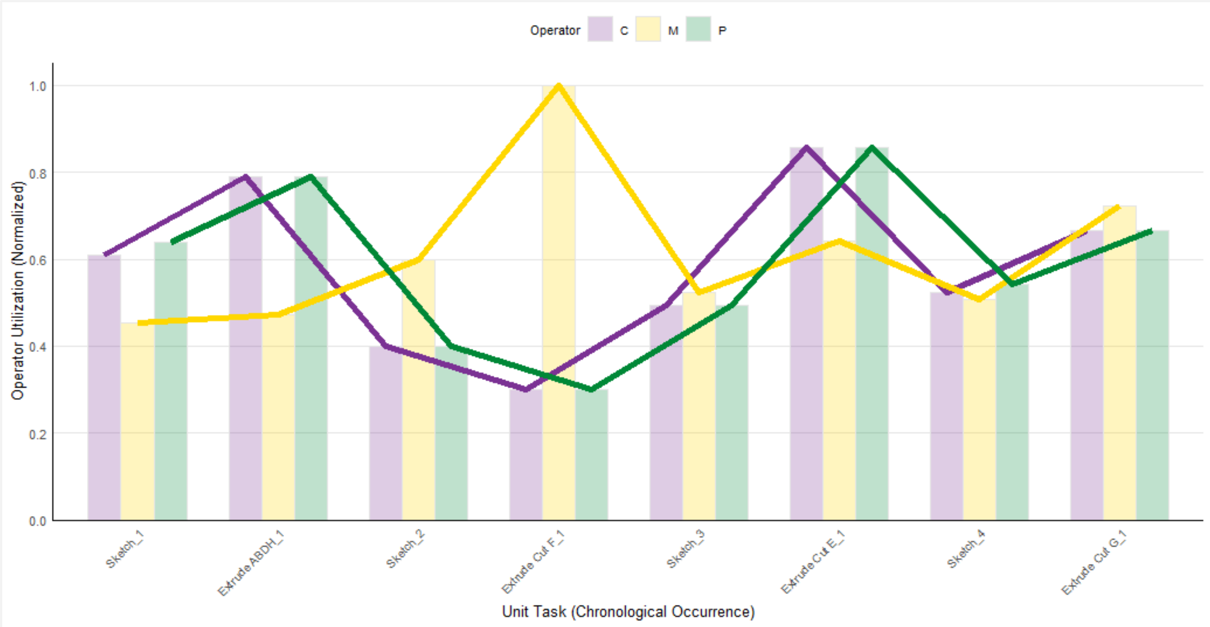

Figure 92 Low performer: Average utilisation of operators in each unit task – Isometric condition

In contrast, the orthographic condition (Figure 93) showed more frequent crossovers and larger task-to-task reallocations of utilisations, consistent with the longer and more iterative unit-task sequence. Early in the task execution, *Extrude LN_1* was characterised by a clear M increase (0.7–0.75) accompanied by a drop in P and C (0.35–0.4), followed immediately by a rebound to high P and C utilisation in *Sketch_2*, indicating rapid switching between execution-heavy and perceptual/cognitive-heavy demands. Across the mid portion of the task with orthographic condition (*Sketch_3–Sketch_5* and the subsequent *extrude/extrude cut* steps), utilisations remained mostly in a moderate band with repeated small reallocations. A notable difference emerged in the later refinement/rework phase (around *Extrude Cut P_1* → *Chamfer K_1* → *Delete Extrude Cut PK_1*), where P and C reached their highest levels (near 1.0) while M fluctuated strongly, including a marked collapse in M utilisation during the *delete* step (near zero), suggesting cognitively and perceptually dominated rework and verification. Overall, the orthographic projection elicited a less uniform changes in utilisations across unit tasks, with more repeated switching between M-dominant and P/C-dominant demands than observed in the more compact sequence when modelling from the isometric condition.

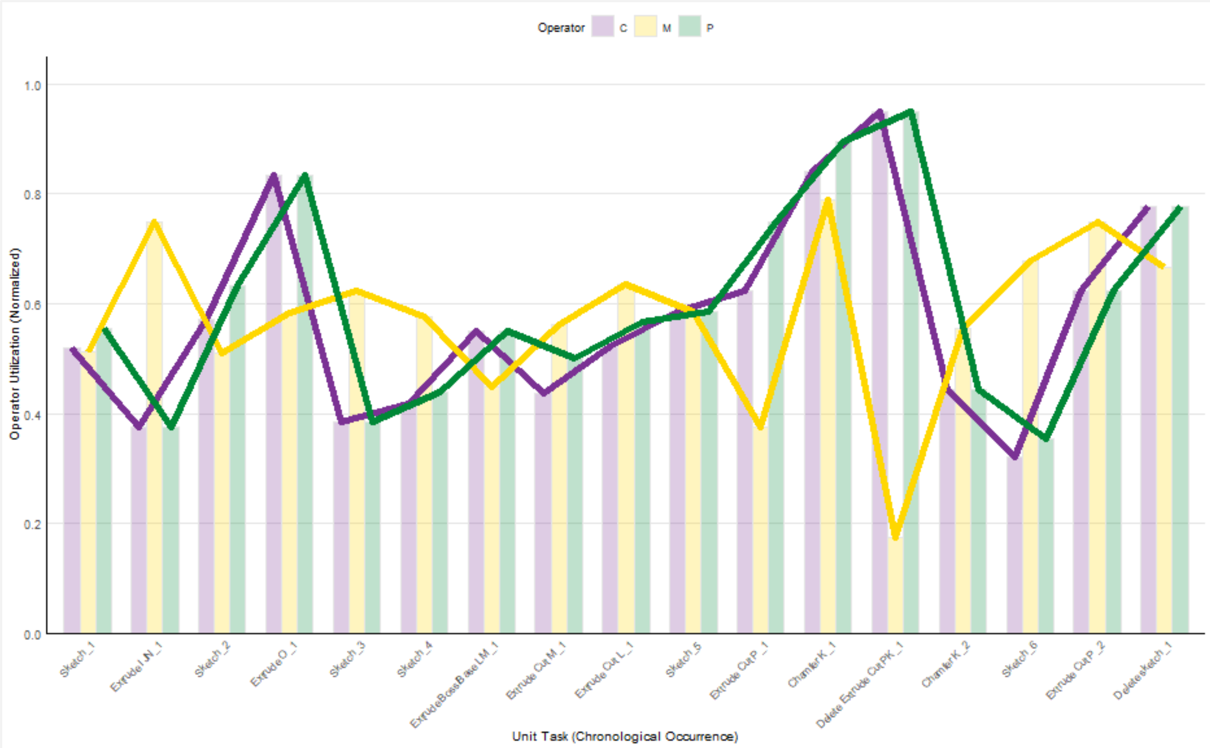

Figure 93 Low performer: Average utilisation of operators in each unit task – Orthographic condition

Across the subtask (CAD operator) clusters

The operator utilisation averages for the Low performer again showed that the dominant operator depends on the subtask type, while the projection condition shifted the balance between operator systems (Table 93).

Table 93 Low performer's Average operator utilisation across subtasks

	Navigating		**Referencing**		**Drawing**		**Constraining**		**Dimensioning**		**Manipulating**		**Deleting**	
Operator	**Iso**	**Ortho**	**Iso**	**Ortho**	**Iso**	**Ortho**	**Iso**	**Ortho**	**Iso**	**Ortho**	**Iso**	**Ortho**	**Iso**	**Ortho**
P	0.57	0.43	0.33	0.86	0.41	0.45	0.25	-	0.44	0.47	0.94	0.98	-	0.43
C	0.57	0.43	0.33	0.86	0.41	0.45	0.25	-	0.44	0.47	0.87	0.98	-	0.43
M	0.90	0.81	0.86	0.69	0.78	0.72	0.75	-	0.62	0.58	0.91	0.92	-	0.57

The highest information-processing demand in both projections occurred during *manipulating*, where utilisations were uniformly high across operators. In the isometric condition, *manipulating* already showed near-ceiling utilisations (P=0.94, C=0.87, M=0.91), and in the orthographic condition this pattern remained with P and C further increasing (P=0.98, C=0.98) while M stayed similarly high (0.92). Thus, for the Low performer, modelling from the orthographic projection amplified P and C system activity during *manipulation*, while maintaining very high M activity.

In contrast, *navigating* shifted toward a more motor-intensive execution in orthographic condition. Modelling from the orthographic projection was associated with higher M (0.81 vs. 0.70) alongside lower P and C (0.43 vs. 0.57 for both), indicating increased execution demands and reduced perceptual/cognitive utilisation during *navigation* compared to the isometric condition. *Referencing* showed the opposite pattern and the strongest projection-related shift: orthographic modelling markedly increased P and C utilisation (0.86 vs. 0.33 for both) while M decreased (0.69 vs. 0.86). This suggests that, for the Low performer, extracting information in the orthographic condition relied more on perceptual/cognitive processing and less on motor activity than in the isometric condition.

For *drawing*, modelling from the orthographic condition led to slightly higher utilisations across systems, indicating broadly similar demands with only small redistribution (a modest reduction in M alongside slightly higher P and C utilisations). *Dimensioning* also increased modestly in the orthographic condition, again suggesting similar operator demands with a small shift toward higher P and C, and slightly lower M. Finally, two subtasks appeared in only one projection: *constraining* was present only in isometric, while *deleting* appeared only in the orthographic condition, consistent with additional rework in the latter case.

Overall, for this Low performer, moving from isometric to orthographic projections was associated with more motor-intensive navigating, higher P and C demand during *manipulating*, and a strong increase in perceptual and cognitive utilisation during *referencing*, while *drawing* and *dimensioning* changed only modestly and orthographic modelling additionally included *deleting*-related rework.

Appendix B

Participant #1

Table 94 P#1: characteristics and CAD performance outputs

	Variable/Metric	Value
Experience & Expertise	Professional working experience	77
	Engineering activities	3
	Using CAD software	3
	CAD skill	5
	CAD proficiency	5
Spatial skill	Spatial visualisation	11
	Mental rotation	12
	Spatial orientation	12
CAD performance	Duration	7.39
	Dimensional accuracy	0.95
	Volume	0.97
	Surface area	0.99
	Shape	0.79
	Overall performance	0.50

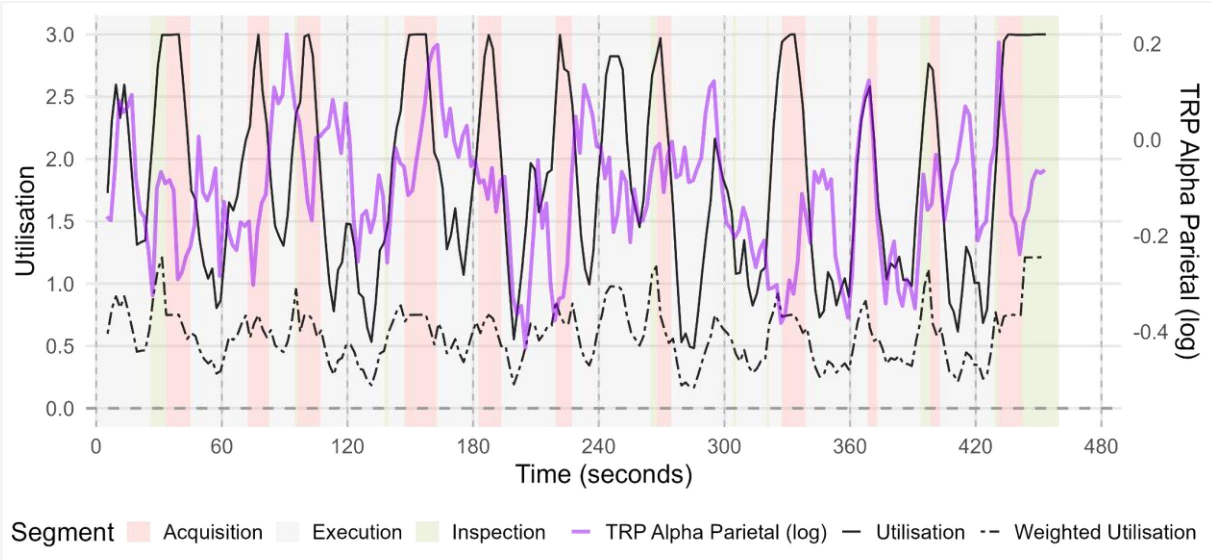

Figure 94 P#1: Dynamic changes in CL across moving windows (10s, 2s)

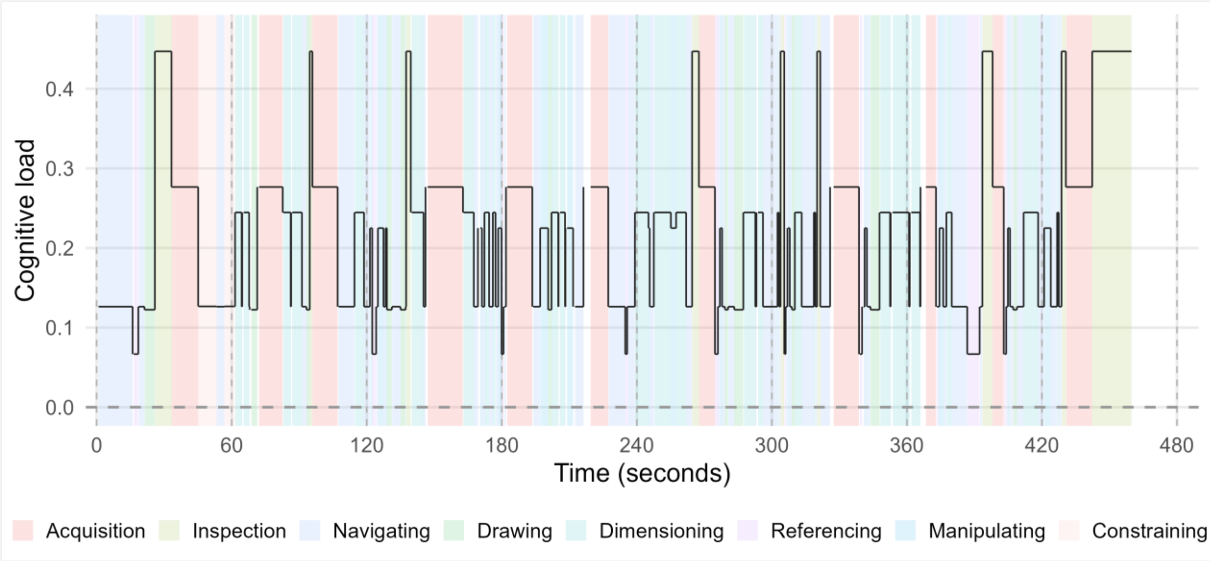

Figure 95 P#1: Dynamic changes in CL indicated by global values of weighted utilisations

Participant #3

Table 95 P#3: characteristics and CAD performance outputs

	Variable/Metric	Value
Experience & Expertise	Professional working experience	22
	Engineering activities	2
	Using CAD software	2
	CAD skill	2
	CAD proficiency	3
Spatial skill	Spatial visualisation	10
	Mental rotation	10
	Spatial orientation	11
CAD performance	Duration	12
	Dimensional accuracy	1
	Volume	0.95
	Surface area	0.98
	Shape	0.75
	Overall performance	0.31

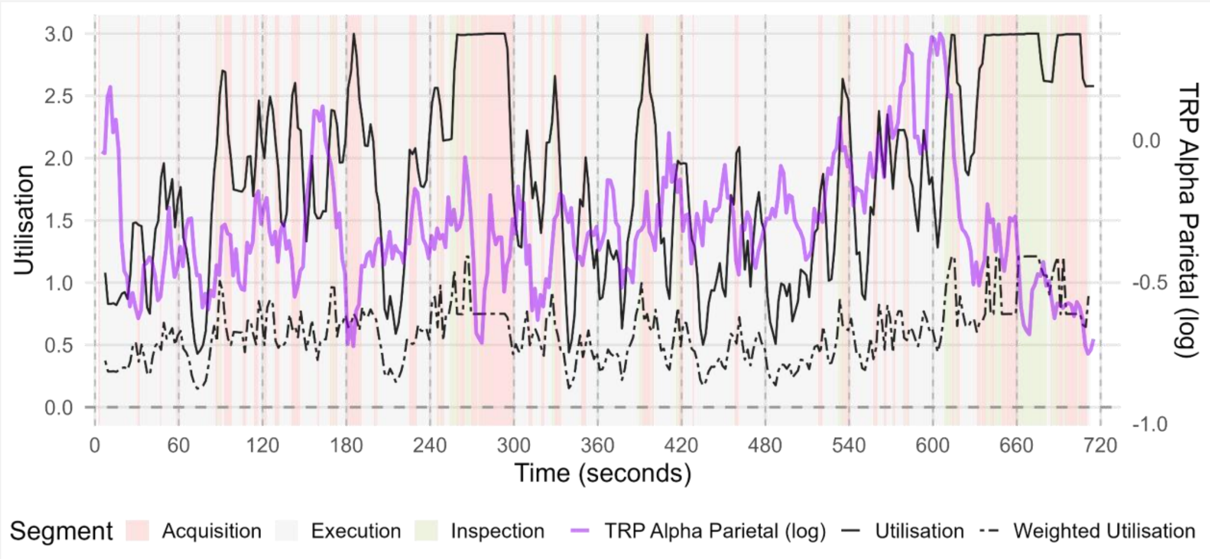

Figure 96 P#3: Dynamic changes in CL across moving windows (10s, 2s)

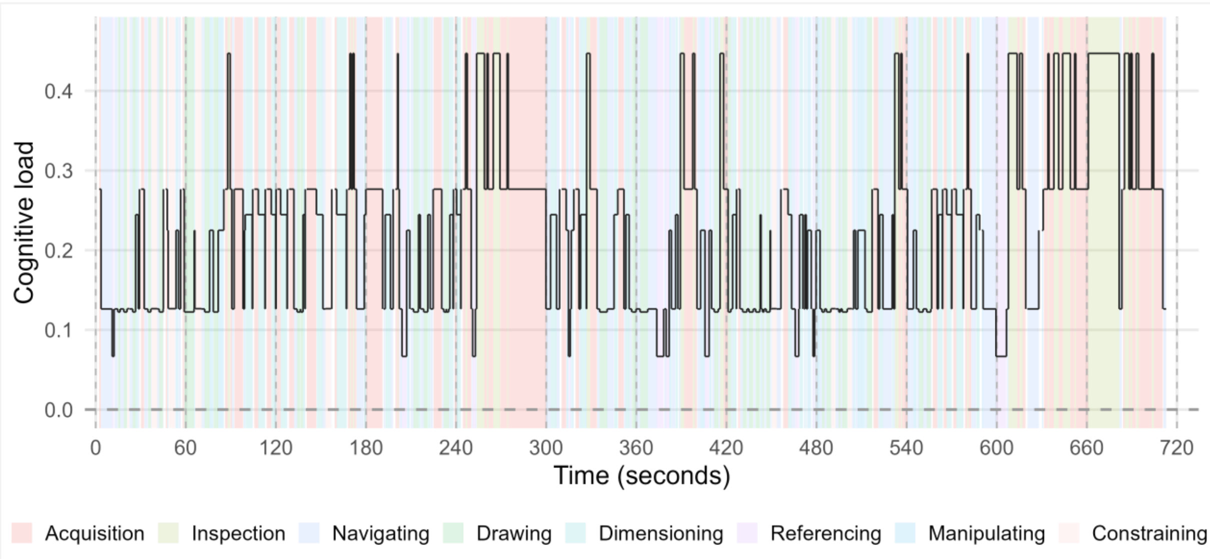

Figure 97 P#3: Dynamic changes in CL indicated by global values of weighted utilisations

Participant #4

Table 96 P#4: characteristics and CAD performance outputs

	Variable/Metric	Value
Experience & Expertise	Professional working experience	76
	Engineering activities	3
	Using CAD software	3
	CAD skill	4
	CAD proficiency	4
Spatial skill	Spatial visualisation	10
	Mental rotation	11
	Spatial orientation	12
CAD performance	Duration	7.04
	Dimensional accuracy	1
	Volume	1
	Surface area	1
	Shape	0.99
	Overall performance	0.57

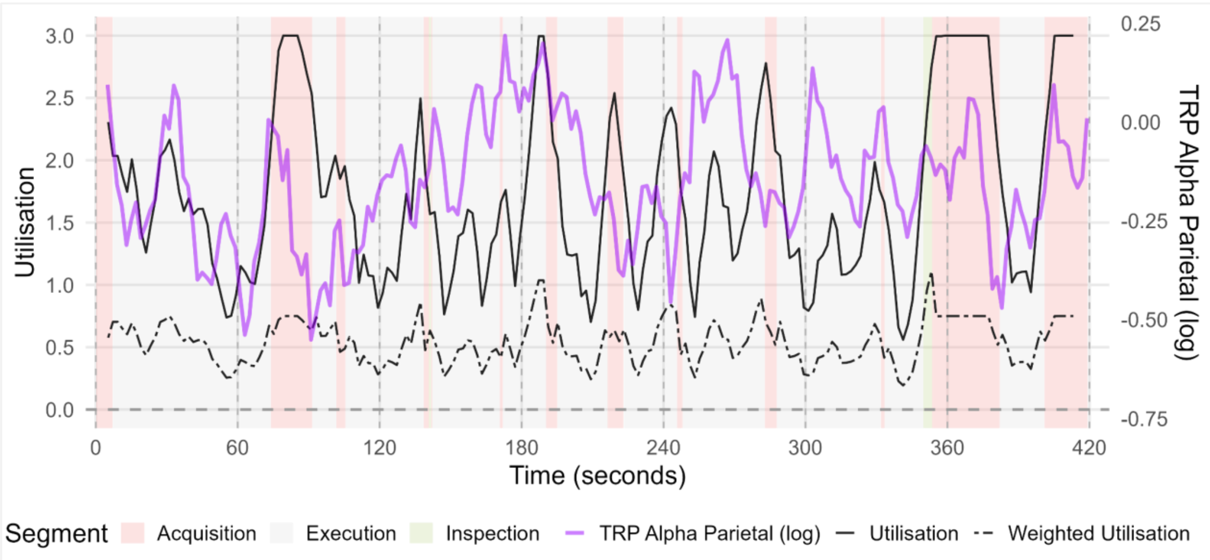

Figure 98 P#4: Dynamic changes in CL across moving windows (10s, 2s)

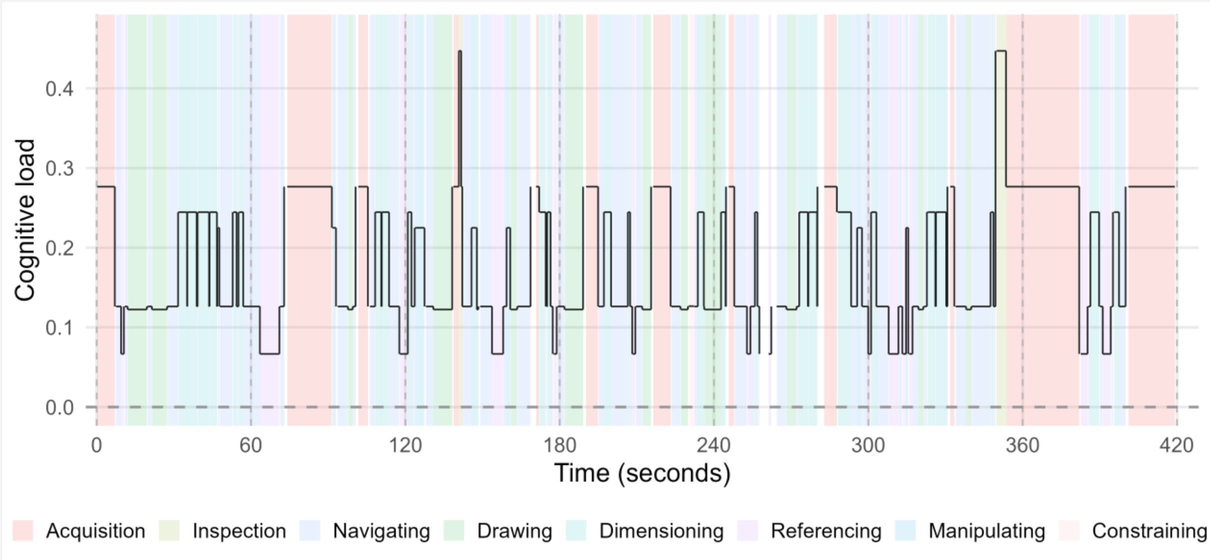

Figure 99 P#4: Dynamic changes in CL indicated by global values of weighted utilisations

Participant #5

Table 97 P#5: characteristics and CAD performance outputs

	Variable/Metric	Value
Experience & Expertise	Professional working experience	18
	Engineering activities	2
	Using CAD software	2
	CAD skill	3
	CAD proficiency	4
Spatial skill	Spatial visualisation	12
	Mental rotation	9
	Spatial orientation	11
CAD performance	Duration	7.16
	Dimensional accuracy	1
	Volume	1
	Surface area	1
	Shape	1
	Overall performance	0.56

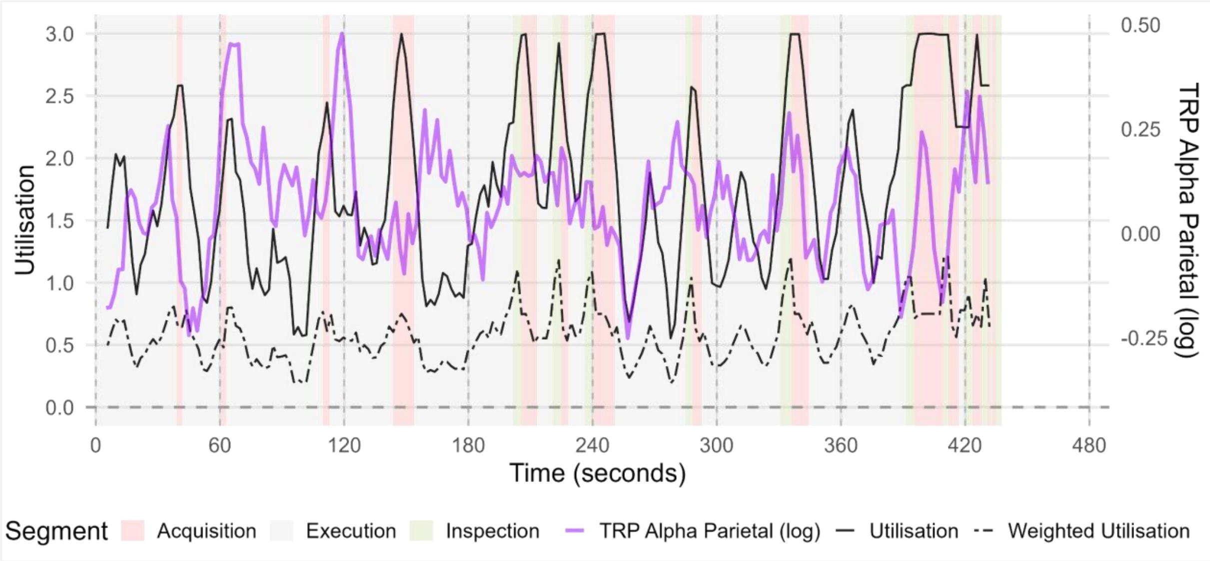

Figure 100 P#5: Dynamic changes in CL across moving windows (10s, 2s)

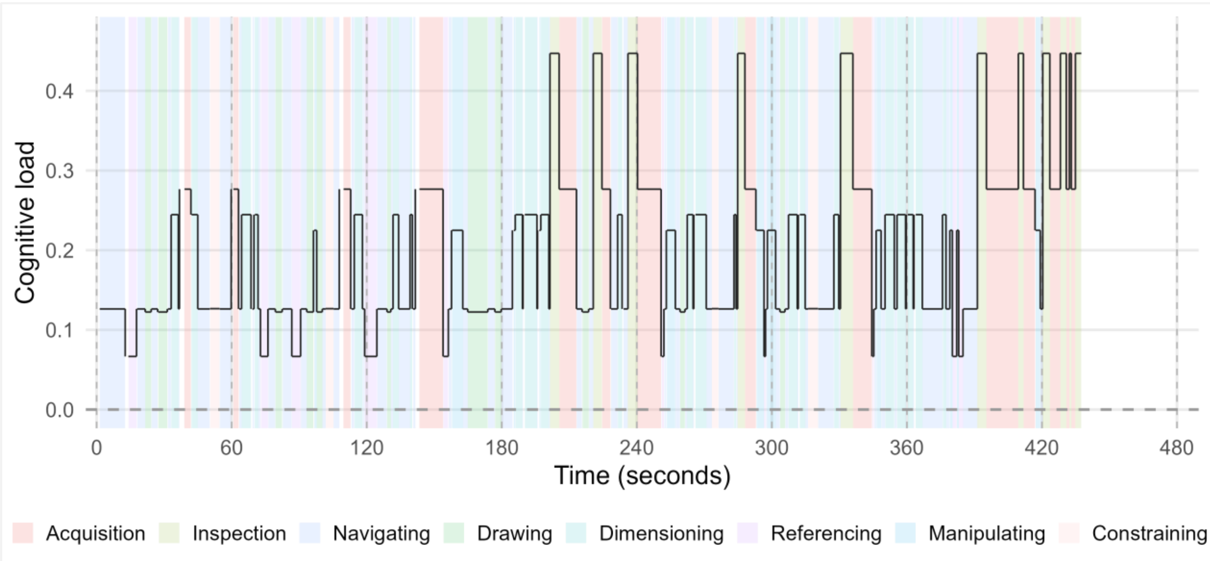

Figure 101 P#5: Dynamic changes in CL indicated by global values of weighted utilisations

Participant #6

Table 98 P#6: characteristics and CAD performance outputs

	Variable/Metric	Value
Experience & Expertise	Professional working experience	24
	Engineering activities	2
	Using CAD software	2
	CAD skill	4
	CAD proficiency	4
Spatial skill	Spatial visualisation	9
	Mental rotation	11
	Spatial orientation	8
CAD performance	Duration	8.48
	Dimensional accuracy	1
	Volume	1
	Surface area	1
	Shape	0.99
	Overall performance	0.47

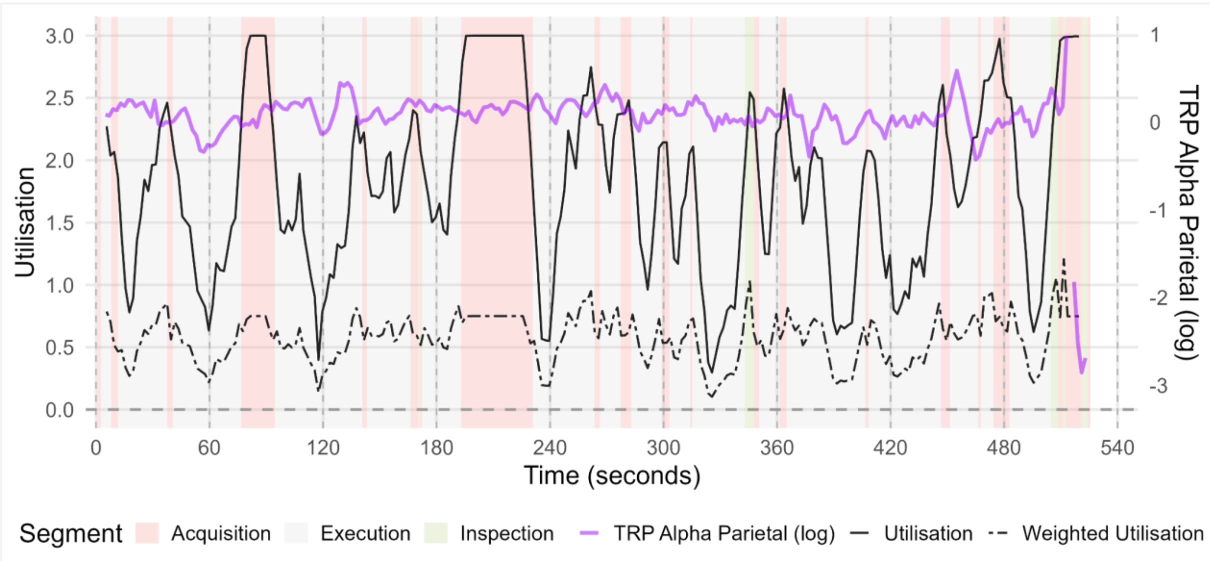

Figure 102 P#6: Dynamic changes in CL across moving windows (10s, 2s)

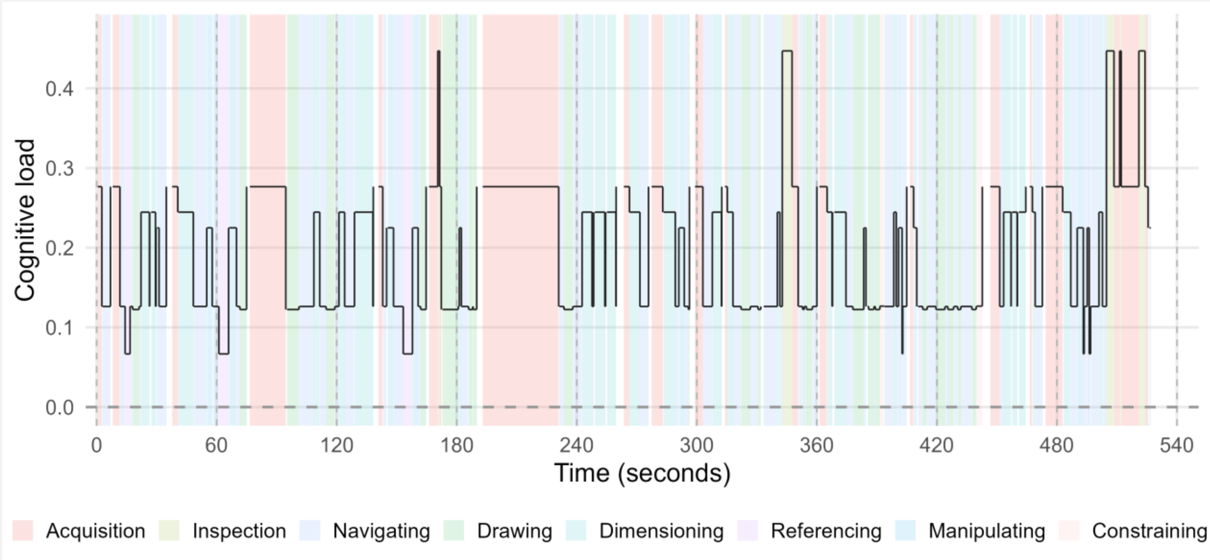

Figure 103 P#6: Dynamic changes in CL indicated by global values of weighted utilisations

Participant #7

Table 99 P#7: characteristics and CAD performance outputs

	Variable/Metric	Value
Experience & Expertise	Professional working experience	17
	Engineering activities	2
	Using CAD software	2
	CAD skill	3
	CAD proficiency	3
Spatial skill	Spatial visualisation	10
	Mental rotation	8
	Spatial orientation	12
CAD performance	Duration	15
	Dimensional accuracy	0.90
	Volume	1.13
	Surface area	1.02
	Shape	0.33
	Overall performance	0.21

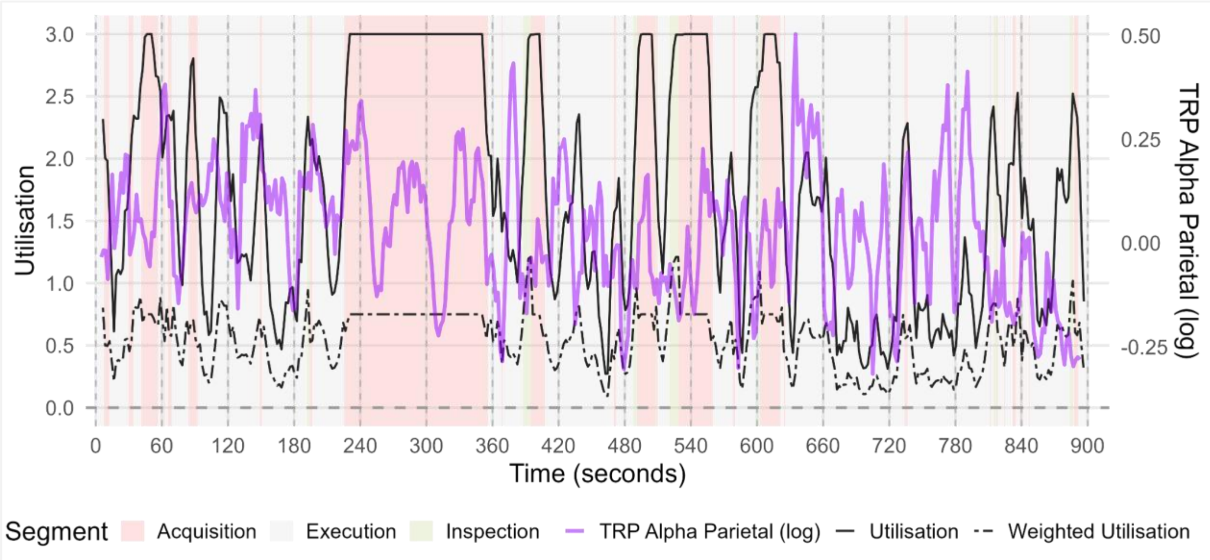

Figure 104 P#7: Dynamic changes in CL across moving windows (10s, 2s)

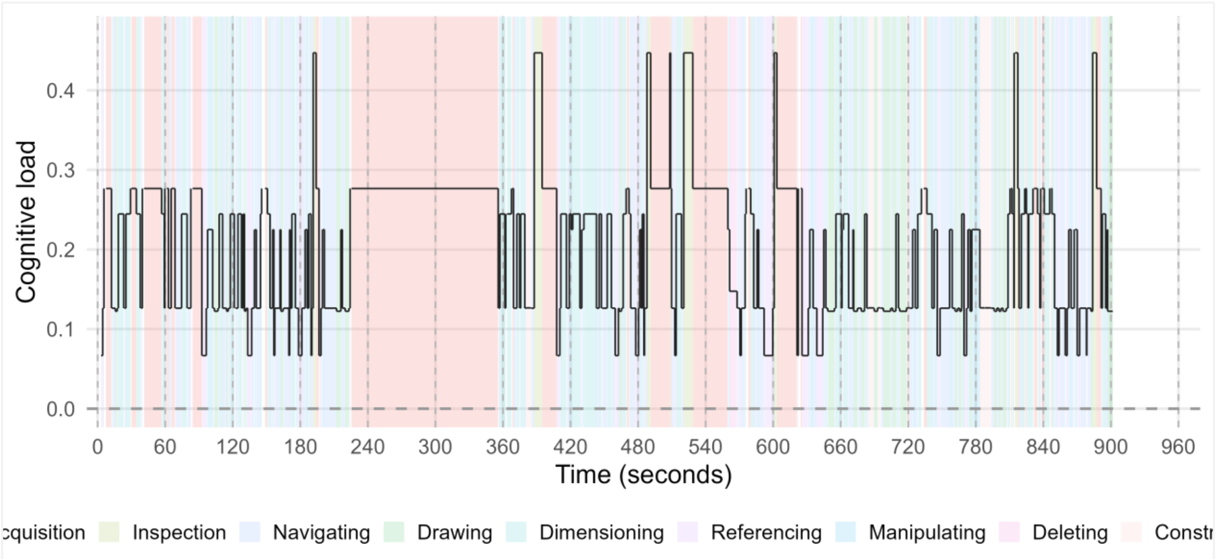

Figure 105 P#7: Dynamic changes in CL indicated by global values of weighted utilisations

Participant #8

Table 100 P#8: characteristics and CAD performance outputs

	Variable/Metric	Value
Experience & Expertise	Professional working experience	19
	Engineering activities	3
	Using CAD software	3
	CAD skill	3
	CAD proficiency	3
Spatial skill	Spatial visualisation	6
	Mental rotation	6
	Spatial orientation	11
CAD performance	Duration	15
	Dimensional accuracy	0.38
	Volume	0.98
	Surface area	0.82
	Shape	0.45
	Overall performance	0.18

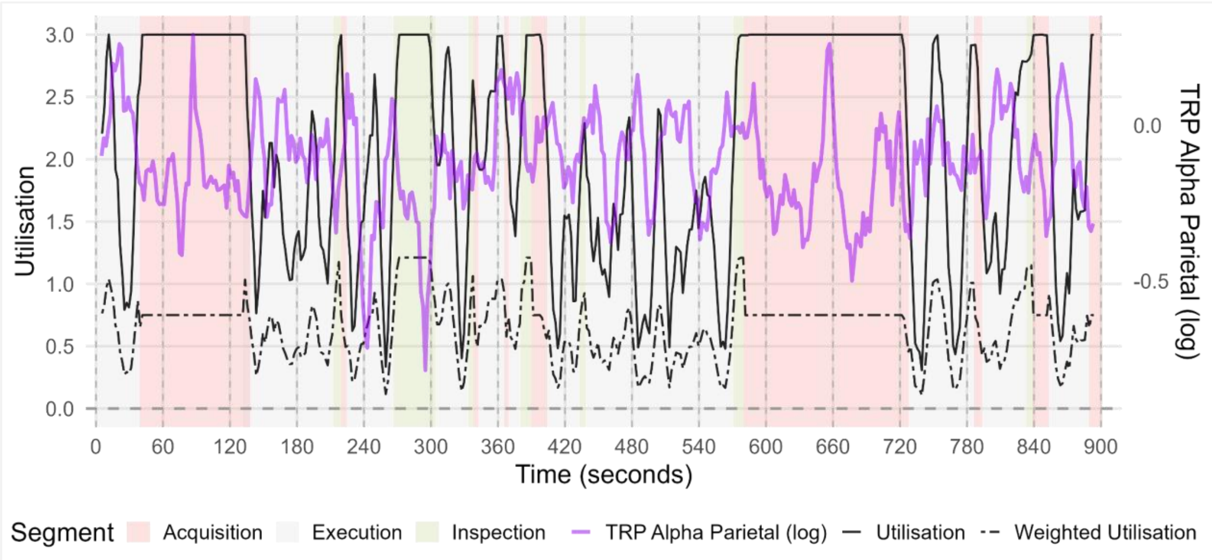

Figure 106 P#8: Dynamic changes in CL across moving windows (10s, 2s)

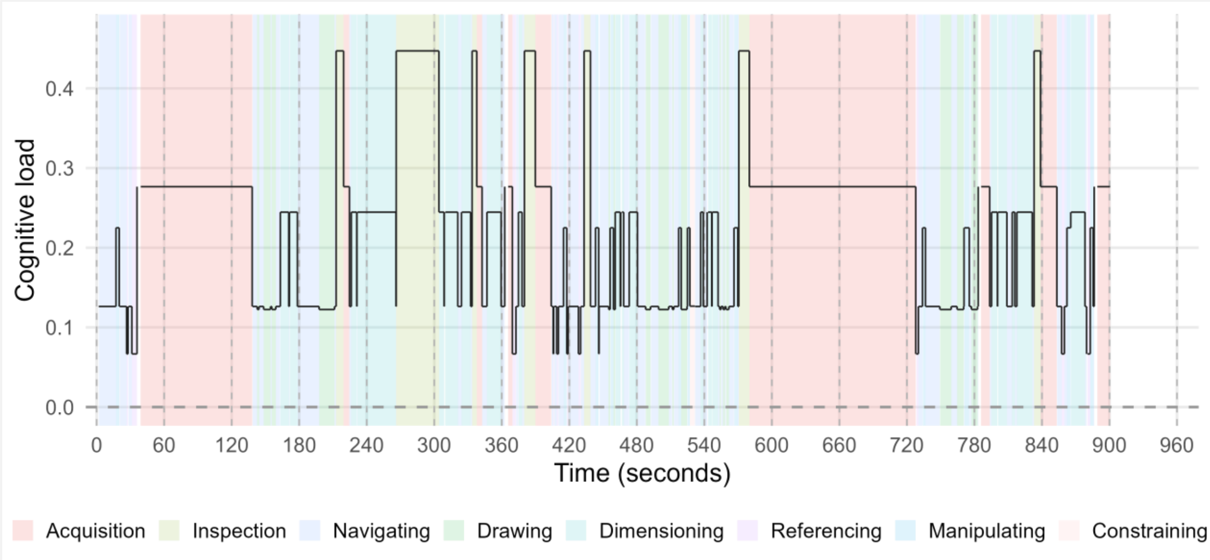

Figure 107 P#8: Dynamic changes in CL indicated by global values of weighted utilisations

Participant #10

Table 101 P#10: characteristics and CAD performance outputs

	Variable/Metric	Value
Experience & Expertise	Professional working experience	19
	Engineering activities	4
	Using CAD software	4
	CAD skill	4
	CAD proficiency	3
Spatial skill	Spatial visualisation	7
	Mental rotation	8
	Spatial orientation	9
CAD performance	Duration	15
	Dimensional accuracy	0.62
	Volume	0.98
	Surface area	0.84
	Shape	0.70
	Overall performance	0.21

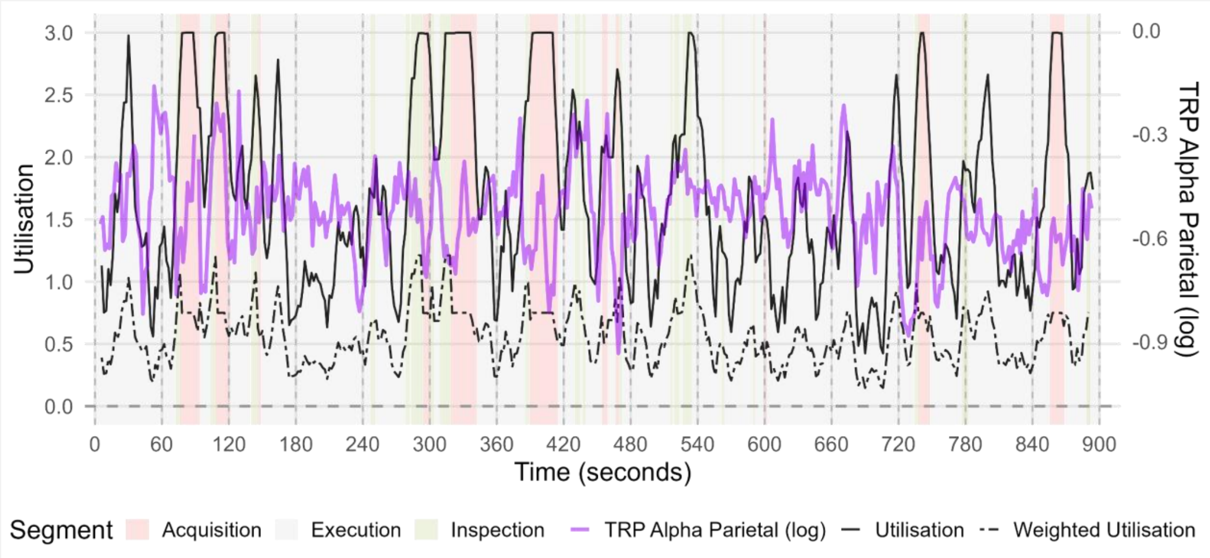

Figure 108 P#10: Dynamic changes in CL across moving windows (10s, 2s)

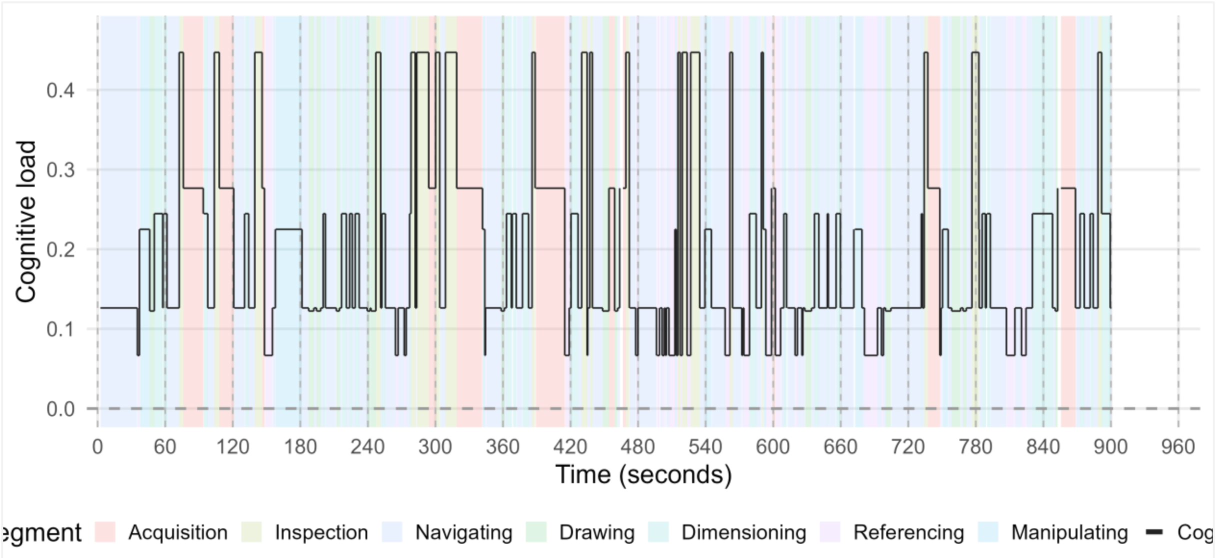

Figure 109 P#10: Dynamic changes in CL indicated by global values of weighted utilisations

Participant #11

Table 102 P#11: characteristics and CAD performance outputs

	Variable/Metric	Value
Experience & Expertise	Professional working experience	20
	Engineering activities	3
	Using CAD software	4
	CAD skill	3
	CAD proficiency	3
Spatial skill	Spatial visualisation	9
	Mental rotation	7
	Spatial orientation	10
CAD performance	Duration	10.59
	Dimensional accuracy	1
	Volume	1
	Surface area	1
	Shape	1
	Overall performance	0.38

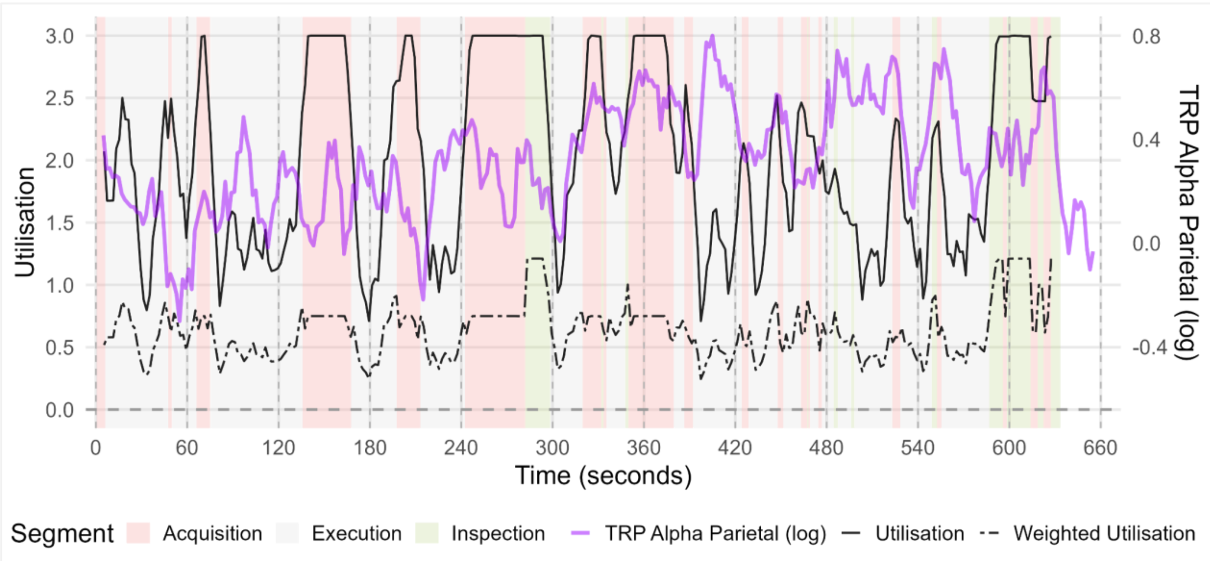

Figure 110 P#11: Dynamic changes in CL across moving windows (10s, 2s)

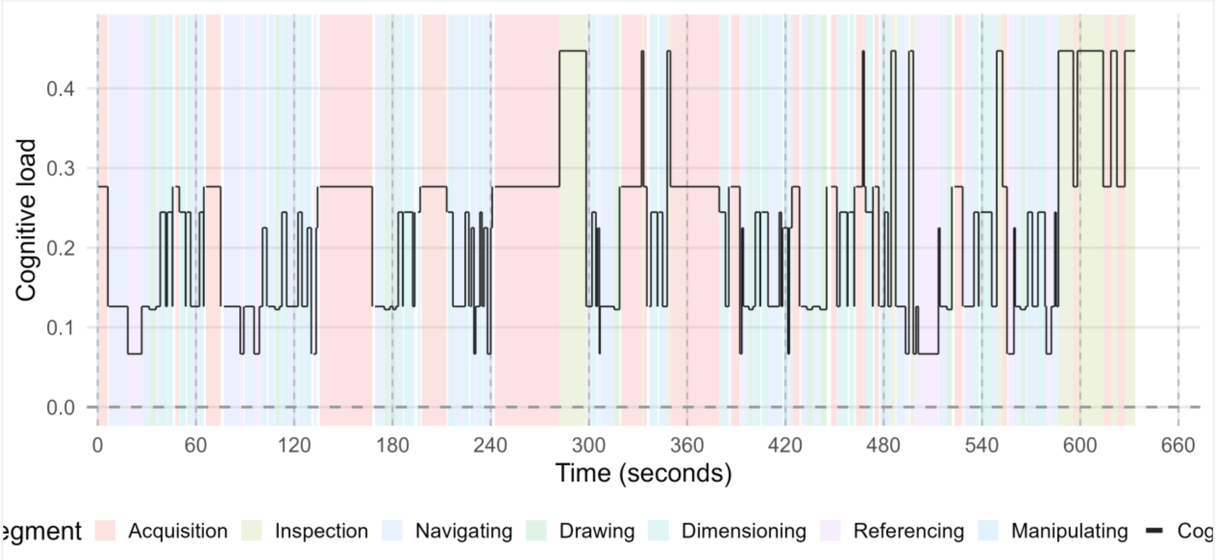

Figure 111 P#11: Dynamic changes in CL indicated by global values of weighted utilisations

Participant #12

Table 103 P#12: characteristics and CAD performance outputs

	Variable/Metric	Value
Experience & Expertise	Professional working experience	22
	Engineering activities	2
	Using CAD software	2
	CAD skill	3
	CAD proficiency	1
Spatial skill	Spatial visualisation	5
	Mental rotation	10
	Spatial orientation	12
CAD performance	Duration	15
	Dimensional accuracy	0.24
	Volume	0.90
	Surface area	0.70
	Shape	0.65
	Overall performance	0.17

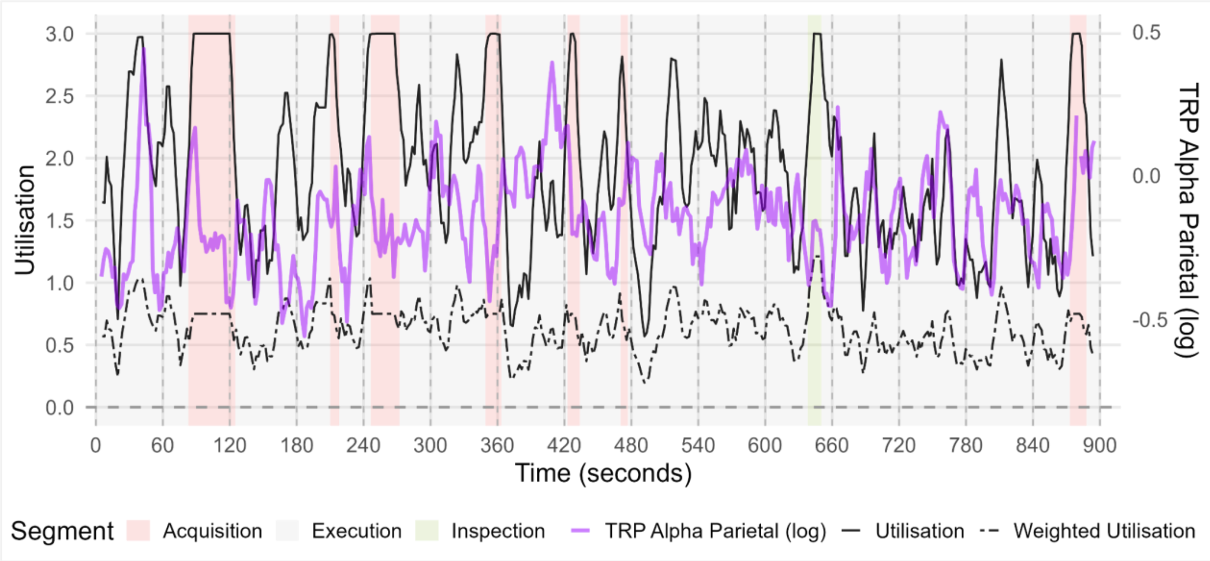

Figure 112 P#12: Dynamic changes in CL across moving windows (10s, 2s)

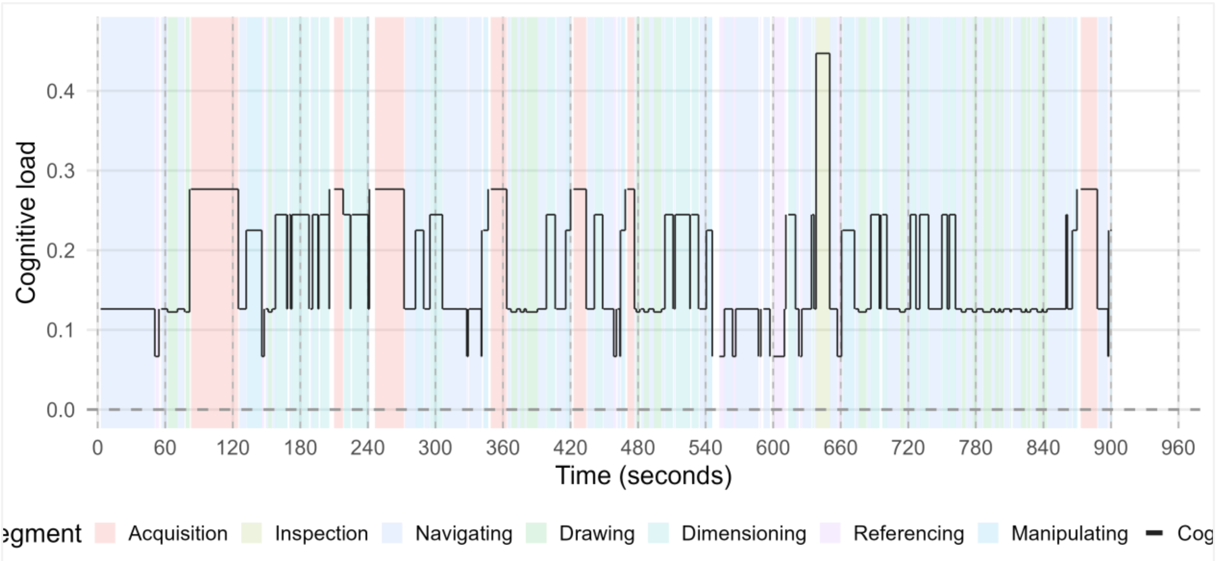

Figure 113 P#12: Dynamic changes in CL indicated by global values of weighted utilisations

Participant #13

Table 104 P#13: characteristics and CAD performance outputs

	Variable/Metric	Value
Experience & Expertise	Professional working experience	18
	Engineering activities	1
	Using CAD software	1
	CAD skill	2
	CAD proficiency	2
Spatial skill	Spatial visualisation	9
	Mental rotation	9
	Spatial orientation	12
CAD performance	Duration	15
	Dimensional accuracy	0.52
	Volume	1.22
	Surface area	0.97
	Shape	0.66
	Overall performance	0.20

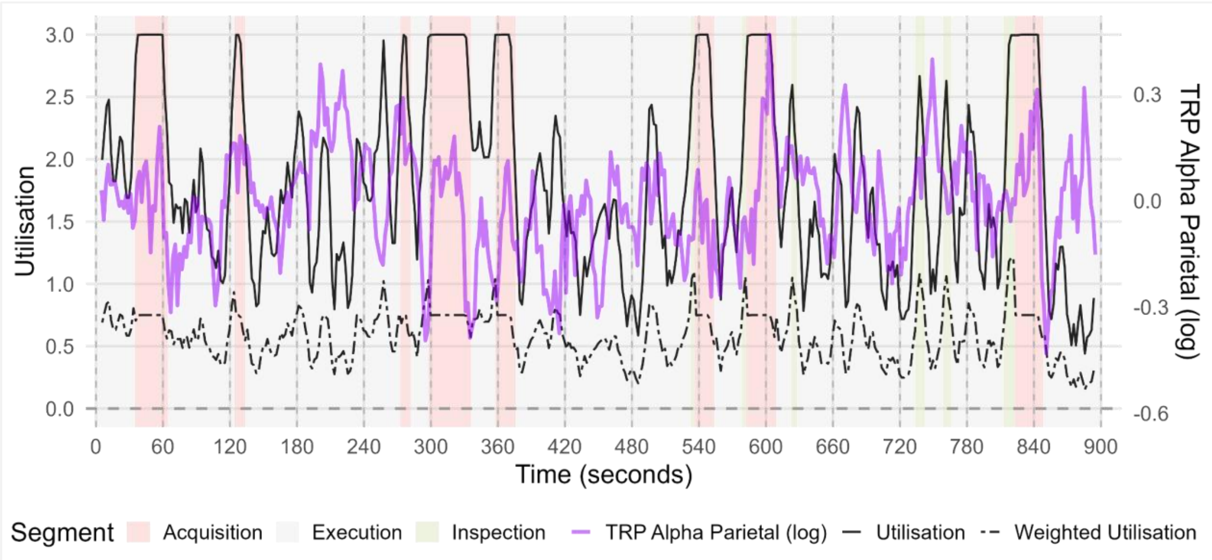

Figure 114: P#13: Dynamic changes in CL across moving windows (10s, 2s)

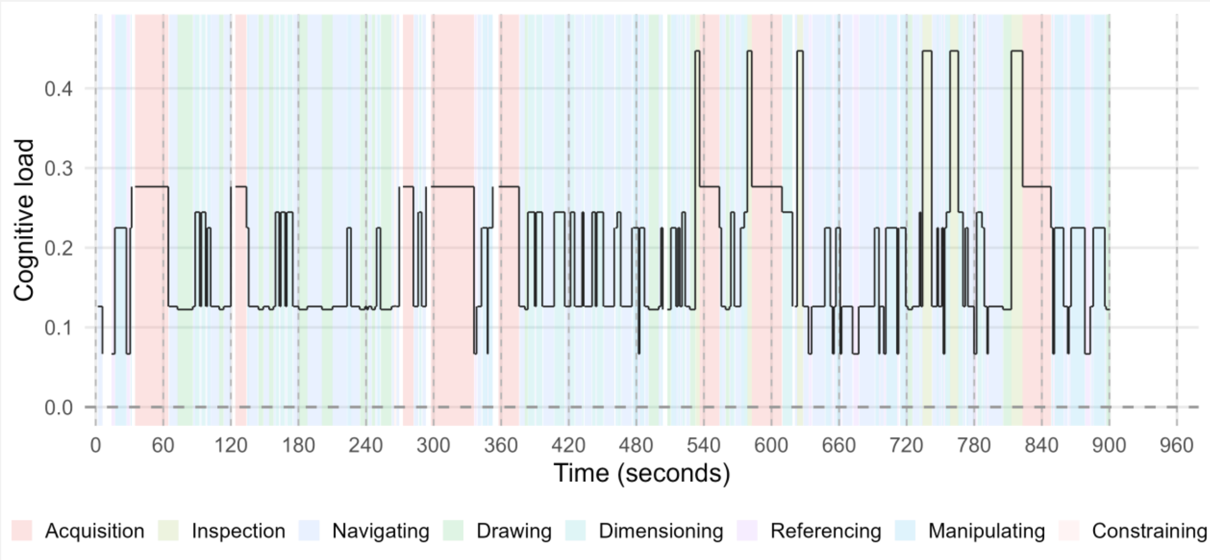

Figure 115 P#13: Dynamic changes in CL indicated by global values of weighted utilisations

Participant #14

Table 105 P#14: characteristics and CAD performance outputs

	Variable/Metric	Value
Experience & Expertise	Professional working experience	36
	Engineering activities	3
	Using CAD software	5
	CAD skill	4
	CAD proficiency	4
Spatial skill	Spatial visualisation	6
	Mental rotation	8
	Spatial orientation	11
CAD performance	Duration	10.35
	Dimensional accuracy	0.95
	Volume	1
	Surface area	1
	Shape	1
	Overall performance	0.38

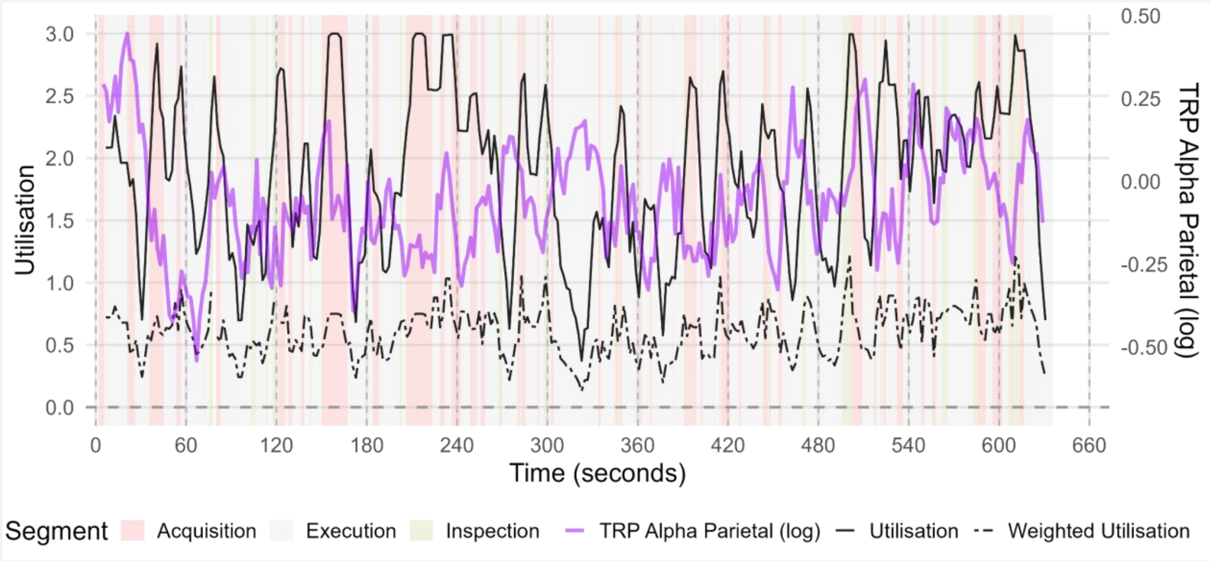

Figure 116 P#14: Dynamic changes in CL across moving windows (10s, 2s)

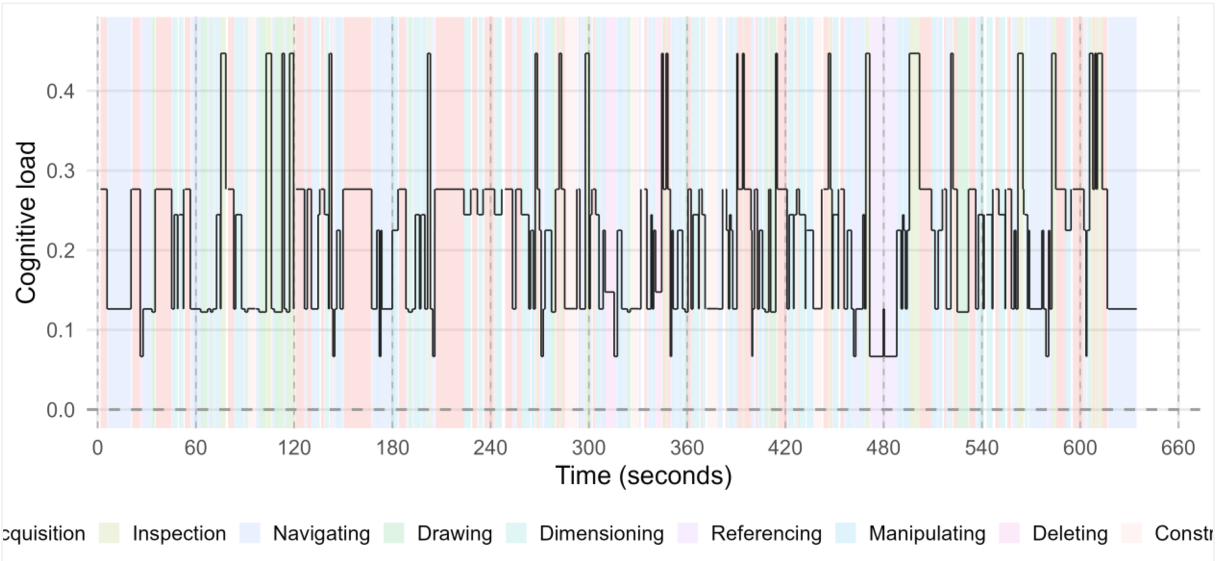

Figure 117 P#14: Dynamic changes in CL indicated by global values of weighted utilisations

Participant #15

Table 106 P#15: characteristics and CAD performance outputs

	Variable/Metric	Value
Experience & Expertise	Professional working experience	32
	Engineering activities	5
	Using CAD software	5
	CAD skill	5
	CAD proficiency	4
Spatial skill	Spatial visualisation	9
	Mental rotation	7
	Spatial orientation	8
CAD performance	Duration	11.34
	Dimensional accuracy	0.95
	Volume	1
	Surface area	1
	Shape	0.94
	Overall performance	0.34

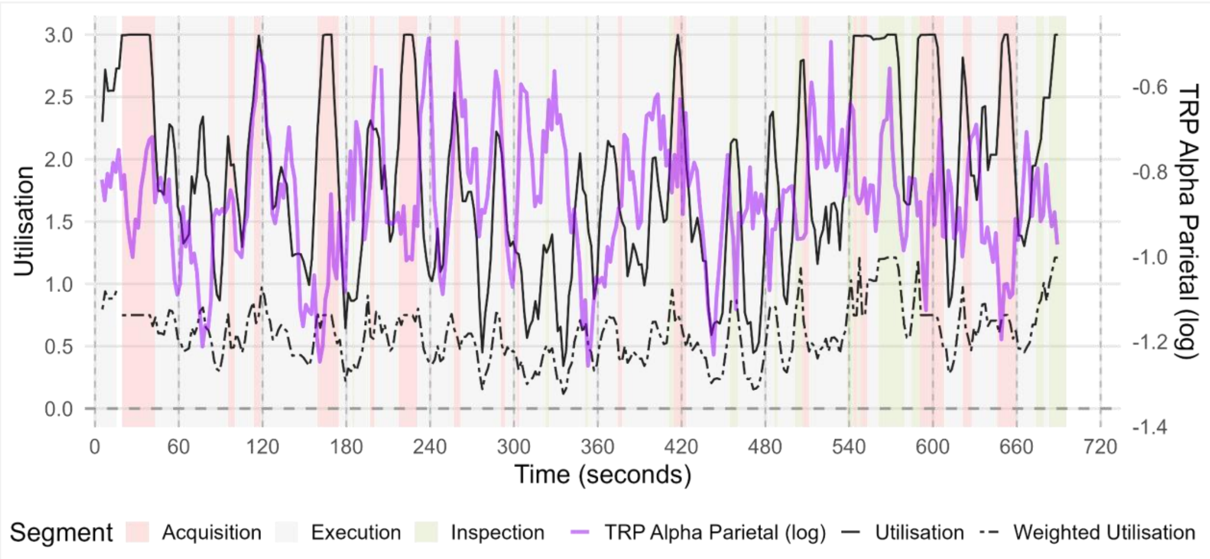

Figure 118 P#15: Dynamic changes in CL across moving windows (10s, 2s)

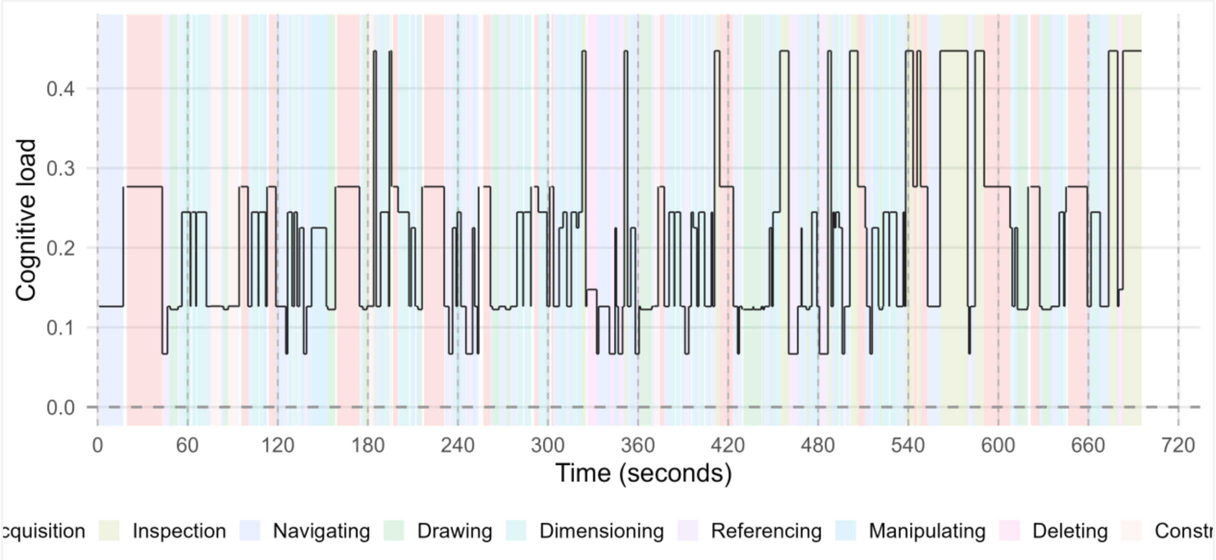

Figure 119 P#15: Dynamic changes in CL indicated by global values of weighted utilisations

Participant #16

Table 107 P#16: characteristics and CAD performance outputs

	Variable/Metric	Value
Experience & Expertise	Professional working experience	14
	Engineering activities	3
	Using CAD software	4
	CAD skill	3
	CAD proficiency	3
Spatial skill	Spatial visualisation	8
	Mental rotation	7
	Spatial orientation	8
CAD performance	Duration	12.54
	Dimensional accuracy	0.90
	Volume	1.01
	Surface area	1.01
	Shape	0.85
	Overall performance	0.30

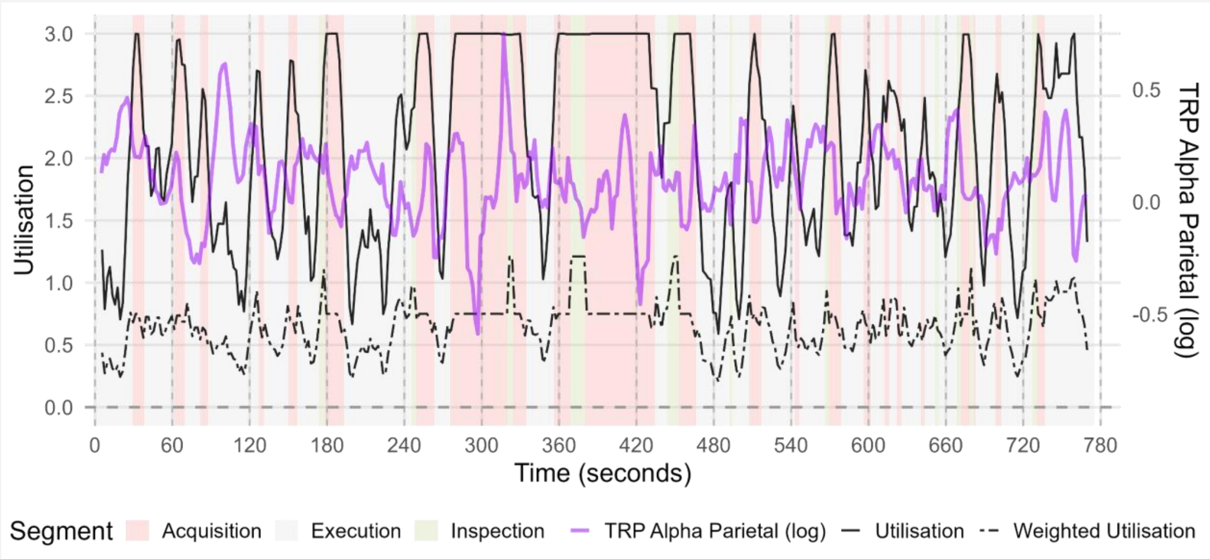

Figure 120 P#16: Dynamic changes in CL across moving windows (10s, 2s)

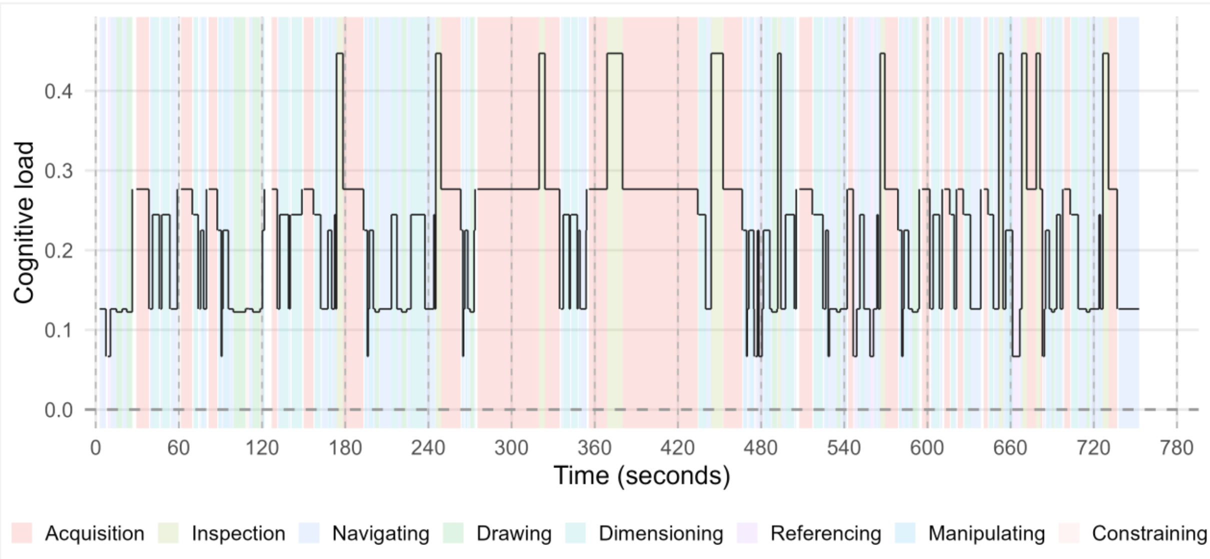

Figure 121 P#16: Dynamic changes in CL indicated by global values of weighted utilisations

Participant #17

Table 108 P#17: characteristics and CAD performance outputs

	Variable/Metric	Value
Experience & Expertise	Professional working experience	48
	Engineering activities	2
	Using CAD software	2
	CAD skill	2
	CAD proficiency	2
Spatial skill	Spatial visualisation	11
	Mental rotation	5
	Spatial orientation	6
CAD performance	Duration	13.38
	Dimensional accuracy	0.86
	Volume	0.93
	Surface area	0.93
	Shape	0.83
	Overall performance	0.27

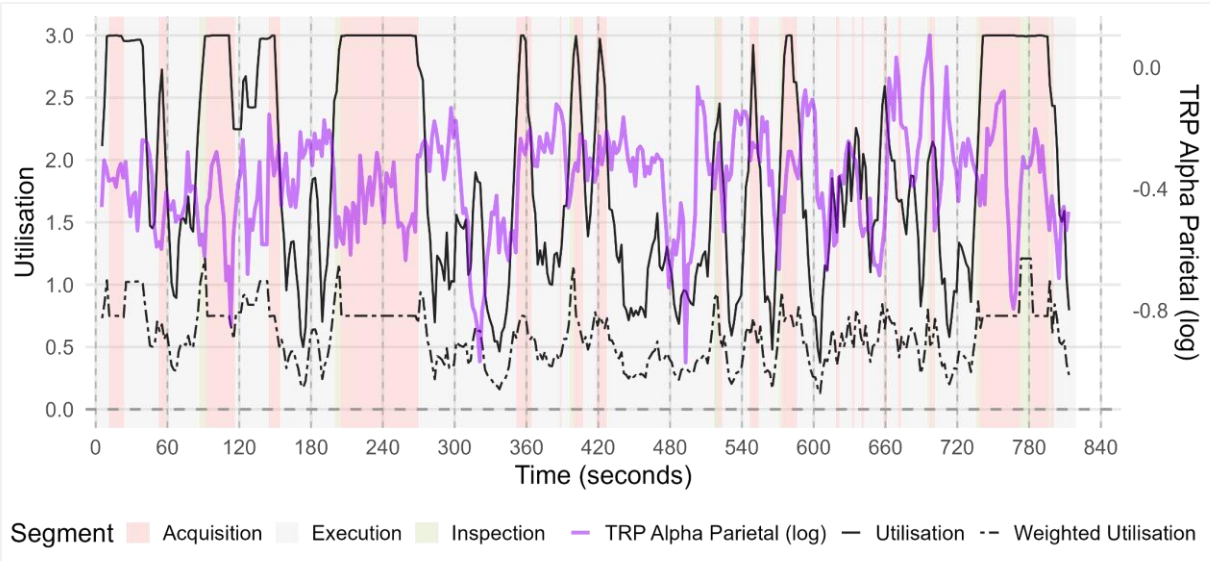

Figure 122 P#17: Dynamic changes in CL across moving windows (10s, 2s)

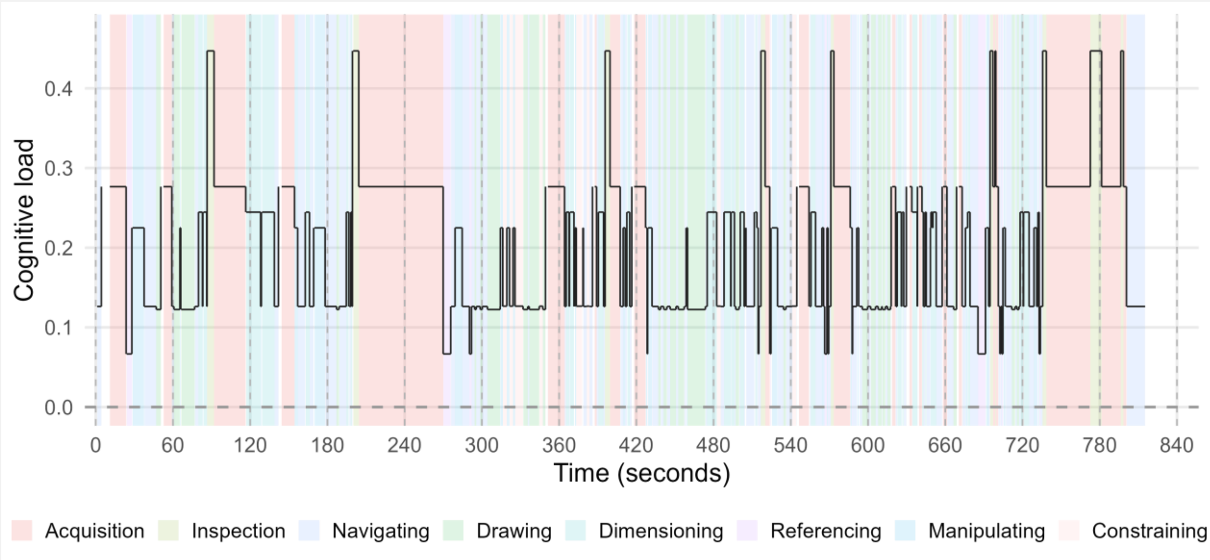

Figure 123 P#17: Dynamic changes in CL indicated by global values of weighted utilisations

Participant #18

Table 109 P#18: characteristics and CAD performance outputs

	Variable/Metric	Value
Experience & Expertise	Professional working experience	30
	Engineering activities	2
	Using CAD software	2
	CAD skill	3
	CAD proficiency	3
Spatial skill	Spatial visualisation	11
	Mental rotation	7
	Spatial orientation	8
CAD performance	Duration	13.25
	Dimensional accuracy	1
	Volume	1
	Surface area	1
	Shape	1
	Overall performance	0.30

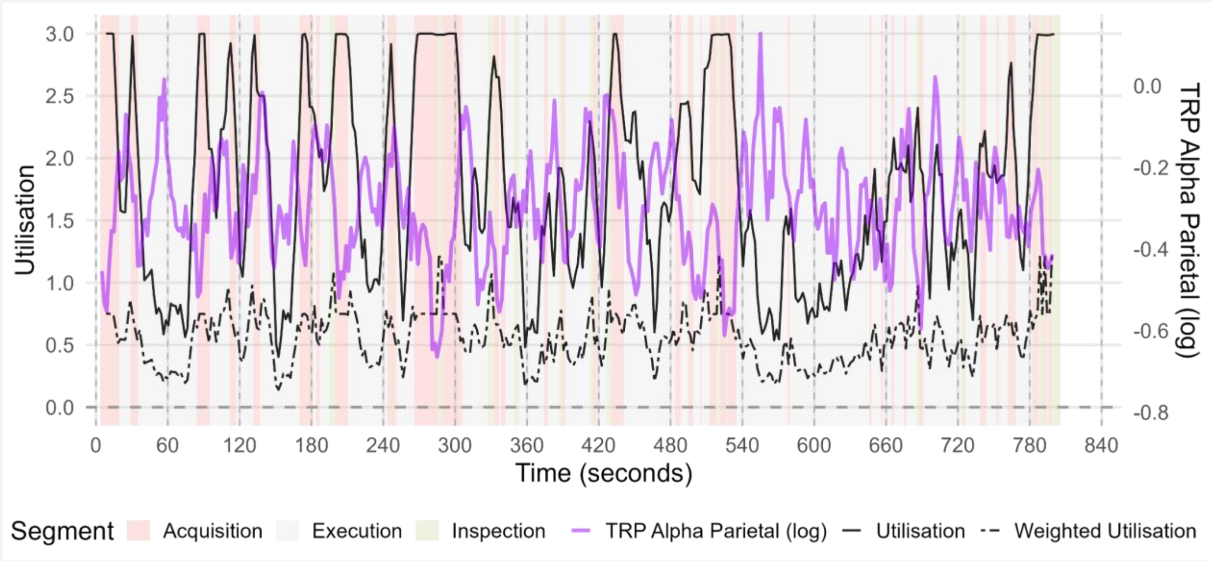

Figure 124 P#18: Dynamic changes in CL across moving windows (10s, 2s)

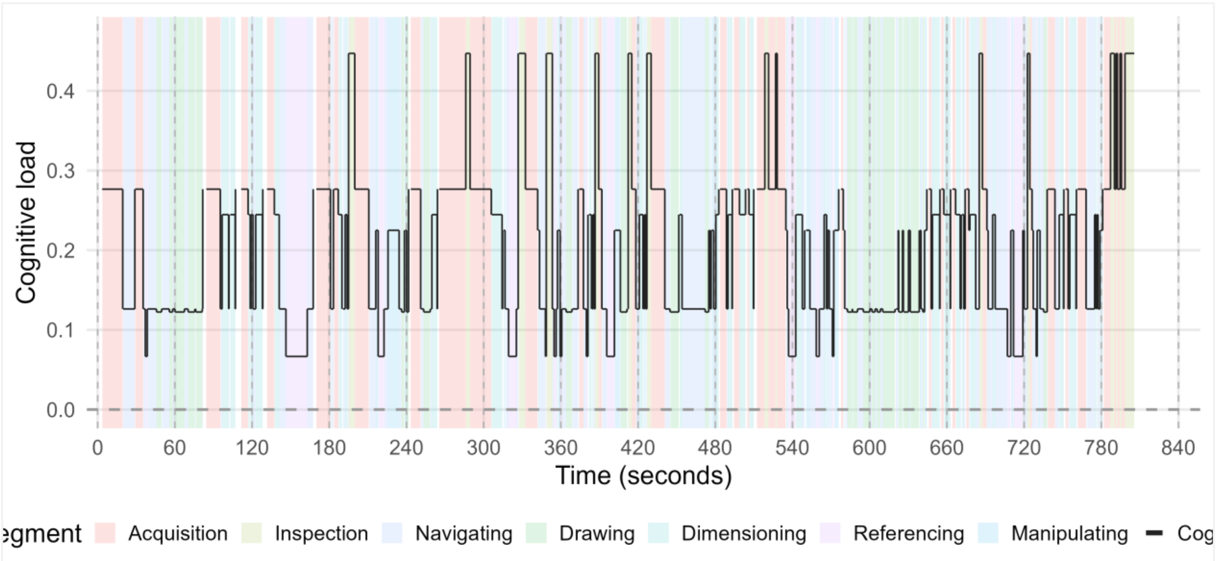

Figure 125 P#18: Dynamic changes in CL indicated by global values of weighted utilisations

Participant #19

Table 110 P#19: characteristics and CAD performance outputs

	Variable/Metric	Value
Experience & Expertise	Professional working experience	12
	Engineering activities	2
	Using CAD software	5
	CAD skill	4
	CAD proficiency	4
Spatial skill	Spatial visualisation	8
	Mental rotation	7
	Spatial orientation	8
CAD performance	Duration	14.21
	Dimensional accuracy	0.95
	Volume	1.01
	Surface area	1.01
	Shape	0.99
	Overall performance	0.28

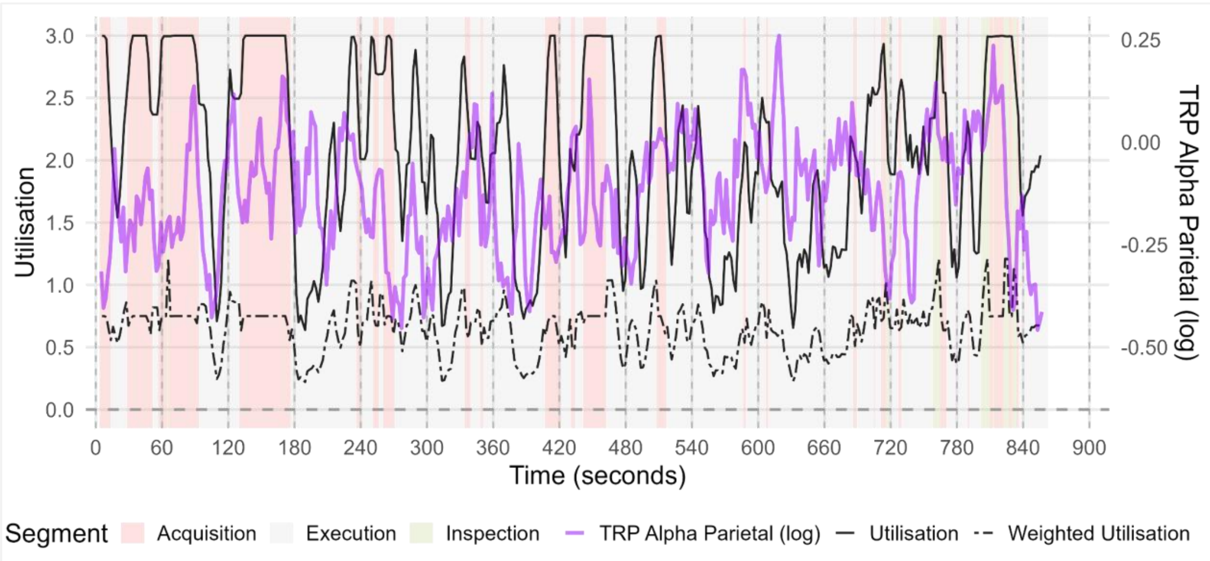

Figure 126 P#19: Dynamic changes in CL across moving windows (10s, 2s)

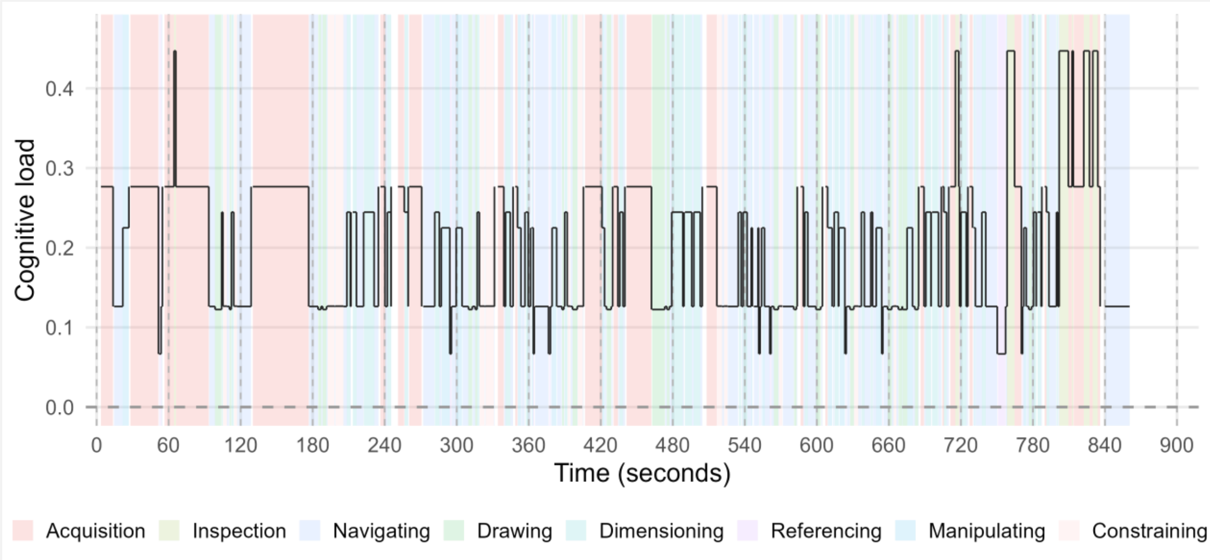

Figure 127 P#19: Dynamic changes in CL indicated by global values of weighted utilisations

Participant #20

Table 111 P#20: characteristics and CAD performance outputs

	Variable/Metric	Value
Experience & Expertise	Professional working experience	15
	Engineering activities	2
	Using CAD software	2
	CAD skill	3
	CAD proficiency	4
Spatial skill	Spatial visualisation	11
	Mental rotation	5
	Spatial orientation	6
CAD performance	Duration	12.35
	Dimensional accuracy	0.95
	Volume	1.01
	Surface area	1.01
	Shape	0.87
	Overall performance	0.31

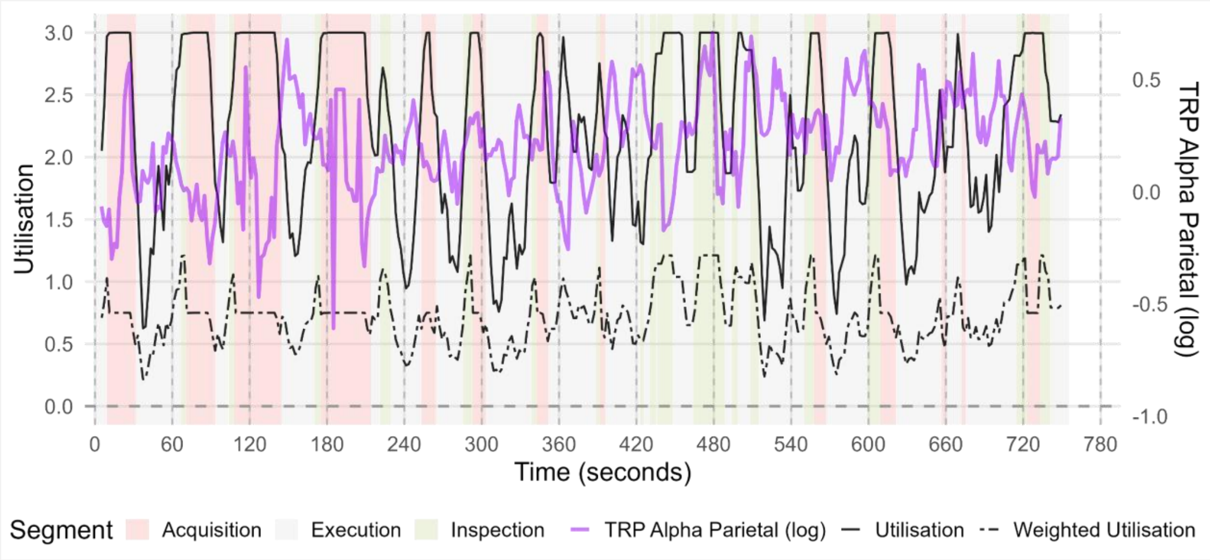

Figure 128 P#20: Dynamic changes in CL across moving windows (10s, 2s)

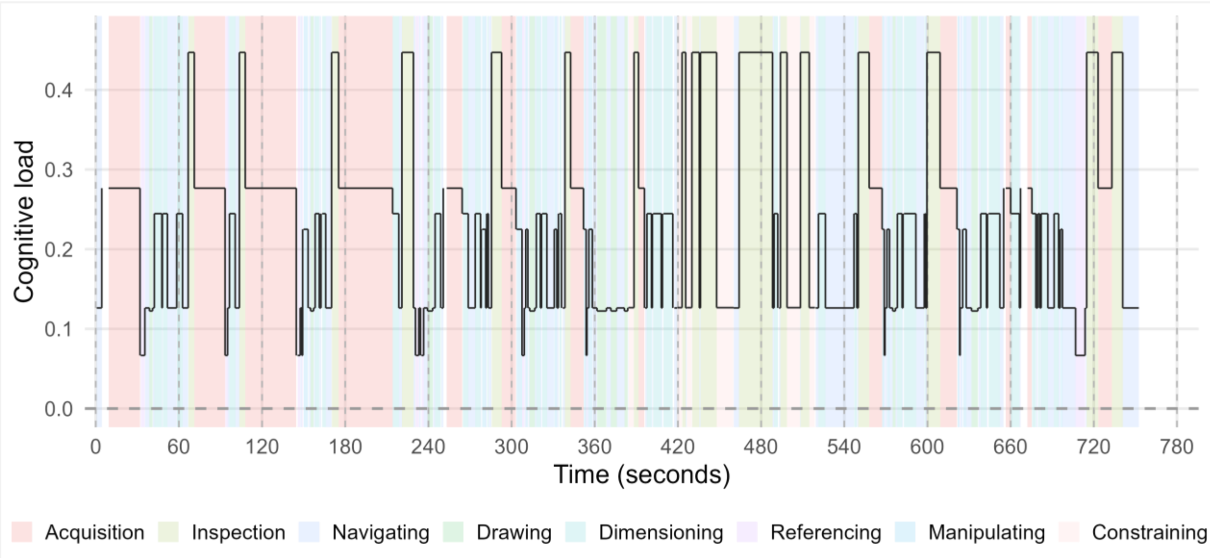

Figure 129 P#20: Dynamic changes in CL indicated by global values of weighted utilisations

Participant #21

Table 112 P#21: characteristics and CAD performance outputs

	Variable/Metric	Value
Experience & Expertise	Professional working experience	60
	Engineering activities	3
	Using CAD software	3
	CAD skill	4
	CAD proficiency	4
Spatial skill	Spatial visualisation	5
	Mental rotation	7
	Spatial orientation	12
CAD performance	Duration	10.25
	Dimensional accuracy	0.95
	Volume	1.01
	Surface area	1.01
	Shape	0.89
	Overall performance	0.37

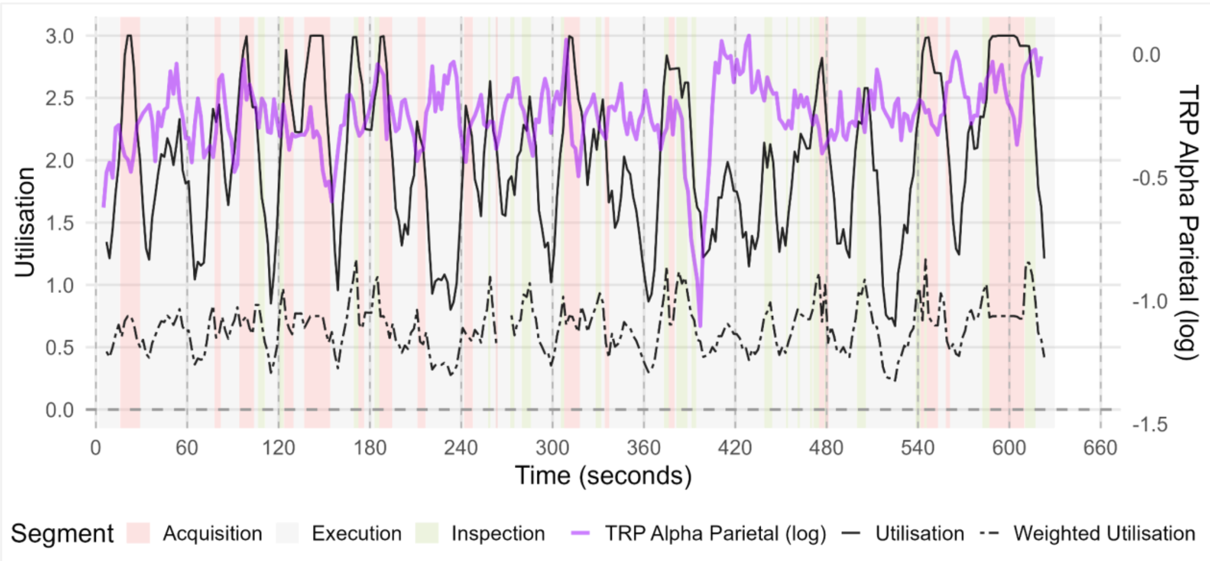

Figure 130 P#21: Dynamic changes in CL across moving windows (10s, 2s)

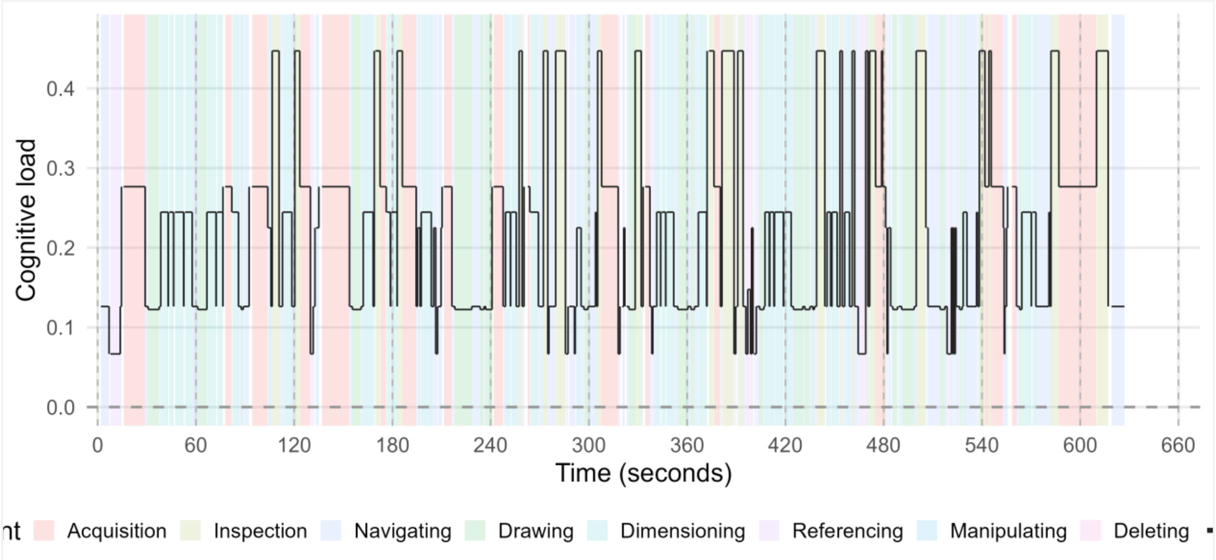

Figure 131 P#21: Dynamic changes in CL indicated by global values of weighted utilisations

Participant #22

Table 113 P#22: characteristics and CAD performance outputs

	Variable/Metric	Value
Experience & Expertise	Professional working experience	19
	Engineering activities	3
	Using CAD software	3
	CAD skill	3
	CAD proficiency	3
Spatial skill	Spatial visualisation	11
	Mental rotation	8
	Spatial orientation	3
CAD performance	Duration	15
	Dimensional accuracy	0.95
	Volume	0.99
	Surface area	0.99
	Shape	0.90
	Overall performance	0.26

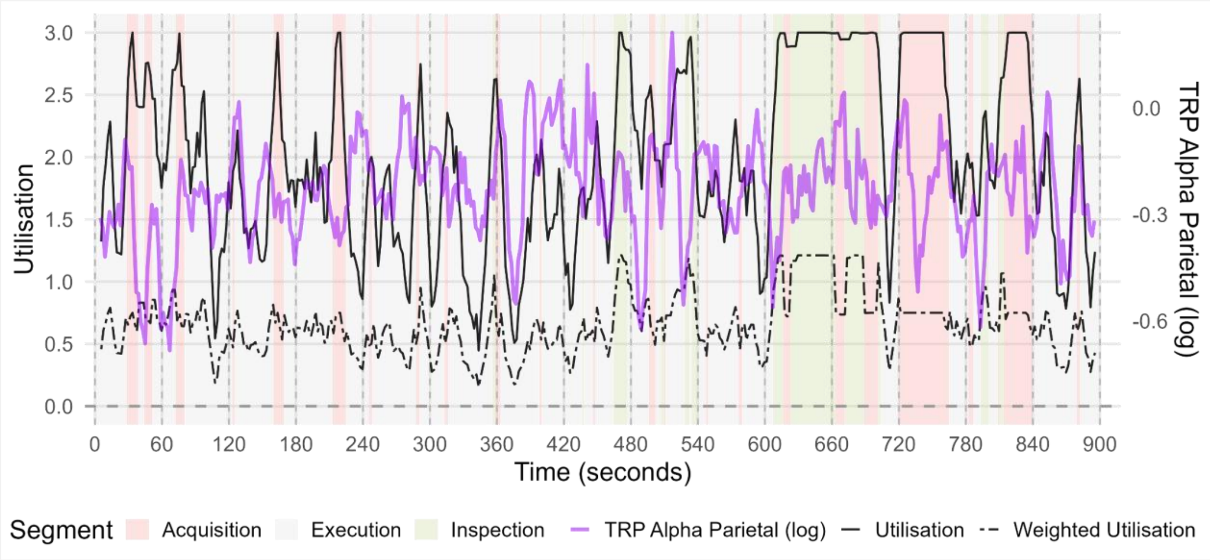

Figure 132 P#22: Dynamic changes in CL across moving windows (10s, 2s)

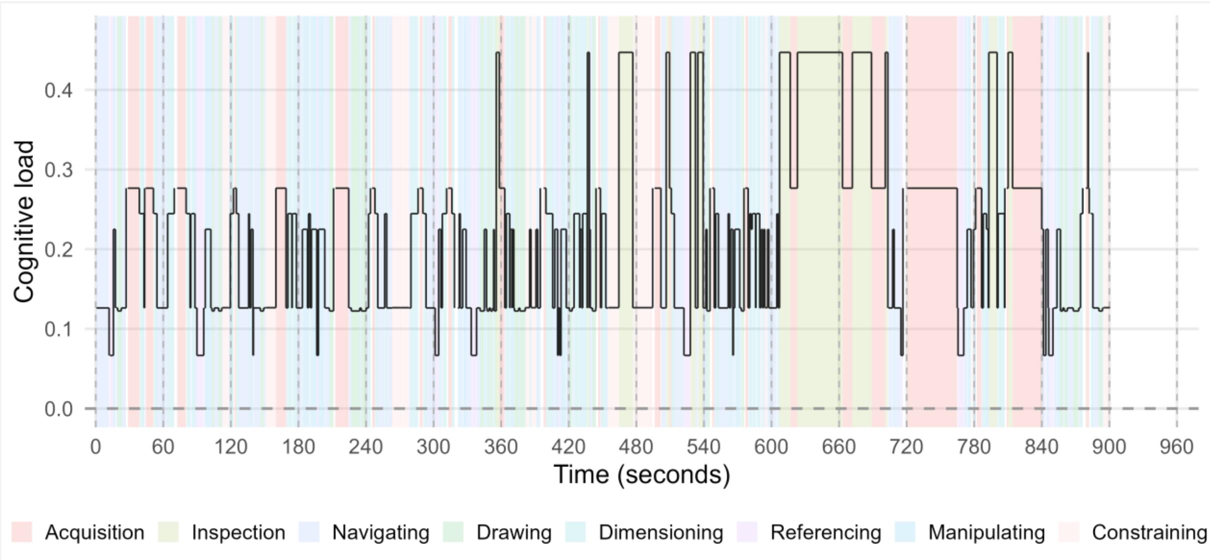

Figure 133 P#22: Dynamic changes in CL indicated by global values of weighted utilisations

Participant #23

Table 114 P#23: characteristics and CAD performance outputs

	Variable/Metric	Value
Experience & Expertise	Professional working experience	39
	Engineering activities	3
	Using CAD software	3
	CAD skill	2
	CAD proficiency	3
Spatial skill	Spatial visualisation	8
	Mental rotation	9
	Spatial orientation	9
CAD performance	Duration	15
	Dimensional accuracy	0.14
	Volume	0.19
	Surface area	0.37
	Shape	0.87
	Overall performance	0.11

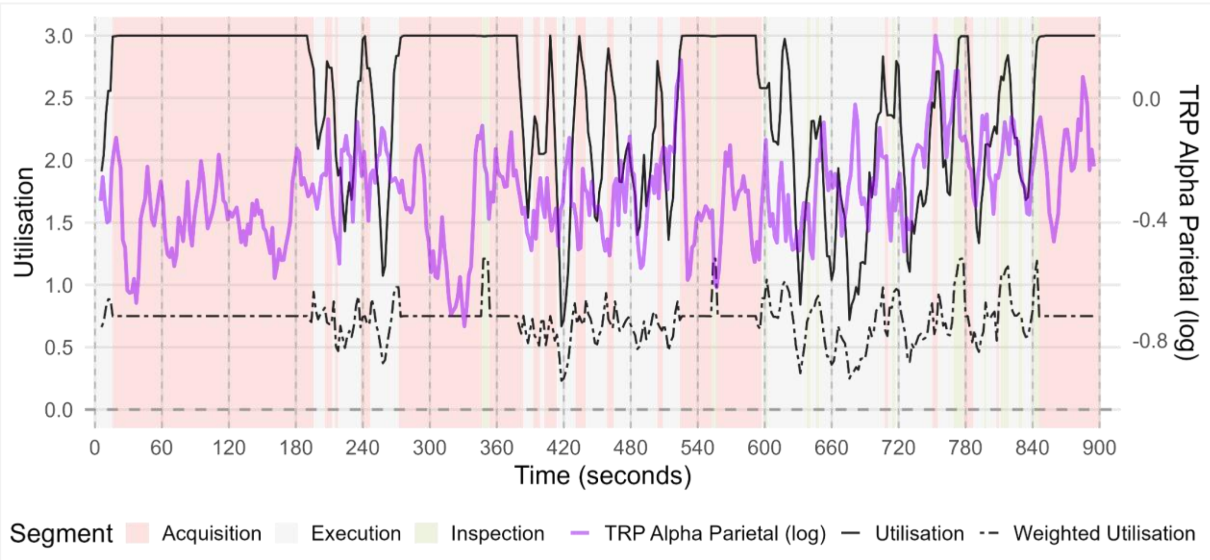

Figure 134 P#23 Dynamic changes in CL across moving windows (10s, 2s)

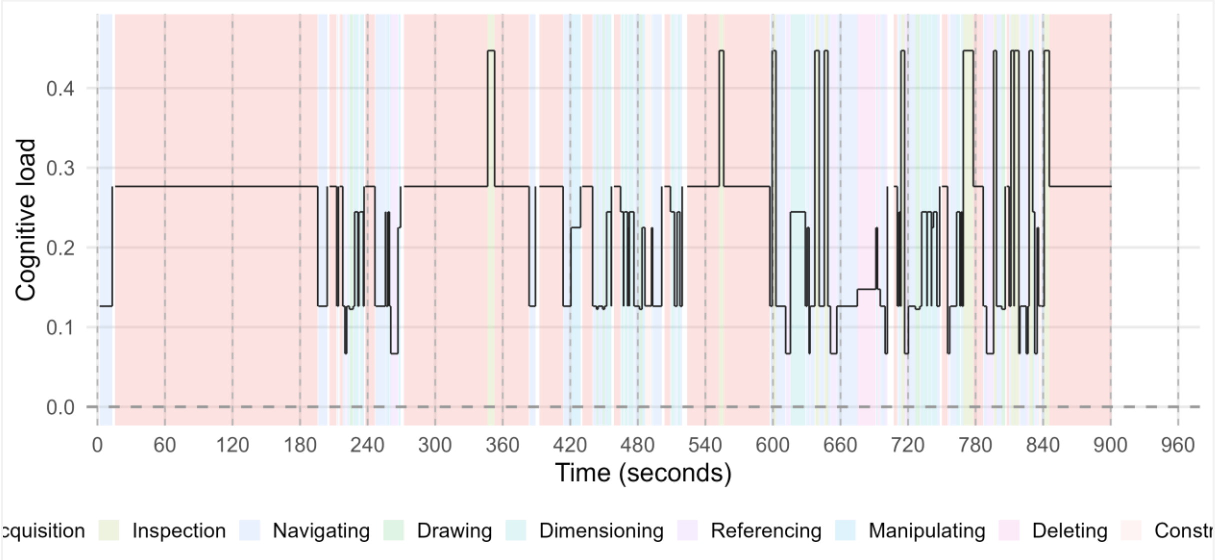

Figure 135 P#23: Dynamic changes in CL indicated by global values of weighted utilisations

Participant #24

Table 115 P#24: characteristics and CAD performance outputs

	Variable/Metric	Value
Experience & Expertise	Professional working experience	15
	Engineering activities	1
	Using CAD software	3
	CAD skill	2
	CAD proficiency	3
Spatial skill	Spatial visualisation	10
	Mental rotation	11
	Spatial orientation	10
CAD performance	Duration	15
	Dimensional accuracy	0.67
	Volume	1.07
	Surface area	0.97
	Shape	0.88
	Overall performance	0.23

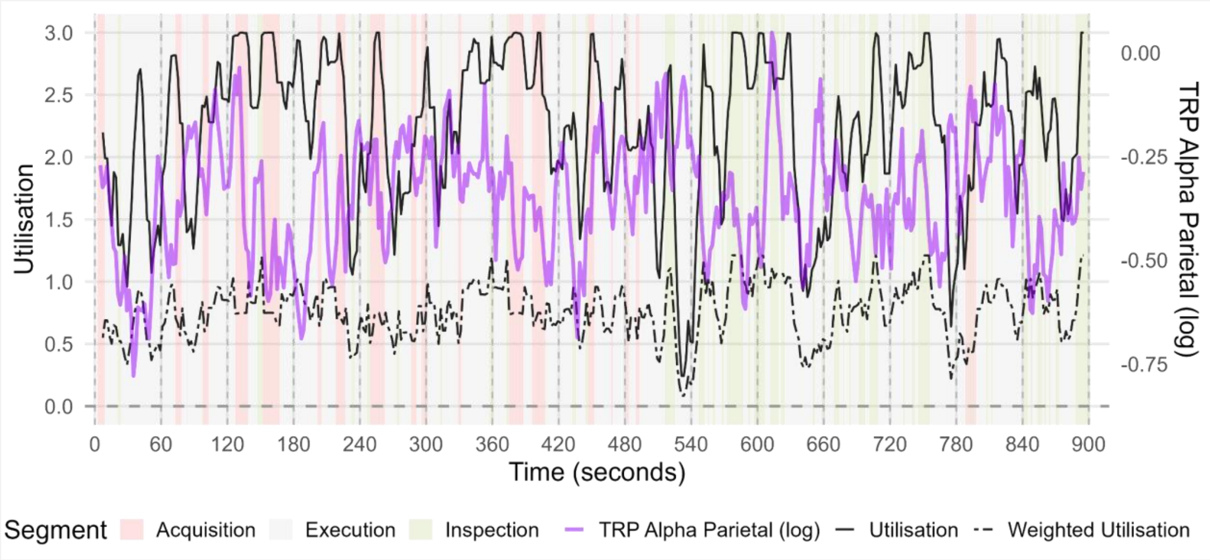

Figure 136 P#24: Dynamic changes in CL across moving windows (10s, 2s)

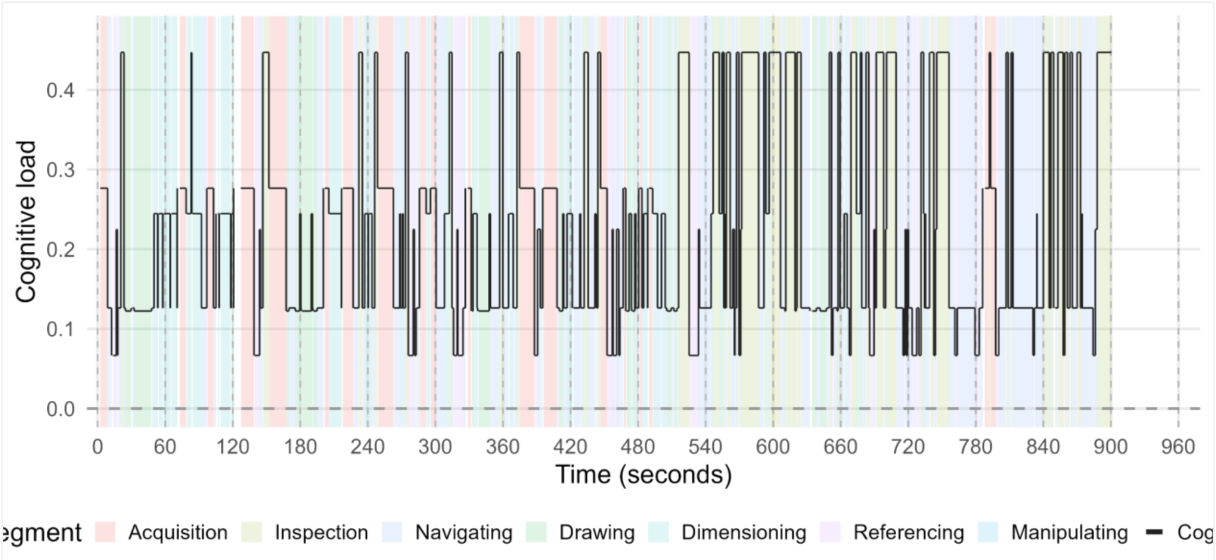

Figure 137 P#24: Dynamic changes in CL indicated by global values of weighted utilisations

Appendix C

Table 116 Metrics for quantifying alignment in dynamic changes of CL indicated by parietal alpha TRP and utilisations

Participant	Pair	Spearman	Kendall	CCF peak	Lag steps	Lag seconds	Agree rate	F1	DTW distance
P#1	TRP vs U	-0.0798	-0.0488	0.113	-4	-8	0.326	0.222	82.906
	TRP vs Uw	-0.102	-0.0644	0.141	-3	-6	0.317	0.156	87.696
	U vs Uw	0.936	0.849	NA	NA	NA	0.808	0.556	54.871
P#3	TRP vs U	0.215	0.140	0.186	0	0	0.377	0.451	137.534
	TRP vs Uw	0.159	0.106	0.148	5	10	0.337	0.286	154.198
	U vs Uw	0.922	0.816	NA	NA	NA	0.762	0.557	88.977
P#4	TRP vs U	-0.114	-0.0732	0.204	5	10	0.337	0.171	72.767
	TRP vs Uw	-0.131	-0.0829	0.119	5	10	0.337	0.146	70.302
	U vs Uw	0.953	0.881	NA	NA	NA	0.859	0.780	43.370
P#5	TRP vs U	-0.0633	-0.0383	0.123	5	10	0.315	0.233	90.904
	TRP vs Uw	-0.0781	-0.0477	0.0739	5	10	0.305	0.209	95.851
	U vs Uw	0.942	0.8656	NA	NA	NA	0.869	0.651	48.289
P#6	TRP vs U	-0.183	-0.122	0.0244	5	10	0.353	0.192	130.145
	TRP vs Uw	-0.190	-0.127	0.0983	4	8	0.337	0.136	132.506
	U vs Uw	0.910	0.821	NA	NA	NA	0.818	0.602	59.681
P#7	TRP vs U	-0.0640	-0.0427	0.0162	-5	-10	0.365	0.135	143.601
	TRP vs Uw	-0.00508	-0.00285	0.0437	3	6	0.338	0.203	139.196
	U vs Uw	0.940	0.861	NA	NA	NA	0.836	0.0156	90.603
P#8	TRP vs U	0.329	0.231	0.277	0	0	0.408	NA	147.586
	TRP vs Uw	0.255	0.176	0.261	-1	-2	0.417	0.226	155.921
	U vs Uw	0.836	0.746	NA	NA	NA	0.840	NA	137.951
P#9	TRP vs U	0.0167	0.008887	0.149	5	10	0.339	0.315	183.727
	TRP vs Uw	-0.0181	-0.0123	0.112	5	10	0.318	0.202	188.891
	U vs Uw	0.961	0.893	NA	NA	NA	0.878	0.685	54.007
P#10	TRP vs U	0.0155	0.0111	0.124	5	10	0.420	0.206	97.631
	TRP vs Uw	0.0274	0.0188	0.103	5	10	0.385	0.106	130.544
	U vs Uw	0.898	0.799	NA	NA	NA	0.772	0.248	92.180
P#11	TRP vs U	-0.00281	7.10E-05	0.0989	-5	-10	0.315	0.146	140.033
	TRP vs Uw	-0.0113	-0.00836	0.150	-5	-10	0.303	0.191	158.219
	U vs Uw	0.934	0.868	NA	NA	NA	0.897	0.404	102.592
P#13	TRP vs U	-0.106	-0.0700	0.0385	-5	-10	0.297	0.124	137.524
	TRP vs Uw	-0.0816	-0.0525	0.0339	-5	-10	0.301	0.151	146.302
	U vs Uw	0.944	0.873154	NA	NA	NA	0.939	0.205	54.469
P#14	TRP vs U	0.0745	0.0533	0.0855	-2	-4	0.3	0.194	118.369
	TRP vs Uw	0.0258	0.0208	0.0756	-2	-4	0.319	0.180	136.546
	U vs Uw	0.853	0.727	NA	NA	NA	0.777	0.393	94.416
P#15	TRP vs U	0.0242	0.0172	0.0541	-5	-10	0.340	0.221	136.407
	TRP vs Uw	0.0257	0.0169	-0.0202	-5	-10	0.349	0.107	145.429
	U vs Uw	0.919	0.819	NA	NA	NA	0.856	0.351	68.523
P#16	TRP vs U	0.200	0.134	0.292	3	6	0.385	0.384	143.784
	TRP vs Uw	0.156	0.104	0.292	4	8	0.322	0.171	171.302
	U vs Uw	0.853	0.749	NA	NA	NA	0.733	0.0876	158.597
P#17	TRP vs U	0.133	0.0917	0.135	-2	-4	0.354	0.210	144.476
	TRP vs Uw	0.158	0.110	0.135	-1	-2	0.354	0.163	141.535
	U vs Uw	0.921	0.833	NA	NA	NA	0.827	0.0593	85.048
P#18	TRP vs U	0.291	0.191	0.337	1	2	0.376	0.443	138.962
	TRP vs Uw	0.244	0.162	0.298	2	4	0.381	0.456	137.821
	U vs Uw	0.932	0.836	NA	NA	NA	0.811	0.671	87.402
P#19	TRP vs U	-0.0679	-0.0488	0.126	5	10	0.338	0.129	122.210
	TRP vs Uw	-0.0517	-0.0370	0.103	5	10	0.319	0.200	207.855
	U vs Uw	0.843	0.745	NA	NA	NA	0.782	0.141	194.760
P#20	TRP vs U	0.227	0.156	0.241	1	2	0.374	0.486	139.971
	TRP vs Uw	0.118	0.0761	0.0943	4	8	0.355	0.187	157.507
	U vs Uw	0.818	0.705	NA	NA	NA	0.823	0.257	141.434
P#21	TRP vs U	0.0797	0.0565	0.0958	5	10	0.339	0.230	144.260
	TRP vs Uw	0.00290	0.00185	0.116	4	8	0.365	0.148	160.237
	U vs Uw	0.836	0.726	NA	NA	NA	0.786	0.262	100.395
P#22	TRP vs U	-6.45E-04	-0.00108	0.139278	5	10	0.285	0.124	162.826
	TRP vs Uw	0.0497	0.0320	0.0876	-3	-6	0.310	0.236	206.772
	U vs Uw	0.910	0.815	NA	NA	NA	0.809	0.427	171.624
P#23	TRP vs U	0.255	0.175	0.296	4	8	0.344	NA	322.933
	TRP vs Uw	0.0238	0.0169	0.0677	5	10	0.366	0.08	801.663
	U vs Uw	0.558	0.498	NA	NA	NA	0.865	NA	698.244
P#24	TRP vs U	0.111	0.0757	0.200	2	4	0.345	0.225	170.518
	TRP vs Uw	0.095097	0.0677	0.187	2	4	0.327	0.169	157.298
	U vs Uw	0.804	0.702	NA	NA	NA	0.795	0.640	128.296

BIOGRAPHY

Fanika Lukačević was born in 1995 in Nova Gradiška (Croatia), where she finished primary and secondary school. In 2014, she enrolled in the study of mechanical engineering at the University of Zagreb Faculty of Mechanical Engineering and Naval Architecture (UNIZAG-FSB). She obtained a bachelor's degree in mechanical engineering in 2017 and graduated in 2020 with a master's degree in mechanical engineering with the specialisation in Product Design and Development. For excellence in education during the undergraduate part of the study, she was awarded the UNIZAG-FSB "Davorin Bazjanac" award. During her graduate studies, she worked at Hilti AG (Schaan, Liechtenstein) as part of an industrial internship; six months in the logistics department (Schaan, Liechtenstein) and seven months in the IT department (Buchs, Switzerland). She spent five months on a student exchange at Luleå Technical University (Luleå, Sweden). She was awarded the Rector's Award of the University of Zagreb for individual scientific work in the interdisciplinary field of research.

Since July 2020, Fanika has been employed as a Research and Teaching Assistant at the Department of Design and Product Development (UNIZAG-FSB). In parallel, she has been enrolled as a PhD student in a double-degree PhD programme in mechanical engineering, established between UNIZAG-FSB and Politecnico di Milano. As part of her doctoral studies, she spent comparable periods at both institutions and fulfilled study obligations at both universities. Fanika has assisted in teaching six courses related to computer-aided design and product development at UNIZAG-FSB, with an average annual teaching load of 165 hours. She has also worked on scientific and educational projects, including Team Adaptability for Innovation-Oriented Product Development (TAIDE) and Data-driven Methods and Tools for Design Innovation (DATA-MATION). Since 2018, she has participated in the organisation of the DESIGN conference. Her research interests include understanding human performance in computer-aided design, with a focus on digital technologies, human-computer interaction, and cognitive ergonomics. She co-authored seven journal and nine conference papers.

ŽIVOTOPIS

Fanika Lukačević rođena je 27. rujna 1995. u Novoj Gradiški, gdje je završila osnovnu i srednju školu (Gimnazija Nova Gradiška). Studij strojarstva na Fakultetu strojarstva i brodogradnje Sveučilišta u Zagrebu (FSB) upisala je 2014. Diplomu prvostupnice strojarstva stekla je 2017., a kao magistra inženjerka strojarstva diplomirala je 2020. na usmjerenju Konstruiranje i razvoj proizvoda. Dodijeljena joj je nagrada FSB-a "Davorin Bazjanac" za izvrsnost u studiranju tijekom preddiplomskoga dijela studija. Tijekom diplomskoga studija radila je u tvrtki Hilti AG (Schaan, Lihtenštajn) u okviru industrijske prakse; šest mjeseci u odjelu logistike (Schaan, Lihtenštajn) te sedam mjeseci u IT odjelu (Buchs, Švicarska). Provela je pet mjeseci na studentskoj razmjeni na Tehničkom sveučilištu Luleå (Luleå, Švedska). Nagrađena je Rektorovom nagradom Sveučilišta u Zagrebu za individualni znanstveni rad u interdisciplinarnom području istraživanja.

Od srpnja 2020. Fanika je zaposlena na Katedri za konstruiranje i razvoj proizvoda FSB-a. Paralelno je upisana na doktorski studij u okviru dvojnoga doktorata iz područja strojarstva, uspostavljenoga između Fakulteta strojarstva i brodogradnje Sveučilišta u Zagrebu i sveučilišta Politecnico di Milano. Kao dio doktorskoga studija, provela je razdoblja usporedivoga trajanja na objema institucijama te ispunila studijske obveze na oba sveučilišta. Fanika je sudjelovala u izvođenju nastave na šest kolegija povezanih s računalom potpomognutim konstruiranjem i razvojem proizvoda na FSB-u, s prosječnim godišnjim nastavnim opterećenjem od 165 sati. Također je radila na znanstvenim i obrazovnim projektima, uključujući *Team Adaptability for Innovation-Oriented Product Development* (TAIDE) i *Data-driven Methods and Tools for Design Innovation* (DATA-MATION). Od 2018. godine sudjeluje u organizaciji konferencije DESIGN. Njezini istraživački interesi uključuju razumijevanje ljudske izvedbe u računalom potpomognutom konstruiranju, s naglaskom na digitalne tehnologije, interakciju čovjeka i računala te kognitivnu ergonomiju. Koautorica je sedam radova objavljenih u znanstvenim časopisima i devet znanstvenih radova objavljenih u zbornicima skupova.

BIBLIOGRAPHY

Journal papers

Lukačević, Fanika; Becattini, Niccolò; Škec, Stanko
Identifying the electroencephalography features for measuring cognitive load in computer-aided design // *Journal of mechanical design*, 147 (2025), 12; MD-24-1740, 13. doi: 10.1115/1.4068746

Lukačević, Fanika ; Becattini, Niccolò ; Perišić, Marija Majda ; Škec, Stanko
Differences in engineers' brain activity when CAD modelling from isometric and orthographic projections // *Scientific reports*, 13 (2023), 1-16. doi: 10.1038/s41598-023-36823-9

Trpčić, Marija; Perišić, Marija Majda; Lukačević, Fanika; Škec, Stanko
Accuracy analysis of extraoral 3D scanning in dental prosthetic // *Acta stomatologica Croatica*, 57 (2023), 4; 339-352. doi: 10.15644/asc57/4/5

Horvat, Nikola ; Martinec, Tomislav ; Lukačević, Fanika ; Perišić, Marija Majda ; Škec, Stanko
The potential of immersive virtual reality for representations in design education // *Virtual reality*, 26 (2022), 1227-1244. doi: 10.1007/s10055-022-00630-w

Lukačević, Fanika ; Škec, Stanko ; Martinec, Tomislav ; Štorga, Mario
Challenges of utilising sensor data acquired by smart products in product development activities // *Acta Polytechnica Hungarica*, 19 (2022), 4; 165-187. doi: 10.12700/APH.19.4.2022.4.9

Lukačević, Fanika ; Škec, Stanko ; Törlind, Peter ; Štorga, Mario
Identifying subassemblies and understanding their functions during a design review in immersive and non-immersive virtual environments // *Frontiers of engineering management*, 8 (2021), 412-428. doi: 10.1007/s42524-020-0099-z

Lukačević, Fanika ; Škec, Stanko ; Perišić, Marija Majda ; Horvat, Nikola ; Štorga, Mario
Spatial perception of 3D CAD model dimensions and affordances in virtual environments // *IEEE access,* 8 (2020), Access-2020-40098, 18. doi: 10.1109/ACCESS.2020.3025634

Conference papers

Lukačević, Fanika; Becattini, Niccolò; Škec, Stanko
Investigating relationships between performance and workload in CAD tasks // Proceedings of the Design Society, Volume 5: ICED25 / Cascini, Gaetano (ur.).
Dallas (TX): Cambridge University Press, 2025. pp. 2161-2170. doi: 10.1017/pds.2025.10230

Lukačević, Fanika; Becattini, Niccolò; Škec, Stanko
Comparing engineering designers' brain activity in visuospatial reasoning tasks // Design Computing and Cognition'24: Volume 2: DCC24 / Gero, John S. (ur.).
Cham: Springer Cham, 2024. pp. 186-203. doi: 10.1007/978-3-031-71922-6_13

Lukačević, Fanika; Becattini, Niccolò; Škec, Stanko
Engineering designers' CAD performance when modelling from isometric and orthographic projections // Proceedings of the Design Society, Volume 4: DESIGN 2024 / Štorga, Mario; Škec, Stanko; Martinec, Tomislav et al. (ur.).
Cambridge University Press, 2024. pp. 653-662. doi: 10.1017/pds.2024.68

Martinec, Tomislav; Valjak, Filip; Horvat, Nikola; Goudswaard, Mark; Nygård Ege, Daniel; Ballantyne, Robert; Berg, Martin Francis; Glaser, Tobias; Grosse, Cornelius et al.
Virtual design hackathons: a data collection framework // Proceedings of the Design Society, Volume 4: DESIGN 2024 / Štorga, Mario; Škec, Stanko; Martinec, Tomislav et al. (ur.).
Cambridge University Press, 2024. pp. 45-54. doi: 10.1017/pds.2024.7

Lukačević, Fanika; Becattini, Niccolò; Škec, Stanko
EEG-based cognitive load indicators in CAD modelling tasks of varying complexity // Proceedings of the Design Society, Volume 3: ICED23.
Cambridge University Press, 2023. pp. 1545-1554. doi: 10.1017/pds.2023.155

Lukačević, Fanika ; Li, Shumin ; Becattini, Niccolò ; Škec, Stanko
Comparing EEG Brain Power of Mechanical Engineers in 3D CAD Modelling from 2D and 3D Representations // Proceedings of the Design Society, Volume 2: DESIGN22.
Cambridge University Press, 2022. pp. 901-910. doi: 10.1017/pds.2022.92

Martinec, Tomislav ; Škec, Stanko ; Lukačević, Fanika ; Štorga, Mario
Modelling Proportions and Sequences of Operations in Team Design Activities // Proceedings of the Design Society, Volume 1: ICED21.
Cambridge University Press, 2021. pp. 2187-2196. doi: 10.1017/pds.2021.480

Horvat, Nikola ; Škec, Stanko ; Martinec, Tomislav ; Lukačević, Fanika ; Perišić, Marija Majda Identifying the effect of reviewers' expertise on design review using virtual reality and desktop interface // Proceedings of the Design Society, Volume 1: DESIGN Conference. Cambridge University Press, 2020. pp. 187-196. doi: 10.1017/dsd.2020.304

Horvat, Nikola ; Škec, Stanko ; Martinec, Tomislav ; Lukačević, Fanika ; Perišić, Marija Majda Comparing Virtual Reality and Desktop Interface for Reviewing 3D CAD Models // Proceedings of the Design Society, Volume 1: International Conference on Engineering Design. Cambridge University Press, 2019. pp. 1923-1932. doi: 10.1017/dsi.2019.198