2019

Load-inducing factors in instructional design: Process-related advances in theory and assessment

Dissertation

submitted by

Maria Wirzberger

born on December 15, 1983
in Lohr am Main

for obtaining the academic degree
doctor rerum naturalium (Dr. rer. nat.)
in Psychology

at the
Faculty of Behavioural and Social Sciences
of the TU Chemnitz

Submitted: October 26, 2018
Defended: January 22, 2019

Advisors
Prof. Dr. Günter Daniel Rey
Prof. Dr. Josef F. Krems

Without the knowledge of human cognitive processes, instructional design is blind.

(Sweller, Ayres, & Kalyuga, 2011, p. v)

Acknowledgements

Going through the process of this thesis has been a journey with many challenges, crises, and lots of emotional ups and downs. I am deeply grateful for the warm and kind support from all the people that helped me during this time. I could not have done this without you, you are my true heroes!

First of all, I thank my advisor Professor Günter Daniel Rey, who always trusted my ability to manage things, encouraged me to approach the next steps in my career ahead of time, was proud of my success, always had an open ear for my ideas, and never lost the kindness to deal with my worries. I also thank my co-advisor Professor Josef F. Krems, who provided just the right input at the right points, shared insights from his broad experience in cognitive research, and finally set the schedule to get this thesis finished. My very special grateful thanks go to Dr. Jelmer P. Borst, who made me feel very welcome during my research stay in Groningen, was a great supervisor, kindly supported me in completing the missing pieces for my thesis, introduced me to the joys of MATLAB, and prevented me from getting lost in the parameter setting desert of ACT-R. I further thank my “three musketeers” in Chemnitz, my valued colleagues and friends René Schmidt, Alexander Skulmowski, and Robert Herms, with whom I had the pleasure to do amazing research, participate in inspiring discussions, and who also pulled me out of the one or the other emotional crisis. Further, thanks to Shirin Esmaeili Bijarsari, who was a great and inspiring student assistant and co-author, supported me in conceptual issues, data collection, and manuscript revisions, and offered me her friendship. Thanks as well to my co-authors and colleagues Steve Nebel, Maik Beege, and Sascha Schneider, who provided critical comments on my studies and manuscripts that helped me to improve my work, and to Nico Münch, who always offered me kind and patient technical support with a smile on his face. Many thanks to my scholar Ronja Joschko, whom I had the pleasure to mentor for one year, for devoting her time to proofread my synopsis and her valuable advice for improvement. Moreover, I am very grateful for the opportunity to have been part of the Research Training Group “CrossWorlds” and thank the spokesmen Professor Maximilian Eibl and Professor Peter Ohler for accepting my application, appreciating my work, and providing me valuable career-related advice.

On my journey, I further had the pleasure to connect with great people who provided encouragement to pursue my goals. First of all, I give my special thanks to Dr. Falk Lieder, who challenged my mind, believed in my great potential, chose me to work with him, and thus motivated me to keep on track and cross the finish line. I further thank Professor Ute Schmid, who is a great mentor and role model, strengthened my trust in my abilities, and fostered the growth of my professional network. Grateful thanks to Professor Verena Nitsch, who guided my steps in organizing my first conference symposium, supported my career development from an early stage with her experience and advice, and encouraged me to reach for the stars. Moreover, special thanks to my academic advisors Professor Manfred Thüring and Professor Dietrich Manzey, who already provided encouragement during my master's degree, gave me helpful advice from the start of my dissertation project, and valued my scientific potential. I also thank Professor Nele Russwinkel, who kindly introduced me to ACT-R, established the contact to the Cognitive Modeling Group in Groningen, and fostered my personal growth during my master thesis. Many thanks to Professor Georg Jahn, who made me aware of the "hidden networks" in academia, appreciated my opinion and my ideas, and inspired me with his energy and enthusiasm. In addition, many thanks to Jun.-Professor Anne Kerstin Reimers, who supported me with her experience in the challenges of an academic career, her open ear for my worries, and her friendship.

I also want to thank a few people outside the "academic circle", who were real treasures on my way. First of all, very grateful thanks to Susanne Meisel, who gave me comfort with her friendship when I started to despair, always provided me an open ear and time for a walk and encouraged me to take a break from work sometimes. I further thank Elke and Frank Hiemann, who were great and caring neighbors that took interest in my progress and my life. I also thank Heinrich Bohaboj, who provided the necessary spiritual guidance to keep me connected to myself and my surroundings. Moreover, I thank Els van Dongen and Rudi Kamminga, who were lovely and encouraging hosts during my stay in Groningen.

Finally, I express my greatest, very special thanks to the most important people in my life. I am deeply grateful for all the emotional and organizational support that I received from my parents Bernadette and Stefan Wirzberger and my brother Daniel Wirzberger during this demanding and challenging time. Without all your great love and care, it would have been impossible to realize all these steps. Thank you for being part of my life!

Chemnitz, October 26, 2018

Maria Wirzberger

Summary

This thesis addresses ongoing controversies in cognitive load research related to the scope and interplay of resource-demanding factors in instructional situations on a temporal perspective. In a novel approach, it applies experimental task frameworks from basic cognitive research and combines different methods for assessing cognitive load and underlying cognitive processes. The first experimental study (N = 96) involves a basal learning task related to processes of working memory updating. Distinct facets of cognitive load are manipulated simultaneously with reference to the number, distance, and repetition of presented letters. Reaction times and errors in updating and recall steps of the task indicate the individual and combined influence of the varied features and individual aptitude variables and further emphasize the processual nature of schema acquisition. Within the second experimental study (N = 116) participants complete an abstract symbol learning task with different levels of task complexity according to the number of included elements. At five predefined stages over the task, interruptions are induced by an embedded visual search task. From the continuous monitoring of performance efficiency, a logarithmically decreasing change of invested cognitive resources seems plausible. Divergent effects of the induced interruptions related to conditions of task complexity hint on the activation of distinct cognitive strategies. By extending these behavioral results with a cognitive modeling approach based on the cognitive architecture ACT-R, underlying cognitive processes and mechanisms could be inspected in more detail. From the obtained insights, the potential for deconstructing and formalizing effects of increased task complexity on a cognitive level emerges. Furthermore, the time-related reconsideration of the cognitive load framework receives support on a neural level. The third experimental study (N = 123) involves a dual-task setting that requires participants to learn visually presented symbol combinations while memorizing auditory presented number sequences. Cognitive load during the learning task is addressed by secondary task performance, prosodic speech parameters, and physiological markers. In addition, the robustness of the acquired schemata is tested by a transfer task that requires participants to apply the obtained symbol combinations. The observed pattern of evidence supports the idea of a logarithmically decreasing progression of cognitive load with increasing schema acquisition. It further hints on robust and stable transfer performance under enhanced transfer demands. Taken together, the evidence obtained in this thesis emphasizes a process-related reconceptualization of the existing theoretical cognitive load framework and underlines the importance of a multimethod-approach to continuous cognitive load assessment. On a practical side, it informs the development of

adaptive algorithms and the learner-aligned design of instructional support and thus leverages a pathway towards intelligent educational assistants.

Zusammenfassung

Die vorliegende Dissertation nähert sich aktuellen Kontroversen in der Forschung zur kognitiven Beanspruchung in Lehr-Lernsituationen im Zusammenhang mit der Abgrenzung und dem Zusammenspiel ressourcenbeanspruchender Faktoren unter einer zeitbezogenen Perspektive. In einem neuartigen Forschungsansatz werden zu diesem Zweck experimentelle Aufgaben aus der kognitiven Grundlagenforschung angewendet und verschiedene Methoden zur Erfassung der kognitiven Beanspruchung und der Betrachtung zugrunde liegender kognitiver Prozesse kombiniert. Die erste experimentelle Studie ($N = 96$) beinhaltet eine basale Lernaufgabe im Zusammenhang mit Prozessen des Working Memory Updating. Definierte Facetten der kognitiven Beanspruchung werden simultan anhand der Anzahl, dem Abstand und der Wiederholung präsentierter Buchstaben manipuliert. Reaktionszeiten und Fehler in Updateschritten und finaler Wiedergabe im Zuge der Aufgabe zeigen den individuellen und kombinierten Einfluss der variierten Merkmale und individueller Charakteristika der Lernenden und unterstreichen zusätzlich den prozessualen Charakter des Schemaerwerbs. In der zweiten experimentellen Studie ($N = 116$) absolvieren die Teilnehmenden eine abstrakte Symbollernaufgabe mit unterschiedlichen Komplexitätsstufen, die durch die Anzahl der enthaltenen Elemente determiniert werden. Zu fünf vordefinierten Zeitpunkten im Aufgabenverlauf erfolgen Unterbrechungen durch eine eingebettete visuelle Suchaufgabe. Auf Basis der kontinuierlichen Erfassung der Leistungseffizienz erscheint eine logarithmisch abnehmende Veränderung der investierten kognitiven Ressourcen plausibel. Unterschiedliche Effekte der induzierten Unterbrechungen in den Bedingungen der Aufgabenkomplexität deuten auf die Aktivierung unterschiedlicher kognitiver Strategien hin. Mit der Erweiterung der verhaltensbezogenen Befunde um einen kognitiven Modellierungsansatz, basierend auf der kognitiven Architektur ACT-R, können die zugrunde liegenden kognitiven Prozesse und Mechanismen genauer untersucht werden. Die gewonnenen Erkenntnisse bieten das Potenzial zur Dekonstruktion und Formalisierung von Effekten erhöhter Aufgabenkomplexität auf kognitiver Ebene. Gleichzeitig stützen diese eine zeitbezogene Neubetrachtung des Rahmenmodells kognitiver Beanspruchung auf neuronaler Ebene. Die dritte experimentelle Studie ($N = 123$) nutzt einen Dual-Task-Ansatz, bei dem die Teilnehmenden visuell präsentierte Symbolkombinationen lernen, während sie sich gleichzeitig auditiv präsentierte Zahlenreihen merken sollen. Die kognitive Beanspruchung während der Lernaufgabe wird durch die Sekundäraufgabenleistung, prosodische Sprachparameter und physiologische Marker erfasst. Darüber hinaus wird die Robustheit der erworbenen Schemata durch eine Transferaufgabe geprüft, welche die Anwendung der zuvor erlernten Symbolkombinationen erfordert. Das

resultierende Evidenzmuster stützt die Idee eines logarithmisch abnehmenden Verlaufs der kognitiven Beanspruchung mit zunehmendem Schemaerwerb und deutet auf eine robuste und stabile Transferleistung auch unter erhöhten Aufgabenanforderungen hin. Zusammenfassend betonen die in der vorliegenden Dissertation gewonnenen Erkenntnisse eine prozessgeleitete Rekonzeptualisierung des bestehenden theoretischen Rahmenmodells der kognitiven Beanspruchung und unterstreichen zusätzlich die Bedeutung eines multimethodischen Ansatzes zur kontinuierlichen Erfassung der kognitiven Beanspruchung. Auf praktischer Seite lassen sich zentrale Hinweise für die Entwicklung adaptiver Algorithmen sowie eine an den Lernenden orientierte Gestaltung instruktionaler Prozesse ableiten, welche den Weg zu intelligenten Lehr-Lernsystemen eröffnen.

Overview on included original content

The following table lists the original content that is used in the indicated chapters of the submitted thesis.

Chapter	Original content
Article 1	Wirzberger M., Beege M., Schneider S., Nebel S. & Rey G.D. (2016), One for all?! Simultaneous examination of load-inducing factors for advancing media-related instructional research, *Computers & Education, 100*, 18-31. doi: 10.1016/j.compedu.2016.04.010
Article 2	Wirzberger, M., Esmaeili Bijarsari, S., & Rey, G. D. (2017). Embedded interruptions and task complexity influence schema-related cognitive load progression in an abstract learning task, *Acta Psychologica, 179*, 30-41. doi: 10.1016/j.actpsy.2017.07.001
Article 3	Wirzberger, M., Herms, R., Esmaeili Bijarsari, S., Eibl, M., & Rey, G. D. (2018). Schema-related cognitive load influences performance, speech, and physiology in a dual-task setting: A continuous multi-measure approach. *Cognitive Research: Principles and Implications,* 3:46. doi: 10.1186/s41235-018-0138-z

Content

Synopsis

1 Introduction

1.1 Practical significance

Recent advances in computer-based technology offer the potential to explore innovative solutions in learning and training contexts. The arising scope relates to various benefits from both the educating and the educated perspective – given that emerging challenges are considered in a sufficient way. When designing intelligent educational systems, the most important goal persists in providing each learner the opportunity to achieve the best possible learning outcome with appropriate effort. A sophisticated approach to enhanced cognitive skill acquisition can be achieved by tailoring instructional support to individual learners' needs, which also increases motivation and encourages sustained performance. For instance, during learning activities, an adaptive system could align the amount and speed of the presented content or the degree and scope of instructional feedback. At the same time, learners' cognitive resources should not be overloaded due to the variety of occupied modalities and provided interactive features. In consequence, for providing adequate feedback, such systems need sufficient input related to both performance and cognitive resource supply. While performance can be inspected via tracking learners' task-related progress, the actual pattern of invested cognitive resources needs to be derived from an enhanced scope of learner-related information. These can result from including additional channels, such as behavior or psychophysiological signals, as well as supporting evidence by computational models on task-related cognitive processes. Arising challenges in the development of adaptive educational systems firstly involve issues of adequate assessment. They address the accurate learner state recognition that requires intelligent algorithms for correctly interpreting the acquired signals. Secondly, system behavior needs to be adjusted continuously to meet learners' needs as sophisticated as possible. Motivated by both issues, this thesis explores the pattern of cognitive resource investment related to task performance by monitoring variations in cognitive demands over the task with a novel combination of sensitive measures related to performance, speech, physiological reactions, and computational cognitive modeling. On this account, it contributes evidence relevant to developing dynamic recognition algorithms underneath intelligent educational technologies.

1.2 Theoretical background

Approaching the subject on a theoretical level, instruction-related cognitive demands need to be considered and monitored carefully. A well-established theory in this field is the Cognitive Load Theory (Sweller, Ayres, & Kalyuga, 2011; Sweller, Van Merrienboër, & Paas, 1998), which addresses the construct of cognitive load in terms of working memory resources required to perform a certain task in a given situational context (Kalyuga & Plass, 2018). These demands relate to the ergonomic concept of strain (Beckmann, 2010; Kalyuga, 2011; Manzey, 1998), as they constitute a subjective experience that each individual learner has to cope with. The theory looks back on a history of about 30 years of active research with broad impact on conducive instructional design in a variety of domains. Upon its core assumptions, it resides on vested models of memory (Anderson, 1983; Atkinson & Shiffrin, 1971; Baddeley, 1992) that indicate limited working memory resources in terms of both duration and capacity of stored information. Besides a temporal duration of 20 to 30 s (e.g., Wickens, Hollands, Banbury, & Parasuraman, 2013), according to more recent research, the number of simultaneously available elements would reside around four (Cowan, 2010). By contrast, long-term memory resources provide nearly infinite storage capacity and duration and thus can be used to establish permanent knowledge structures. According to Schweppe and Rummer (2014), both memory systems are strongly interconnected, as working memory resources represent the activated part of long-term memory that holds the attentional focus. In line with evidence from schema theory (Anderson, 1983), the emerging organized knowledge structures are described as schemata that involve both declarative and procedural components (Gagné & Dick, 1983). Existing schemata influence how learners manage certain learning content and can be modified with new knowledge. In a recent review summarizing the evidence on schemata, Gosh, and Gilboa (2014) describe associative network structures, the foundation in multiple episodes, a lack of unit detail, and adaptability as core characteristics of the schema concept. With reference to established models of learning and skill acquisition (Anderson, 1982; Fitts & Posner, 1967; Ebbinghaus, 1964) higher amounts of resource investment are plausible in earlier process stages when knowledge structures still have to be established. The initially declarative knowledge becomes increasingly procedural and automated with task progression and in consequence, demands less cognitive resources.

Since its first description in the 1980s, the Cognitive Load Theory underwent several stages of refinement in terms of the postulated facets of cognitive load in learning situations. In the beginning, it focused on the prevention of harmful effects from task-irrelevant aspects, referred to as extraneous cognitive load (Sweller, 1988). This kind of resource demands arises from an inappropriate instructional presentation, for instance, due to inherent demands to split the attention between relevant sources of information (Ginns, 2006; Schroeder & Cenkci, 2018). On a broader level, it further involves interfering situational aspects of the learning context such as the prevalence of competing goals (Gerjets, Scheiter, & Schorr, 2003) by additional or interrupting tasks. The next stage of development expanded the focus on a load-inducing facet relevant to the learning task: the complexity of the used learning material in relation to existing previous knowledge (Sweller & Chandler, 1994). Along with defining this facet as intrinsic cognitive load, the concept of element interactivity was introduced, which emphasizes the interrelation of information elements as source of complexity (Chen, Kalyuga, & Sweller, 2017; Ngu, Phan, Yeung, & Chung, 2018). Characteristically, changes in element interactivity are related to the nature of what is learned (Sweller, 2010) and its amount should be kept at a manageable level for the individual learner to foster optimal learning outcomes. On this account, intrinsic cognitive load offers a toehold for adaptive learning procedures tailored to learners' expertise. A step ahead, in addition to extraneous and intrinsic cognitive load, a further source of cognitive load with beneficial effects for learning was introduced, primarily on theory-based accounts to explain so far unexplainable patterns of evidence (Paas, Tuivonen, Tabbers, & Van Gerven, 2003; Sweller et al., 1998). This so-called germane cognitive load emerges from learning-related processes of schema acquisition and automation and operates under the assumption of highly motivated learners that devote all available resources to these processes. Extraneous, intrinsic, and germane cognitive load were assumed to contribute independently and additively to overall cognitive resource demands in learning contexts (Sweller et al., 1998). This triarchic model of cognitive load can further be connected to the Cognitive Theory of Multimedia Learning (Mayer, 2009), another influential explanatory framework in instructional design. Although there is no exact mapping, essential cognitive processing relates to intrinsic cognitive load, as it deals with the selection and representation of relevant learning material in working memory. Extraneous cognitive processing is caused by suboptimal instructions, reminding of extraneous cognitive load, while generative cognitive processing corresponds to germane cognitive load by an active organization and integration of learning contents as well as learners' level of motivation (Kalyuga, 2011; Mayer, 2009).

Although the outlined facets have been broadly discussed in the corresponding literature, issues regarding the proper empirical assessment and psychometric separation persist. Such raise doubts on the originally postulated assumption of their purely additive interplay, as well as the independence of the later introduced germane cognitive load facet. To address the arising concerns, suggestions emerged to reformulate germane cognitive load as germane resources invested to deal with task-relevant intrinsic cognitive load (Kalyuga, 2011; Sweller, 2010, 2018) contrary to extraneous resources to deal with task-irrelevant extraneous cognitive load. Kalyuga and Singh (2016) argue even more towards a strict re-reduction of the framework into a two-component model that merely differentiates facilitative (productive) and impairing (unproductive) cognitive load factors and completely subsumes germane cognitive load under the facet of intrinsic cognitive load. By contrast, Seufert (2018) emphasizes the reasonability to retain the separation of intrinsic and germane cognitive load when considering aspects of self-regulation. In this context, both intrinsic and extraneous cognitive load facets represent task affordances imposed by the learning material, whereas germane cognitive load refers to learner-based decisions. She further criticizes the mainly static and deterministic perspective on cognitive load and clearly outlines the benefits of a more dynamic view on changes in cognitive load during learning. Following de Jong (2010), the separate consideration of the facets of intrinsic and germane cognitive load receives further confirmation as both represent distinct ontological categories. Whereas intrinsic cognitive load is related to the static complexity of the presented material, germane cognitive load refers to dynamic cognitive processes. Galy, Cariou, and Melan (2012) also outline the asymmetric nature of these facets that are supposed to act on different components of the cognitive system. According to Schnotz and Kürschner (2007), decreasing levels of cognitive load with increasing expertise are indeed plausible, which further advocates to adopt a process perspective on cognitive load, as claimed by Beckmann (2010).

Taken together, the existing literature reveals the lack of a time-related perspective in instructional cognitive load research. Such drives the demand of a processual reconceptualization of the three-component model that quantifies temporal changes resulting from schema-acquisition across the task. Based on this position, as documented in Wirzberger, Beege, Schneider, Nebel, and Rey (2016), Wirzberger, Esmaeili Bijarsari, and Rey (2017), and Wirzberger, Herms, Esmaeili Bijarsari, Eibl, and Rey (2018), the current thesis follows distinct levels of inspection of the outlined cognitive load facets with focus on their interplay during the learning process. Inspired by the concept of the zone of proximal development (Vygotski, 1963), the goal of intelligent adaptive systems would then be to keep the resulting cognitive load pattern at a manageable level for each learner at all stages.

1.3 Cognitive load assessment

According to Sweller (2018), the Cognitive Load Theory has originally been developed as a theoretical construct to explain experimentally obtained results, with little attempt to actually measure cognitive load. Nevertheless, since its initial description a variety of cognitive load measures has emerged (Paas et al., 2003; Sweller et al., 2011; Zheng, 2018). They operate on various parameters that can be categorized into subjective ratings, performance measures, physiological markers, and behavioral indices. Following Chen et al. (2016), while performance measures directly reflect task-related outcomes, behavioral indices hold information that does not directly affect domain-based outcomes. Contrary to physiological markers, the occurrence of behavioral indices can mostly or entirely be controlled by the learner. Related to arising differences in learners' evaluation of the complexity of the presented content and the benefit of the provided instructional support with increasing schema acquisition (Martin, 2018; Schnotz & Kürschner, 2007), a continuous monitoring of cognitive resource demands is advisable, which constitutes a core focus of the current thesis. An overview of the cognitive load indicators used in the related experimental studies is provided in *Table 1*.

Table 1

Cognitive load measures applied by the experimental studies in the included articles

Study	**Subjective ratings**	**Performance**	**Physiology**	**Behavior**
1	Paas (1992)	Reaction times Error rates	-	-
2	Krell (2015)	Efficiency Interruptions	-	-
3	Leppink et al. (2013)	Efficiency Secondary task	HR, SCR	Prosody

Note. Efficiency = correct responses per second, HR = heart rate, SCR = skin conductance response, Prosody = number and duration of silent pauses, phoneme-based articulation rate.

Amongst the earliest attempts to provide insights into cognitive demands arising from learning situations, subjective rating scales comprise a broadly used instrument. In particular, the unidimensional mental effort scale by Paas (1992) offers a convenient and easily usable option, although its informative value as single rating is limited. It requires participants to rate their perceived mental effort on a nine-point Likert scale ranging from "very, very low" to "very, very high". Kalyuga, Chandler, and Sweller (1999) use a modified scale to assess subjective mental load with a rating of instructional difficulty on a seven-point Likert scale ranging from "extremely easy" to "extremely difficult". In line with this procedure, Wirzberger et al. (2016) applied the mental effort scale accompanied by a rating on estimated task difficulty to enhance the scope of the stated predications. A more differentiated questionnaire that aims at assessing experienced mental load and mental effort is provided by Krell (2015). While mental load refers to cognitive demands arising from task-related and situational characteristics, mental effort refers to cognitive capacity invested in dealing with them. The questionnaire involves six items for each mental load (e.g., “The tasks were difficult to answer.”) and mental effort (e.g., “I have made an effort at the processing of the tasks.”) that have to be rated on a seven-point Likert scale from “not at all (1)” to “moderately (4)” and “totally (7)”. It was used by Wirzberger, Esmaeili Bijarsari, et al. (2017) to obtain additional insights in invested cognitive resources across conditions. Contrary to the previously reported ratings, the questionnaire developed by Leppink, Paas, Van der Vleuten, Van Gog, and Van Merriënboer (2013) addresses the facets of intrinsic, extraneous, and germane cognitive load separately. It includes three questions on each intrinsic and extraneous cognitive load and four questions referring to germane cognitive load. The first two categories are closely connected to the underlying conceptual definitions by tapping either the complexity of the topic to be learned or the clarity of the instructional explanations (Ayres, 2018). In terms of the latter category, the related questions focus on understanding and knowledge acquisition, which turned out to be a rather controversial issue due to the lack of meaningful results in some studies (e.g., Leppink, Paas, Van Gog, Van der Vleuten, & Van Merriënboer, 2014). The reversed effect pattern for the germane cognitive load facet reported by Wirzberger et al. (2018), who also applied this questionnaire, further supports this critique. Leppink et al. (2014) discuss the related difficulties with reference to the already outlined re-reduced two-factorial cognitive load model. A more recent questionnaire by Klepsch, Schmitz, and Seufert (2017) addresses this issue by including the effort component more explicitly in questions related to germane cognitive load. The authors still emphasize the value of the three-factor model of cognitive load facets from a measurement point of view and state a more general applicability of their questionnaire across a wider range of educational subjects and domains.

Cognitive resource demands are further reflected in a more indirect way in performance-related parameters such as reaction times, error rates, and accuracy. These measures have a broad application and long history of use across a variety of psychological fields and disciplines. According to relevant literature from instructional cognitive load research (Hoffman & Schraw, 2010; Paas & Van Gogh, 2006) opposed to single accuracy or reaction time measures, a combined metric can be used as indicator for the quality of acquired cognitive schemata and thus offers a higher indicative value. Hoffmann and Schraw (2010) compare different approaches for calculating efficiency scores from both performance and effort indicators and outline the dependency of the chosen measure on the nature of the research question and the construct of interest. On this account, Wirzberger, Esmaeili Bijarsari, et al. (2017) and Wirzberger et al. (2018) applied an efficiency measure calculated from correct responses and reaction times according to the likelihood model, to obtain insights in changes across the process of schema acquisition. On methodological accounts, secondary tasks constitute a sophisticated way to shed light on task-related cognitive resource demands. They operate on the rationale that changes in working memory load related to a primary task can be monitored by a secondary task (Sweller, 2018; Kraiger, Ford, & Salas, 1993) and have already been proven reliability and validity in cognitive load assessment (e.g., Korbach, Brünken, & Park, 2017; Park & Brünken, 2018). This measurement approach is inspired by existing dual-task paradigms that apply a variety of tasks ranging from counting or reciting the alphabet to finger tapping, and humming a melody (e.g., D'Espositio, Onishi, Thompson, Robinson, Armstrong, & Grossman, 1996). Already applied secondary tasks in learning contexts include the observation of changes in auditory or visual stimuli (Brünken, Plass, & Leutner, 2004; Brünken, Steinbacher, Plass, & Leutner, 2002), requirements to memorize additional content (Renkl, Gruber, Weber, Lerche, & Schweizer, 2003; Wirzberger et al., 2018), the classification of auditory stimuli while performing a motor learning task (Esmaeili Bijarsari, Wirzberger, & Rey, 2017), or the performance of motor tasks such as tapping a previously learned rhythm by foot (Park & Brünken, 2015). However, choosing an appropriate secondary task that neither interferes with primary task requirements nor lacks sensitivity to observe arising demands still comprises a challenge when applying such task procedure continuously.

Increased cognitive demands also affect physiological states such heart rate, skin conductance, or brain blood flow dynamics. Due to the related characteristic of continuous assessment, they are particularly suited to obtain temporal progression patterns (Zheng & Greenberg, 2018). Amongst the variety of parameters and techniques, Wirzberger et al. (2018) recorded participants' mean normalized skin conductance response, indicating changes in the sympathetic nervous system (Chen, Zhou, & Yu, 2018), and heart rate, accompanying cognitive processing demands (Kennedy & Scholey, 2000). These measures have already proven sensitivity in related research (Chen et al., 2018) and point towards higher demands on cognitive resources by increasing values. However, they provide only an overall evaluation of the prevalent cognitive load level, without specifying different facets. A sophisticated conceptual approach to obtain information on individual cognitive load facets on a neural level was postulated by Whelan (2007). It aligns to existing evidence from functional neuroimaging literature that builds around the measurement of peaks in the blood oxygen level due to neuronal activity. Based on this rationale, he suggests that extraneous cognitive load would correspond in particular to activity in brain regions responsible for sensory processing, such as the posterior parietal association cortex, Broca's area, and Wernicke's area. By contrast, the intrinsic cognitive load component should be associated with activity in brain regions involved in maintaining and manipulating the attentional focus, in particular, the dorsolateral prefrontal cortex. Finally, germane cognitive load is assumed to hold connections to activity in brain regions related to motivation, as highly motivated learners are more likely to devote available cognitive resources solely to processes and strategies of schema acquisition. Corresponding brain regions, in this case, involve the superior frontal sulcus and the intraparietal sulcus. Although this approach offers high explanatory potential, so far it has not been explicitly tested yet due to the high methodological effort and inherent task-related constraints.

In terms of behavioral responses, duration-based parameters from speech signals have proven sensitivity to changing levels of cognitive load (Chen et al., 2016). They can be classified as behavioral, since they show inherent characteristics such as disfluency, articulation rate, content quality, the number of syllables, and the number and duration of pauses regardless of the meaning of the utterance. Existing evidence indicates that increasing levels of cognitive load result in a slower speech tempo as well as more and longer pauses within the speech flow due to necessary planning processes (e.g., Müller, Großmann-Hutter, Jameson, Rummer, & Wittig, 2001). Contrary to existing work that applied speech parameters to capture fixed task demands across shorter time spans (e.g., Yap, 2012), Wirzberger et al. (2018) inspected the phoneme-based features articulation rate, number of silent pauses, and duration of silent pauses with references to task-inherent processual changes during schema acquisition. Related work extended the focus by additional parameters and further applied the acoustic-prosodic features loudness and pitch, and the voice quality features vocal fold frequency and voice amplitude (Herms, 2018; Herms, Wirzberger, Eibl, & Rey, 2018). A comparison of discrete classes of low, medium, and high levels of cognitive load showed statistically significant differences for articulation rate, pause duration, pitch, and voice quality features. The latter indicate less rough or hoarse characteristics of the speech signal with increasing levels of cognitive load.

Considering the outlined characteristics, Korbach et al. (2017) already demonstrated the benefit of combining measures related to behavior and secondary task performance to achieve a continuous cognitive load assessment. Further accounting for the fact that a single measure alone is not sufficient to obtain the underlying pattern of cognitive resource investment in learning situations, Chen et al. (2018) emphasize an even more comprehensive approach on cognitive load assessment. They introduce a multimodal framework that fuses a variety of cognitive load-indicating channels of continuous learner-related information, for instance, physiological signals from skin conductance and electroencephalography and behavioral data streams from speech, gaze patterns, and mouse and keyboard interactions.

1.4 Research focus

The current thesis addresses ongoing controversies in instructional cognitive load research by examining the processes and mechanisms of the interplay between the outlined cognitive load facets. Corresponding to a process-related reconceptualization of the existing three-component model that takes into account temporal changes in resources related to schema acquisition across the task, it applies a combination of continuous approaches for cognitive load assessment. The resulting evidence should provide a foundation for the development of adaptive instructional procedures with learner-aligned instructional support. Considering this background, fading instructional guidance then can tie in with both the level of expertise, reflected in current performance, but also the level of invested cognitive resources, detectable by capturing learners' cognitive load. The inclusion of context-related features has further indicative value, for instance, due to a facilitative use of interruptions to keep learners involved in the task.

2 Experimental studies

On methodological accounts, the novelty of this thesis consists in applying experimental task frameworks from basic cognitive research to inspect the interplay of cognitive load facets in a controlled, internally valid manner. Due to the demonstrated impact of prior knowledge on task performance and cognitive resource demands (e.g., Chen et al., 2017; Rey & Buchwald, 2011; Rey & Fischer, 2013; Seufert, 2018), tasks with no or commonly shared prior knowledge were chosen to keep the arising influences at a constant level. A related joint characteristic of the reported studies constitutes the a priori determination of task complexity according to the number of interacting elements of information (Beckmann, 2010; Chen et al., 2017; Ngu et al., 2018; Sweller & Chandler, 1994). As a particular focus was put on the inspection of the learning process, a set of continuous measures was applied. Whereas data in the first and second study were collected in group-based settings, the third study used individual testing sessions due to the nature of the recorded measures. *Table 2* provides an overview of sample characteristics across the included studies and *Table 3* outlines details of the underlying research designs.

Table 2

Sample characteristics of the experimental studies reported in the included articles

Study	N	M [a]	SD [a]	Range [a]	Gender [b]
1	96	24.35	4.81	18-48	79.17
2	116	23.25	4.34	18-44	80.17
3	123	22.67	3.55	18-34	76.42

Note. [a] Age in years. [b] Percentage of female participants.

2.1 Methods

A main characteristic of the first experimental study (Wirzberger et al., 2016) comprised the investigation of the previously discussed facets of cognitive load in a joint task framework. The adopted paradigm of working memory updating (Ecker, Oberauer, Lewandowsky, & Chee, 2010) can be regarded as condensation of learning-relevant working memory processes, as content-related changes need to be represented correctly over time. In such tasks, an initially presented input undergoes several steps of updating, which involve processes of retrieval, transformation, and substitution that are reflected as well in a more implicit manner in the concluding recall of the final state. Aligned to Ecker et al. (2010), the presently employed task was formed of letter sets and accompanying alphabetic transformations over three (practice phase) or six (test phase) steps. Participants received a new set of letters at the outset of a trial that had to be incremented at one of the positions within each updating step and memorized afterward. As displayed in *Table 3*, facets of intrinsic, extraneous, and germane cognitive load were addressed by controlled, task-inherent variations (Beckmann, 2010) in a 3 x 2 x 2 within-subjects design: Firstly, the number of letters to be memorized simultaneously determined task complexity and resulted from adding or removing one letter around the intermediate difficulty of three letters (Ecker et al., 2010). Secondly, an increased distance between presented letters aligned to the split attention effect (Ginns, 2006; Schroeder & Cenkci, 2018) as means of inappropriate instructional presentation. Thirdly, the repetition of letter sets from a previous training sequence enabled the use of already existing task-related schemata. As individual aptitude variables are known to play an important role in such tasks, the standardized psychological inventory d2-R (Brickenkamp, Schmidt-Atzert, & Liepmann, 2010) provided insights into participants' concentration abilities. Due to the recording of error rates (corrected for inherited errors) and reaction times in both update steps and final recall as dependent variables, a verifying feedback (Shute, 2008) on the percentage of correct responses could be provided after completing the task.

Table 3

Study designs of the experimental studies reported in the included articles

Study	Trials	Session duration	Material	ICL	ECL	GCL
1	6 [a] / 24 [b]	45 min	d2-R Letter sequences	Number of letters (2/3/4)	Increased distance of letters [c]	Repetition of letter sets [c]
2	64 [d]	60 min	OSPAN [e] SSPAN [e] Symbol sequences	Number of symbols (2/3)	Interrupting visual search task	Performance efficiency (cr/rt*1000)
3	64 [d] / 60 [f]	60 min	OSPAN Symbol sequences [g]	Number of symbols (3/4)	Embedded secondary task	Performance efficiency (cr/rt*1000)

Notes. ICL = intrinsic cognitive load, ECL = extraneous cognitive load, GCL = germane cognitive load. [a] Practice phase, [b] Test phase, [c] With vs. without, [d] Learning task, [e] Short versions, [f] Transfer task, [g] An additional classification task applied symbol sequences with distortions to assess transfer demands.

In the subsequent second experimental study (Wirzberger, Esmaeili Bijarsari, et al., 2017), the focus was shifted towards the processual nature of schema acquisition. The task further addressed the issue of potentially occurring interferences of prior knowledge from the previous letter stimuli and used more abstract material to inspect the temporal interplay of cognitive load facets. In more detail, participants had to learn abstract geometric symbol combinations via trial and error by verifying feedback (Shute, 2008) that informed about the correctness of the response and the correct response in terms of errors. In line with the first study, the number of symbols in a defined order that formed a combination represented task complexity as between-subjects independent variable. Under the assumption that the prevalence of distracting tasks with competing goals (Gerjets et al., 2003) represents a common situational constraint in computer-based learning environments, interruptions were induced at five defined stages over the task as further within-subjects independent variable. The emerging effects on performance should further hint on the underlying progress in schema acquisition. Following Wickens

(2002), a resource-demanding perceptual task would be able to cause substantial interference with a cognitive task that involves storage and/or transformation processes in working memory. On this account, the interrupting task itself adopted a visual search paradigm with a sufficient number of geometric symbols similar to the learning task (Trick, 2008), as similarity (Gillie & Broadbent, 1989) and an appropriate task duration (Monk, Trafton, & Boehm-Davis, 2008) should ensure the interrupting potential. To assess the investment of cognitive resources related to schema acquisition in the resulting 2 x 5-factorial mixed design, performance efficiency computed from reaction times and correct responses (Hoffman & Schraw, 2010), was inspected as dependent variable. Participants working memory span (Unsworth, Heitz, Schrock, & Engle, 2005) and perceived mental load and mental effort (Krell, 2015) were obtained before the learning task, whereas the amount of recalled symbol combinations was recorded afterward.

The third experimental study (Wirzberger et al., 2018) extended the abstract visual-motor symbol sequence learning task used in the second study by an embedded auditory-verbal secondary task to enable a closer monitoring of the investment of cognitive resources related to schema acquisition. Distinct input and output modalities were chosen to ensure the occurrence of resource interference merely at a cognitive stage due to the simultaneous processing of task requirements (Wickens, 2002, 2008). The constant interchange of both tasks over time was inspired by the procedure of automated complex working memory span tasks (Redick et al., 2012; Unsworth et al., 2005), which are characterized by the alternating sequence of distractor and target tasks. While the primary task slightly adjusted the task paradigm by Wirzberger, Esmaeili Bijarsari, et al. (2017) in terms of both number and presentation of symbols, the secondary task required participants to memorize and recall a spoken five-digit sequence from start to finish of each trial. Again, task complexity in the primary task varied according to the interrelated number of symbols. In addition to performance parameters from primary and secondary tasks, inspected via combined efficiency measures (Hoffman & Schraw, 2010), cognitive load-related parameters from prosodic speech features and physiological parameters were recorded. Particularly the inspection of varying levels of cognitive load in speech-related characteristics, such as the number and duration of pauses and the articulation rate, comprises a novel and innovative solution in cognitive load research (Herms et al., 2018). Accompanied by more established physiological markers of skin conductance response and heart rate, the study provided an elaborated pattern of multimodal indicators for cognitive load. Beyond a subsequent recall of memorized symbol combinations, a specifically designed transfer task aimed at obtaining the robustness of the acquired schemata. Based on the set of previously learned symbol combinations, it required participants to categorize displayed symbol combinations in terms of their correctness. The task operated on a 2 x 5 factorial mixed design that aligned with the aforementioned between-subjects variation of task complexity due to the number of symbols. In addition, defined levels of distortion of the presented symbols induced increased transfer demands that were inspected in terms of errors (corrected for inherited errors) and reaction times. Again, participants' individual working memory capacity was controlled for by completing a working memory span task (Unsworth et al., 2005) before the learning task.

2.2 Results

Responding to the request of Martin (2018) to apply more complex statistical models to represent the factor time in cognitive load assessment, data analysis across the included experimental studies is characterized by advanced statistical approaches (see *Table 4*). The resulting continuous inspection of learner states further corresponds to Leppink and Van Merrienboër (2015), who advise against the aggregation of repeated measures due to the resulting loss in informative value about individual task-related progressions.

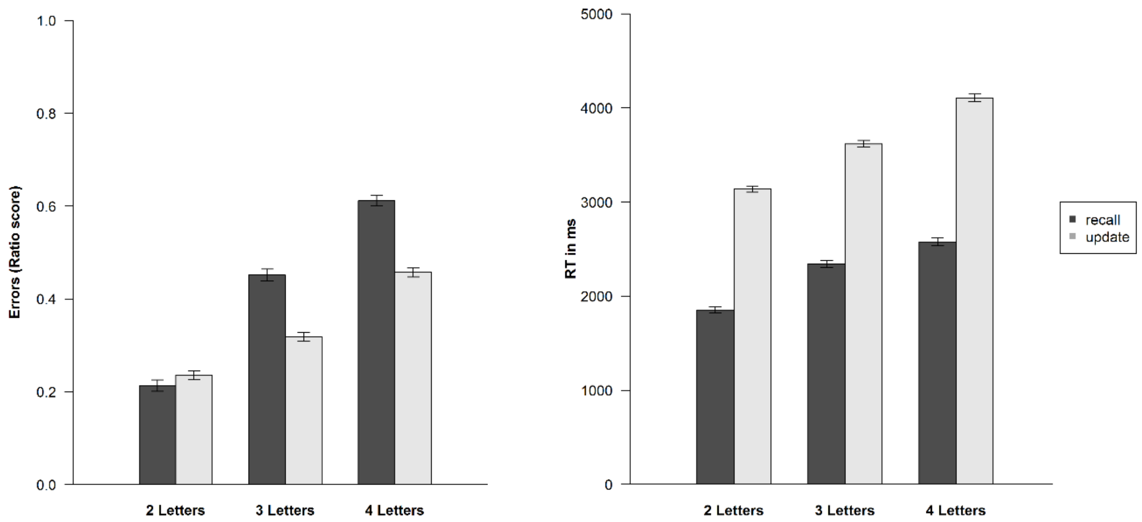

Figure 1. Complexity-related differences in reaction times and errors in both update and final recall stages. Error bars indicate standard errors.

In the first experimental study, results displayed a constant increase in both error rates and reaction times with increasing complexity, as shown in *Figure 1*. The visual impression suggests that participants reacted slower during the update steps but made more errors during the final recall. A significant increase in reaction times, but not error rates, with increased distance between stimuli resulted and the pattern of effects further indicated a benefit of repeated letter sets in terms of error rates and reaction times. In addition, significant two- and three-way-interactions between the examined facets showed up. Besides the beneficial influence of higher individual levels of concentration, task-related progress also fostered an overall increase in performance.

Table 4

Characteristics of data analysis in the experimental studies reported in the included articles

Study	**Participants included**	**Statistical approach**	**Main outcomes**
1	92	Linear mixed effects model [a]	Significant main effects and interactions Significant influence of aptitude and control variables
2	113	Conditional growth curve model [a] ANCOVA [a, b] t-Test	Superiority of logarithmic progression model Differences in interruption effects across conditions
3	103 [c]	Conditional growth curve model [a] Time series regression [d] ANOVA t-Test	Increasing performance under decreasing levels of cognitive load Robust transfer performance even under increased task demands

Notes. [a] Accounting for interindividual variance by inclusion of random intercept, [b] Based on linear mixed effects model, [c] Further exclusions required for secondary task performance (n = 102), speech parameters (n = 102), and physiological parameters (n = 101), [d] Accounting for interindividual variance by normalization on individual baseline.

Inspired by temporal models of learning and skill acquisition (Anderson, 1982; Ebbinghaus, 1964; Fitts & Posner, 1967), plausible linear, quadratic and logarithmic progressions were compared statistically in the second experimental study. The obtained results revealed a nonlinear increasing development of performance efficiency over time that differed between both conditions of task complexity, with superiority for the logarithmic model. In addition, condition differences arose with respect to the impairing influence of the induced interruptions, as a loss in performance was more obvious in the easy task condition. No differences in performance efficiency between easy and difficult task conditions resulted in the interrupting task. In addition, the amount of correctly recalled symbol combinations after completing the learning task was comparable between both conditions.

The pattern of evidence in the third experimental study indicated an increase in performance efficiency over time for both primary and secondary tasks. Performance benefits for the easy task condition were also supported by more correctly recalled symbol combinations after the learning task. In line with evidence from the second study, analyses of continuously recorded parameters were based on logarithmic progression models. Speech-related parameters pointed towards reduced levels of cognitive load with increasing task progress, as the number and mean duration of pauses decreased, and articulation rates increased. In a similar way, physiological parameters displayed decreasing progressions and furthermore showed a repetitive seasonal pattern across subtasks with increases in secondary task-related steps. Results obtained from the transfer task hint on robust and stable performance even under enhanced task demands due to distorted symbols. Reaction times were significantly increased, but generally faster under difficult task conditions.

2.3 Implications

Taken together, the pattern of evidence arising from the outlined experimental studies supports a process-related refinement of the theoretical framework of cognitive load, considering ICL and ECL on a structural and GCL on a processual level. Such also relates to the recently introduced reformulation of GCL as resources dealing with relevant aspects of a learning task (ICL) contrary to independent additivity (Sweller, 2018). On practical accounts, the obtained results provide the basis for developing multimodal models of cognitive load progression as algorithmic base for adaptive instructional support according to learners' individual cognitive resource supply.

Examining the results of the first experimental study in more detail, the prevalence of significant interactions supports existing doubts on the assumption of a purely additive relationship between the described cognitive load facets (Kalyuga, 2011; Sweller, 2018). In extension, the overall improvement in performance across the task emphasizes both the declarative and procedural nature of task-related schemata (Gagnè & Dick, 1983). The significant influence of increased concentration resources aligns with existing evidence on individual aptitude variables (e.g., Wirzberger & Rey, 2018) and emphasizes the importance of considering individual cognitive abilities in the context of learning. Increased demands due to enlarged spatial distance could be compensated by extended reaction times, whereas the overarching effect of increasing task complexity affected both measures without compensation. In addition, the study confirms the prevalence of the previously outlined processes involved in working memory updating (Ecker et al, 2010). These processes are directly reflected in increased reaction times during updating steps compared to the more indirect reflection in error scores during the final recall.

Summarizing the evidence obtained from the second experimental study, the superior logarithmic curve progression receives support from the well-established learning curve from Ebbinghaus (1964). On this account, the assumption that invested cognitive resources decrease at the same pace as learning performance increases receives reasonable corroboration but needs to be explored in more detail. Furthermore, condition effects in resumption performance relate to prior studies on volitional protection against competing goals (Gerjets et al., 2003; Scheiter, Gerjets, & Heise, 2014). Evidence shows that higher levels of task difficulty can shield against distractions from task-irrelevant information. In addition, such desirable difficulties could force people to apply alternative task-related strategies over time, for instance a more time-efficient heuristic encoding procedure that focuses just on a minor set amongst all offered cues.

The demonstrated increasing performance in both primary and secondary task in the third experimental study, accompanied by decreasing progressions in speech-related and physiological parameters, supports the assumption of decreasing levels of cognitive load due to increasing schema acquisition. As already outlined by Kraiger et al. (1993), such pattern hints on dynamics related to primary task automation, as free resources from this task can be increasingly devoted to deal with secondary task requirements. The evident seasonal pattern raises the conclusion that the auditory-verbal modality combination puts higher demands on learners' cognitive resources compared to the visual-motor modality combination. Since one potential explanation refers to the persisting dominance of the visual modality in many task domains, in instructional scenarios predominantly visual cues might be chosen for additional support. Similar to the results of the second experimental study, task performance in the transfer task again suggests a higher investment of cognitive resources with increasing task complexity. Tying in with evidence on desirable difficulties (Bjork & Bjork, 2011) and the zone of proximal development (Vygostki, 1963), to foster optimal learning performance, adaptive task procedures should provide constant challenges to keep learners involved in investing cognitive resources to achieve a robust and stable performance. On a methodological level, the correspondence between the applied measures particularly underlines the benefit to explore the potential of speech-related cognitive load indicators in multimodal learning environments.

2.4 Limitations

Although the first experimental study indicates statistically significant interactions between the addressed facets, these might have resulted due to interference in the experimental manipulation. In particular, the induced spatial distance highly depended on the number of presented letters, as a closer spatial proximity was required if there were more letters on the screen. Moreover, familiarity with the Latin alphabet could not be fully controlled and usually differs even amongst native speakers, which resulted in task-inherent benefits for participants with higher fluency and exposure.

From the obtained measures in the second experimental study, conclusions on the underlying progression of cognitive load result solely by the inversion of the resulting performance curves. Following Martin (2018), combined scores from invested time and obtained performance lack controllability, as participants could have reached equivalent levels of efficiency with different amounts of invested resources or achieved performance. Thus, a continuous monitoring of related resource demands is lacking as well as the further inclusion of motivational aspects. The latter could also have influenced how participants dealt with the task across different conditions of complexity. Since the task required participants to memorize only a few symbol combinations, participants facing increased complexity might have benefitted more from extended time frames due to presentation characteristics. Comparable to the previous study, the group-based testing sessions always involve the prevalence of peer-pressure in task-related timing.

The latter aspect was addressed in the third experimental study due to the use of individual testing sessions, as well as the alignment in terms of stimuli presentation between conditions. However, differences in symbol complexity might have resulted from varying visual characteristics of the used symbols, like the salient edges of a star. These could have fostered benefits in terms of the retentivity of certain symbol combinations. Moreover, task order ambiguities could have occurred, since the secondary task was presented first and interleaved by the primary task. Inspecting the obtained measures more closely, progressions in physiological parameters may hint on the prevalence of an orientation response at the outset of the task, followed by the adjustment with increasing task progress. However, even after the first ten trials, a recognizable progression persists that hints on a modified pattern of cognitive resource investment.

2.5 Future work

The first experimental study mainly indicated a reduction of task complexity, the use of more abstract material without the reliance on previous knowledge, and the focus on the more processual characteristic of schema acquisition. These aspects were addressed in the second experimental study that demonstrated the necessity to continuously monitor the task-related investment of cognitive resources. In addition, the question persisted how robust the acquired schemata would be under conditions of enhanced transfer demands. While the third experimental study could seize the outlined suggestions by applying a multi-method approach to cognitive load assessment, and a subsequent classification task for assessing transfer performance, it still requires the inclusion of motivational characteristics in future studies. Further valuable perspectives could arise from the use of additional parameters such as gaze movements or mouse interaction patterns as well as the transfer into more applied task domains. Moreover, exploring the use of different secondary task paradigms that involve incompatible modality-content matchings and combine auditory-motor and/or visual-verbal channels could extend the informative value in terms of the robustness of the observed patterns. On the level of data analysis, the application of more complex procedures for inspecting stages across the underlying cognitive load progression, such as hidden Markov models (e.g., Visser, Raijmakers, & Molenaar, 2002), could further increase the predictive scope of the obtained insights. For purposes of classifying and interpreting multimodal cognitive load-related signals, the additional use of machine learning approaches can be of value, as demonstrated by Herms (2018).

3 Cognitive modeling

To strengthen and extend evidence obtained from the second experimental study (Wirzberger, Esmaeili Bijarsari, et al., 2017), a cognitive modeling approach using the cognitive architecture ACT-R (Adaptive Control of Thought – Rational; Anderson, 2007) constitutes a further methodological building block of the current thesis. The related purposes are twofold: Firstly, potential explanations for the unexpected effect of induced interruptions should be explored. Secondly, further insights into the cognitive processes behind the postulated facets of cognitive load should be obtained on a neural level by region-of-interest (ROI) predictions (Borst & Anderson, 2017) based on evidence from functional magnetic resonance imaging (fMRI). A pilot version of this model for the easy task condition was reported in Wirzberger, Rey, and Krems (2017) and has been expanded since to both conditions.

As a great strength, a cognitive modeling approach requires a precise formalization of human cognition, since it raises the need to decompose steps within the given task and related cognitive actions. Based upon the close compatibility with vested psychological evidence on human information processing, such offers the opportunity to derive well-founded explanations on behavioral phenomena. The idea of building computational models to explain cognitive phenomena has already been discussed by Wegener (1967), who outlined the indicative value of an electronic simulation of mental processes for deriving and validating the related hypotheses. Under the presumption of an existing analogy between model and psychological processes, it allows studying mental functions under conditions that would be difficult or even impossible to realize in human experiments. The constant interchange between experiment and simulation permits to verify and rethink given hypotheses on behavioral patterns and underlying cognitive strategies, and thus opens up the “black box”. In the context of the Cognitive Load Theory, there have been cognitive modeling accounts as well. Sweller (1988) used the production system PRISM (Langley & Neches, 1981) to explain cognitive load effects in problem-solving. He compared means-end and nonspecific goal strategies by determining the number of statements in working memory, the number of productions to fire, the number of conditions in productions to be matched, and the number of cycles to be executed. His conclusions indicate that the conventional means-end problem-solving strategy puts higher demands on cognitive resources and not necessarily fosters schema acquisition.

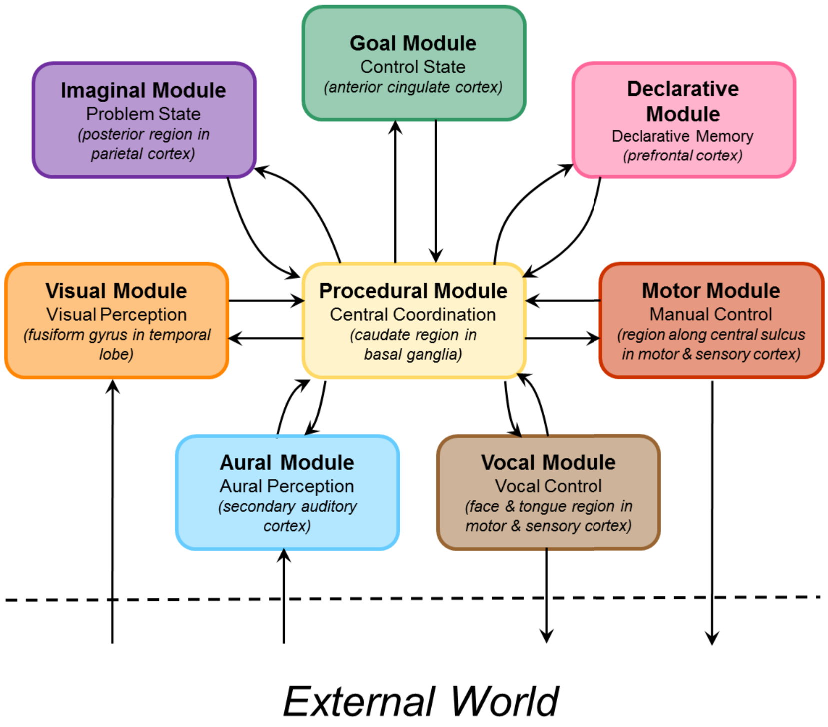

Figure 2. Overview of ACT-R core modules with corresponding brain regions. Based on Borst & Anderson (2015) and Anderson (2007).

Constituting a more prevalent and broadly used production-based approach, ACT-R is particularly characterized by its modular brain-inspired structure that is illustrated in *Figure 2*. The outlined modules represent goal planning (goal module), declarative memory (declarative module), intermediate problem states (imaginal module), action coordination (procedural module), the handling of visual and auditory inputs (visual and aural module), and motor and vocal outputs (motor and vocal module). Borst, Nijboer, Taatgen, van Rijn, and Anderson (2015) validated the mapping between these modules and corresponding ROIs in the human brain by fMRI data. For instance, when a model presses a button, increased activity in the motor module corresponds to activity in the motor cortex devoted to the representation of the hand. Whereas processes in different modules can be executed in parallel, known bottlenecks in information processing are represented by a limited capacitiy of a single information element per module at the same time (e.g., Borst, Taatgen, & van Rijn, 2010; Byrne & Anderson, 2001; Salvucci & Taatgen, 2008).

ACT-R relies on both symbolic and subsymbolic characteristics. The former involve the representation of declarative knowledge via so-called chunks of information and the interaction of defined modules through production rules. The latter constitute activation levels in declarative memory and utility of production rules. Chunks from declarative memory are retrieved based on their level of activation, which is calculated from the history and context of use and has to exceed a defined threshold to be eligible for selection. The full equation[1] for each chunk i involves the components displayed in *Equation 1:*

$$A_i = B_i + \sum_k \sum_j W_{kj} S_{ji} + \sum_l PM_{li} + \varepsilon. \quad (1)$$

The recency and frequency of use of the chunk i is reflected by the base-level activation B_i, W_{kj} represents the amount of activation from source j in buffer k, S_{ji} is the strength of association from source j to chunk i. W_{kj} and S_{ji} are summed over all buffers that provide spreading activation and all chunks in the slot of the chunks in buffer k. P reflects the amount of weighting given to the similarity in slot l and M_{li} represents the similarity between the value l in the retrieval specification and the value in the corresponding slot of chunk i. M_{li} is summed over the slot values of the retrieval specification. The value of ε represents noise, which is composed from an instantaneous component that is computed at the time of a retrieval request, and a permanent component that is associated with each chunk. Base-level activation is calculated as shown in *Equation 2*:

$$B_i = ln\left(\sum_{j=1}^{n} t_j^{-d}\right). \quad (2)$$

[1] Equations on chunk activation, base-level activation and utility relate to content described in the ACT-R reference manual and the tutorial units, available via http://act-r.psy.cmu.edu/software/.

It bases on the number of presentations *n* for the respective chunk i, the time t_j since the *jth* presentation, and a decay parameter *d*. Each time a chunk is presented, its base-level activation is increased, which decays as a power function of the time since that presentation. These decay effects are summed up and then transformed logarithmically.

Production rules consist of a condition part and an action part and are evaluated by the procedural module with regard to the content of the tested buffers. Based on the resulting pattern, a matching production rule is chosen, which triggers the related action. For instance, if the task is to react to a yellow number by key press, the visual module sees a yellow number, and the motor module is not in use, the action of pressing the key can be initiated. If more than one production rule fulfills the constraints, the selection of production i is informed by the subsymbolic cost-benefit mechanism of utility:

$$Probability(i) = \frac{e^{U_i\,/\,\sqrt{2s}}}{\sum_j e^{U_j\,/\,\sqrt{2s}}}. \tag{3}$$

It can be described as displayed in *Equation 3* by summing all productions j with expected utility values U_j that have matching conditions at the point of selection. Based on that, the production with the highest utility is chosen to fire.

3.1 Model concept

Each model run starts with an initial setting of the task goal, which is assumed to result from the previously read instruction. In the following, each learning trial builds upon three task-related steps, displayed in *Figure* 3. At first, the presented symbol is searched and encoded, which is repeated for the second symbol in the case of the difficult condition. This procedure stores an intermediate representation of all encoded visual content in the problem state, for instance, the input symbols 'square – circle' in the difficult condition. Next, the model attempts to retrieve the associated response symbol from declarative memory. In the second step, a response is selected from the provided opportunities on the screen, either according to the retrieved chunk or by random choice in case of no successful retrieval. The final step comprises the search for a visual feedback on the given response and, in the case of a false response, an update of the existing intermediate representation. The final information contains both the input and the correct response parts of the symbol combinations, such as 'square – circle – square' in case of the previous example. In the first trials, there is no sufficiently matching content or no content at all to retrieve, resulting in slower and less accurate responses. After being presented the input symbols several times and retrieving related content from declarative memory, the model performance gets increasingly faster and more accurate due to increasing chunk activation.

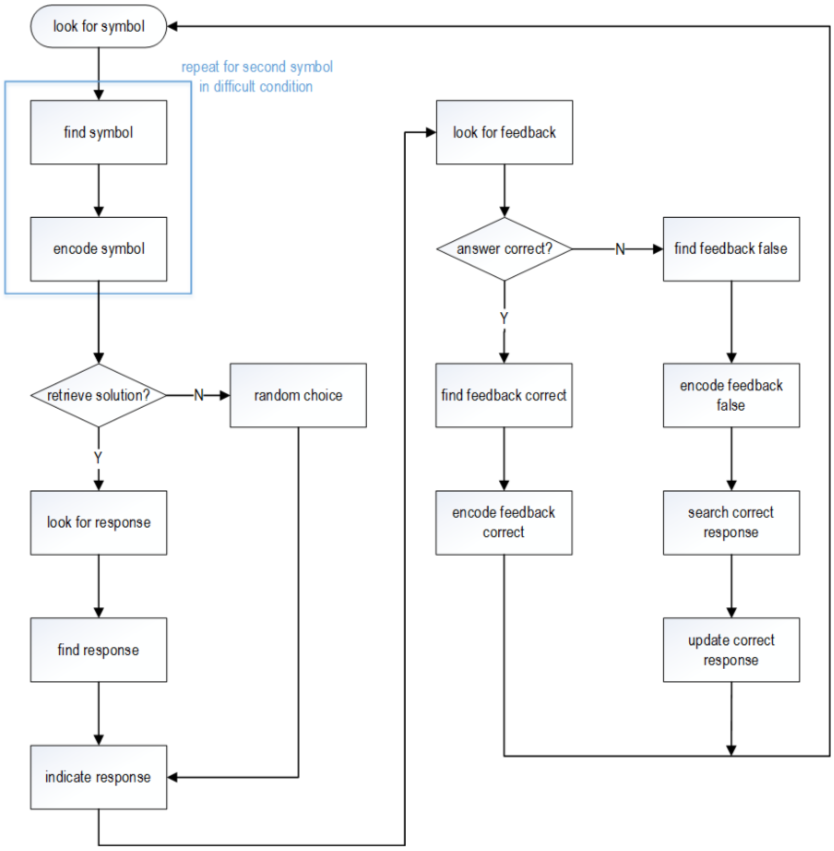

Figure 3. Outline of steps to perform in each the learning trial of the task. Adapted from Wirzberger, Rey, et al. (2017).

To account for the fact that humans sometimes retrieve related but ultimately wrong information from memory – in this case a wrong input-response association – ACT-R includes a partial matching mechanism. Based on initially defined similarities between chunks, a mismatch between request and actual retrieval is calculated. The higher the mismatch, the more the activity of the chunk is penalized (Lebiere, 1999). Increased interactivity between related elements of information (Sweller & Chandler, 1994; Sweller, 2010) is reflected in the spreading activation mechanism (Anderson, 2007) that distributes activation across chunks that share information elements. In the current task, spreading activation particularly effects the difficult task condition: Symbol combinations including the same input symbols, such as 'square – circle' and 'circle – square', obtain equal activation, independent of the correct symbol order.

The steps to be performed within the interrupting task are outlined in *Figure 4*. Following a goal change due to the bottom-up triggered saliency of the interrupting task, the task procedure involves the steps of searching, counting, and responding to the indicated target symbols. Using a color to indicate the task switch followed the model implemented by Wirzberger and

Russwinkel (2015) and represents the immediate attention to the related screen change. Tying in with evidence on pre-attentive and attentive processes in the visual module of ACT-R (Nyamsuren & Taatgen, 2013), the second visual-location request in the visual search is enhanced by additional information on stimulus color that relates to distinct characteristics of the presented symbols. In addition, counting was assumed to constitute a highly trained behavior that occurs almost automatically, thus a simple counting function was applied instead of intermediate retrievals after each counting step. After finishing the counting part, on each of the two response screens the model encodes the requested symbol and attempts to retrieves the potential answer. Again, due to the partial matching mechanism the possibility to retrieve a wrong answer persists.

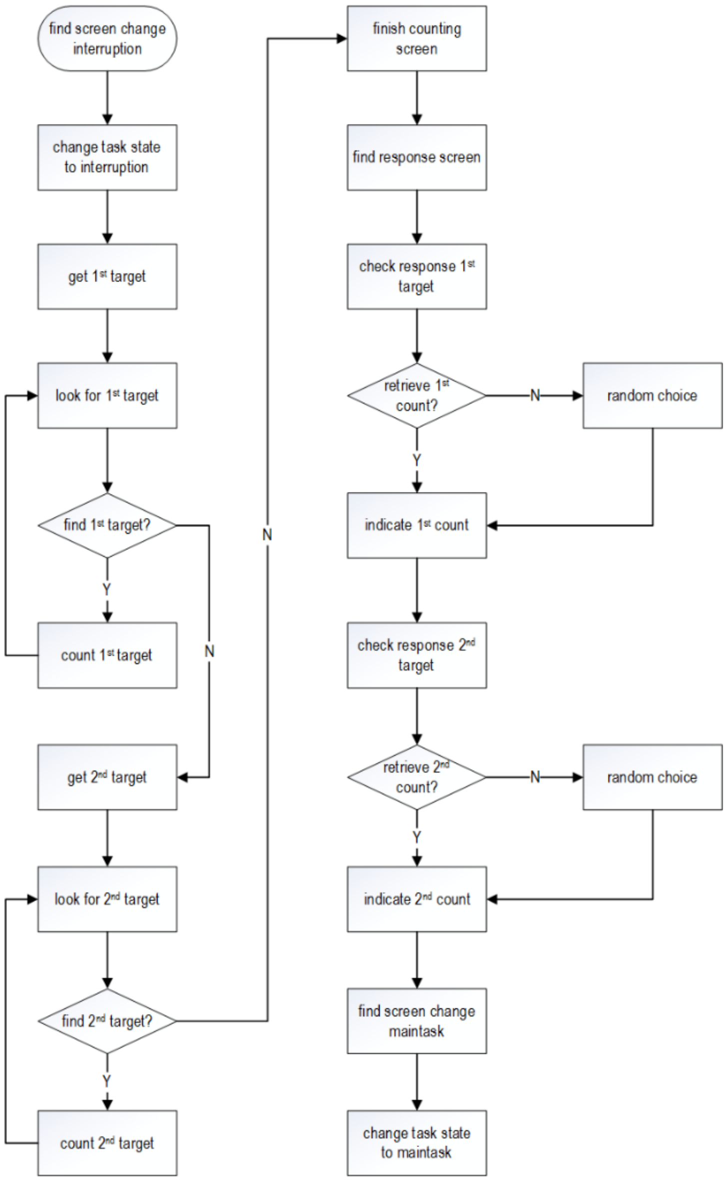

Figure 4. Outline of steps to perform in each occurrence of the interrupting task. Adapted from Wirzberger, Rey, et al. (2017).

Aligning to the available button selection on the screen, the model operates on an increased amount of visual-number finsts. Furthermore, due to the fixed order of the buttons on the screen, they can receive immediate attention without searching through all buttons from the top. When resuming the learning task, in line with Altmann and Trafton (2002) the model attempts to retrieve the previous task goal and thus restores its representation. Emerging interruption effects can be attributed to a decay in the activation of chunks related to the learning task that slows down subsequent retrieval requests (Borst, Taatgen, & van Rijn, 2010, 2015; Trafton, Altmann, Brock, & Minz, 2003). Across both tasks, more specific actions are regarded as more useful and thus receive a slightly higher utility, for instance, productions related to attending or encoding instead of searching around.

3.2 Model comparison

Parameter settings in the reported model are outlined in *Table 5* and correspond to the range of reported standard values (e.g., Anderson, Bothell, Lebiere, &, Matessa, 1998). In addition, the goal chunk related to the learning task received an initial base-level of 70 to account for the fact that participants received a comprehensive instruction on this task before. For the partial matching mechanism, similarities between symbols were set to -1, and for the spreading activation mechanism, content in the imaginal buffer was defined as source for spreading activation upon each retrieval from declarative memory. As already outlined, the increased amount of visual number finsts aligned to the button selection presented on the screen and received a value of 10.

Table 5

Parameter settings related to chunk activation and retrieval time

	:bll[a]	**:mp**[a]	**:mas**[a]	**:ans**[a]	**:rt**[b]	**:lf**[b]
Setting	0.5	0.401	1.7	0.2	0.11	3.8
Description	Base-level decay	Mismatch penalty	Maximum associative strength	Instantaneous noise	Retrieval threshold	Latency factor

Note. [a] Related to chunk activation. [b] Related to retrieval time (including retrieval failure).

Model data based on $n = 100$ model runs in each condition, since it was not the goal to create an exactly mapping model run for each human participant ($n_{\text{easy}} = 55$, $n_{\text{difficult}} = 58$), but rather to obtain robust conclusions from the average model performance. In addition, a close behavioral mapping in terms of interruption performance was not the core focus of the model, thus in the following only comparative results regarding the symbol learning task will be reported in detail. However, it was ensured that no crucial differences between both conditions persisted for the interrupting task. Both human performance and the currently reported model meet this constraint.

When comparing human and model data, beyond a graphical inspection Schunn and Wallach (2005) recommend the combined consideration of numerical goodness-of-fit indices related to the relative trend magnitude and the deviation from the exact location. They suggest R^2 to assess the relative trend magnitude, as it directly refers to the accounted proportion of variance and indicates a better fit by higher values. It is particularly suited to evaluate models with strong correlations to human data. For obtaining the deviation from the exact location, the root mean squared scaled deviation (RMSSD) constitutes a sophisticated approach:

$$RMSSD = \sqrt{\sum_{i=1}^{k} \frac{\left(\frac{m_i - d_i}{s_i / \sqrt{n_i}}\right)^2}{k}} = \sqrt{\sum_{i=1}^{k} \frac{(m_i - d_i)^2 n_i}{k s_i^2}}\,. \tag{4}$$

As obvious from *Equation 4*, the RMSSD scales the deviation between the model mean m_i for each point i and the data mean d_i for each point i by the corresponding standard error of this mean from human data ($s_i \,/\, \sqrt{n_i}$). The latter is calculated by the standard deviation s_i for each data mean i and the number of data values n_i contributing to each data mean d_i, whereas k is the number of points i. On this account, the RMSSD provides a scale invariant measure to evaluate the model fit in units of the standard error, with lower values indicating a better fit.

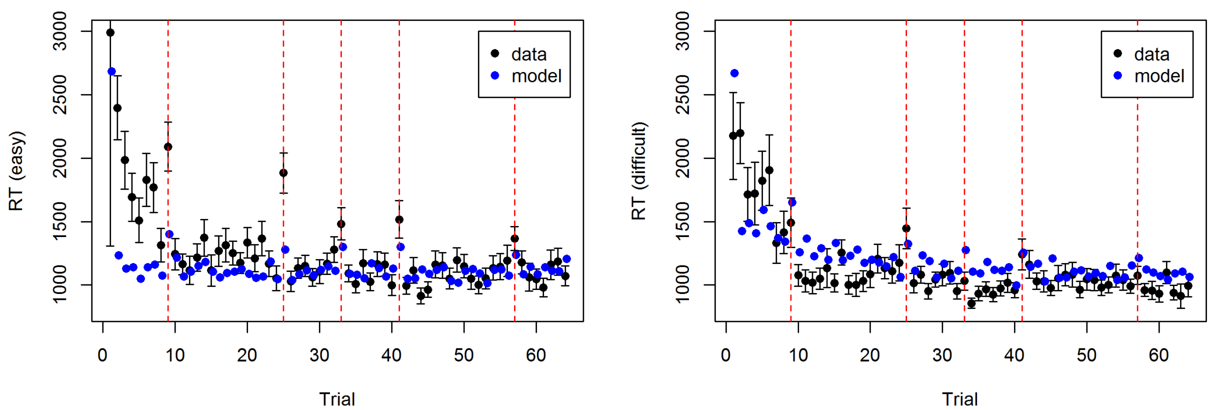

Figure 5. Reaction times for human data and model for the learning task in the easy and difficult condition (correct trials). Error bars indicate standard errors for human data. Red dashed lines indicate trials after an interruption.

In terms of reaction times, comparisons focused only on correctly solved trials. As obvious from *Figure 5*, interruption effects are observable in both conditions for human data, but still more distinctive in the easy task condition, as reported in Wirzberger, Esmaeili Bijarsari, et al. (2017). Standard errors are rather high for the first data point in the easy task condition, as only $n = 2$ observations fulfill the stated constraints. Besides the prevalence of interruption effects in both conditions, the visual inspection indicates that model data can map the initial decrease in reaction times in the difficult task condition, $RMSSD_{difficult} = 2.16$, $R^2_{difficult} = 0.58$. However, the model performs slightly slower than human participants during most of the trials. Apart from a subtler decrease in the beginning, the mapping fits quite well for later trials in the easy task condition, $RMSSD_{easy} = 1.67$, $R^2_{easy} = 0.52$.

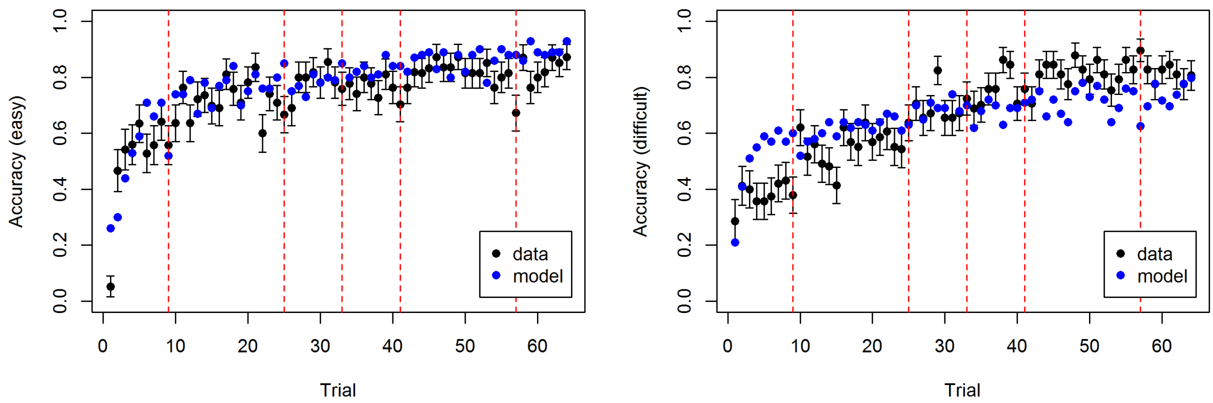

Figure 6. Accuracy for human data and model for the learning task in the easy and difficult condition. Error bars indicate standard errors for human data. Red dashed lines indicate trials after an interruption.

For accuracy, *Figure 6* indicates that the model can map the existing human behavior quite well in the easy task condition, $RMSSD_{easy} = 1.51$, $R^2_{easy} = 0.69$, although it achieves a higher performance in the end and shows a subtler reflection of interruption effects. The model in the difficult task condition learns slower compared to the easy task condition, but still faster than the human participants. However, apart from the nearly perfect location match in the last data points, it cannot fully map the final increase in the human data, $RMSSD_{difficult} = 2.07$, $R^2_{difficult} = 0.57$.

In addition, predefined ROI-predictions were generated (Borst & Anderson, 2017), based upon the previously amplified mapping of activity in ACT-R modules on defined brain regions. The underlying approach uses the recorded start and end times of module activity to simulate a signal comparable to the blood oxygenation level obtainable via fMRI, which shows peaks about 4-6 s after the occurrence of neuronal activity. In the first step, the activity of each inspected module is represented as 0-1 demand function and convolved afterward with the hemodynamic response function, displayed in *Figure 7*. As an example, related to the task of the current model, longer retrieval times due to lower levels of chunk activation would result in increased activity in the declarative module. Such patterns are expectable in early stages of the task, with increased task difficulty, or caused by interruption-related decay, and would be observable by higher peaks in the resulting simulated signal.

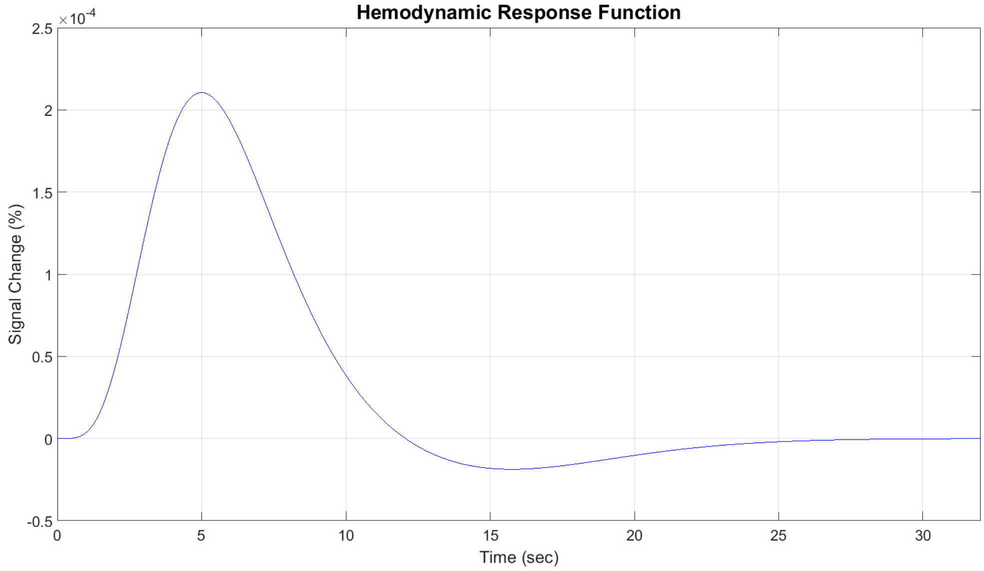

Figure 7. Hemodynamic response function (based on SPM). Adapted from Borst & Anderson (2017).

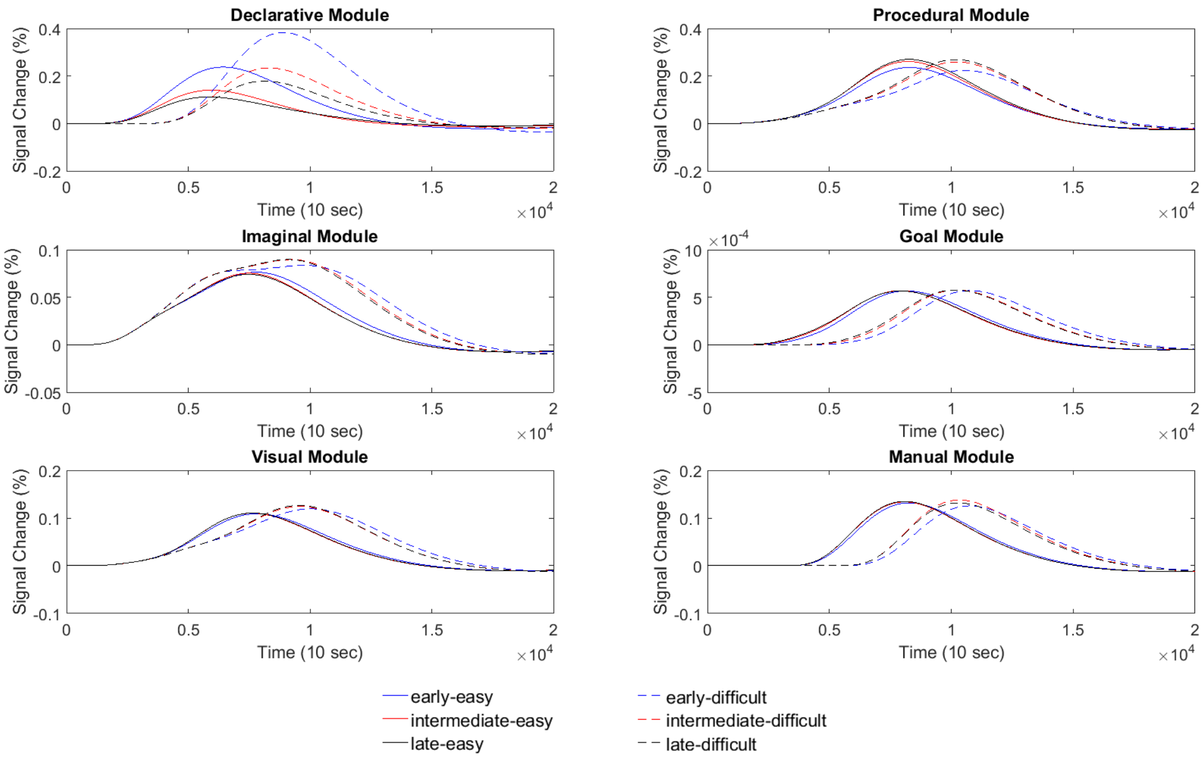

Figure 8. Module activity across different temporal stages of the symbol learning task (excluding resumption trials). Solid lines represent the easy task condition, dashed lines represent the difficult task condition. Blue lines represent trials in the early task stage (n = 20), red lines represent trials in the intermediate task stage (n = 19), and black lines represent trials in the late task stage (n = 20).

Prevalent changes in module activity due to task-inherent learning processes are displayed in *Figure 8.* The curves indicate a decrease in cognitive activity in later task stages in both conditions in the declarative module. The difficult task condition shows a higher level of activity across all stages, with a particularly distinctive peak across early task stages. Resulting activity in the imaginal module exerts a longer duration and shows a slightly increased level in the difficult condition. For the goal, procedural, visual and manual module levels of activity are rather comparable for both conditions, although the peaks occur later in the difficult condition.

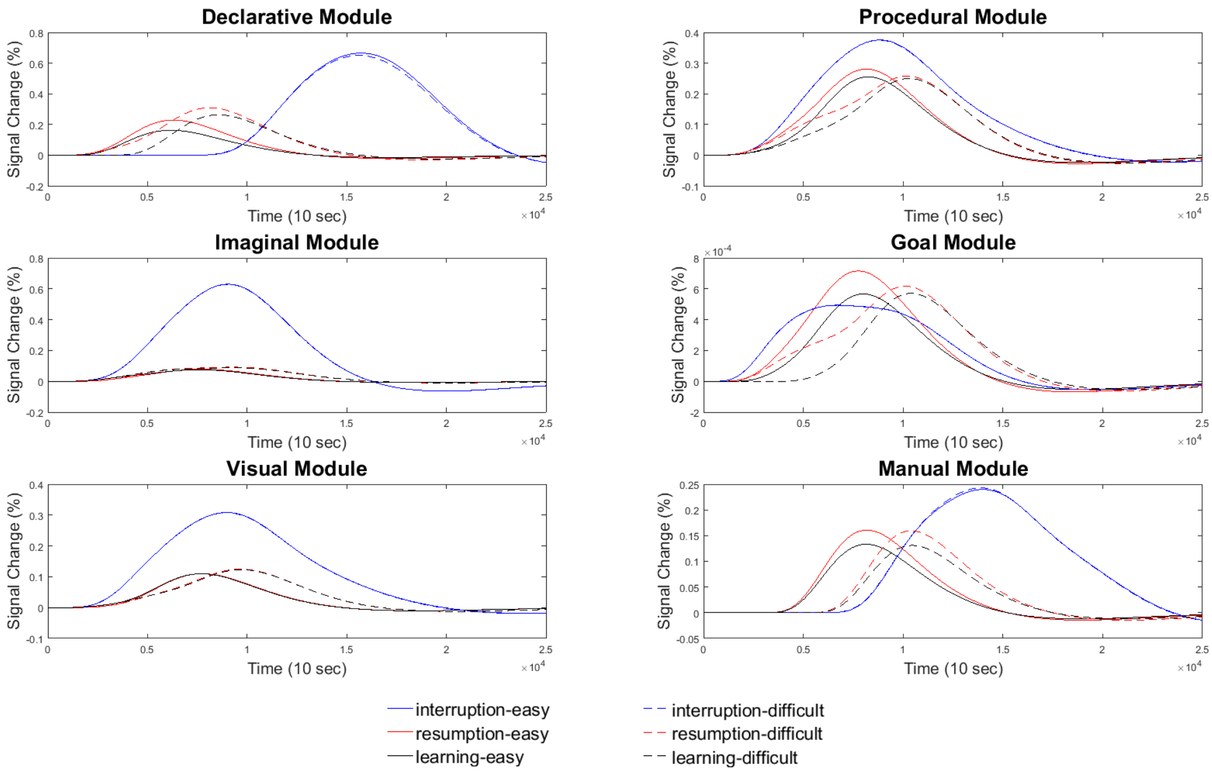

Figure 9. Module activity across interruption, resumption, and learning stages of the task. Solid lines represent the easy task condition, dashed lines represent the difficult task condition. Blue lines represent interruption trials (n = 5), red lines represent resumption trials (n = 5), and black lines represent learning trials (n = 59).

Comparisons between the interrupting task and the learning task are depicted in *Figure 9.* These include a separate visualization of the resumption phase, defined as the first trial that immediately follows the interrupting task. Across all inspected modules, activity levels in the interrupting task do not differ between both task conditions, since the solid and dashed blue lines overlap almost all the time. For the declarative, goal, procedural and manual module, a higher activity across resumption trials compared to the remainder of trials in the learning task results for both conditions. In addition, obvious differences between both task conditions show up during the resumption phase for the goal module and indicate higher levels of activity in the easy task condition. By contrast, no crucial differences between the resumption phase and regular learning trials result for the visual and imaginal modules. Apart from the goal module, the interrupting task always involves a higher level of activity, which peaks later in the declarative and manual modules.

3.3 Implications

The current model comprises a sophisticated approach to explore cognitive processes and mechanisms underlying changes in performance due to the inserted interruptions and task-related progress. In particular, the application of the spreading activation mechanism to map the theoretically introduced concept of element interactivity (Chen et al., 2017; Ngu et al., 2018; Sweller & Chandler, 1994; Sweller, 2010) offers the potential for deconstructing and formalizing effects of increased task complexity on a cognitive level. Inspecting model performance in the easy task condition in more detail, the obvious decrease in human performance in the final learning stage could potentially result from effects of boredom or fatigue. Modeling and explaining such effects would require a different model that also focuses on these effects (Gonzalez, Best, Healy, Kole, & Bourne Jr., 2011). In order to keep the current model focused and as simple as possible, this component was not included. For the difficult condition, model performance hints on an underlying shift in task-related strategies. Due to the small number of learned symbol combinations, over time people might have applied a more heuristic encoding strategy with focus on the first symbol, directly mapping task execution in the easy task condition. Explaining such strategy shift would result in a more complex model on the level of production rules and corresponding selection mechanisms. Taking this into account, the current modeling approach offers potential for future work, first by broadening the scope of the existing model and second by validating this model with new tasks. An additional benefit consists in explaining task order effects resulting from Wirzberger et al. (2018) with an additional ACT-R model that could build on the reported model. However, instead of dealing with an interrupting task, this model would face the constant requirement to simultaneously handle primary and secondary task procedures across both the visual and auditory modality.

As obvious from the ROI-analysis, the model needs to invest a higher amount of declarative memory resources upon each retrieval in the early task stage due to the lack of suitable chunks and lower levels of chunk activation. The smaller level of activity with increasing task progress emphasizes the prevalence of learning effects in both conditions, as existing content in declarative memory receives increasingly higher activation and thus can be retrieved faster and more accurate. In the difficult task condition, invested declarative resources are constantly higher across all stages, which by closer inspection relates to the increased influence of partial matching that penalizes chunk activation and extends retrieval times. It also corresponds well to the previously outlined conceptual approach by Whelan (2007). He attributed increased activity in the dorsolateral prefrontal cortex, a brain-region also connected to the declarative module, to higher levels of intrinsic cognitive load. The later peaks in activity in all modules in the difficult task condition potentially relate to attending and encoding an additional symbol, which, for instance, delays the onset of motor activity related to the response selection. Comparing activity patterns in both learning and interrupting tasks emphasizes the interrupting potential of the visual search task, since the activity in several modules clearly exceeds the activity during symbol learning. However, for both task conditions, task-related demands observable in the goal module are still higher in the learning task, hinting on more complex task-inherent control requirements. Similar activity patterns in both task conditions for the interrupting task reflect the absence of crucial differences between conditions and align to the pattern in human data reported by Wirzberger, Esmaeili Bijarsari, et al. (2017). The observable increased level of activity in the visual module during the interrupting task, which was supposed to trigger extraneous cognitive load, corresponds well to the reported activity in brain regions involved in sensory processing (Whelan, 2007). Finally, observable differences in goal activity during the resumption stage align well with predictions stated by the memory-for-goals model (Altmann & Trafton, 2002). They relate to the demand to rebuild the goal-representation of the learning task after each interruption, which also requires additional production rules, as reflected in increased procedural activity. Increased levels of resumption-related activity in the declarative module should arise from the decay of chunks related to the acquired symbol combinations. Finally, a reasonable explanation for the observable increase in motor-related activity in the resumption stage consists in the relocation of the mouse cursor from a different response screen.

4 Conclusions

The current thesis critically approached existing debates in cognitive load research related to the scope and interplay of distinct resource-demanding facets in instructional situations. Taken together, it emphasizes a process-related reconceptualization of the existing three-component model and underlines the importance of a combined inspection of different cognitive load measures. By extending the experimentally obtained behavioral results with a cognitive modeling approach, underlying cognitive processes and mechanisms could be inspected in more detail. The obtained insights further support the time-related reconsideration of the cognitive load facet framework, even on a neural level. With reference to applications in instructional situations, the resulting evidence can provide a vested foundation for the development of elaborated adaptive instructional procedures, both on the level of underlying algorithms and the design of instructional support. On this account, the research conducted within this thesis leverages a pathway to innovative approaches in the development of intelligent educational assistants.

References

Altmann, E. M., & Trafton, J. G. (2002). Memory for goals: an activation-based model, *Cognitive Science, 26*, 39–83. doi:10.1207/s15516709cog2601_2

Anderson, J. R. (1982). Acquisition of cognitive skill. *Psychological Review, 89*, 369-406. doi:10.1037/0033-295X.89.4.369

Anderson, J. R. (1983). *The architecture of cognition.* Cambridge, MA, USA: Harvard University Press.

Anderson, J. R. (2007). *How can the human mind occur in the physical universe?* New York, NY, USA: Oxford University Press.

Anderson, J. R., Bothell, D., Lebiere, C., & Matessa, M. (1998). An integrated theory of list memory. *Journal of Memory and Language*, *38*, 341–380. doi: 10.1006/jmla.1997.2553

Atkinson, C. & Shiffrin, R. M. (1971). The control of short-term memory. *Scientific American, 225*, 82-90. Stable URL: http://www.jstor.org/stable/24922803

Ayres, P. (2018). Subjective measures of cognitive load. What can they reliably measure?. In R. Zheng, (Ed.). *Cognitive load measurement and application: A theoretical framework for meaningful research and practice* (pp. 287-759). New York, NY, USA: Routledge.

Baddeley, A. D. (1992). Working memory. *Science, 255*, 556-559. doi:10.1126/science.1736359

Beckmann, J. F. (2010). Taming a beast of burden - On some issues with the conceptualisation and operationalisation of cognitive load. *Learning and Instruction, 20*, 250-264. doi:10.1016/j.learninstruc.2009.02.024

Bjork, E. L., & Bjork, R. A. (2011). Making things hard on yourself, but in a good way: Creating desirable difficulties to enhance learning. In M. A. Gernsbacher, R. W. Pew, L. M. Hough, & J. R. Pomerantz (Eds.), *Psychology and the real world: Essays illustrating fundamental contributions to society* (pp. 56–64). New York, NY, USA: Worth Publishers.

Borst, J. P., & Anderson, J. R. (2015). Using the ACT-R cognitive architecture in combination with fMRI data. In *An introduction to model-based cognitive neuroscience* (pp. 339-352). New York, NY, USA: Springer.

Borst, J. P., & Anderson, J. R. (2017). A step-by-step tutorial on using the cognitive architecture ACT-R in combination with fMRI data. *Journal of Mathematical Psychology, 76*, 94-103. doi:10.1016/j.jmp.2016.05.005

Borst, J. P., Nijboer, M., Taatgen, N. A., van Rijn, H., & Anderson, J. (2015). Using data-driven model-brain mappings to constrain formal models of cognition. *PLoS ONE, 10,* e0119673. doi: 10.1371/journal.pone.0119673

Borst, J. P., Taatgen, N. A., & van Rijn, H. (2010). The problem state: A cognitive bottleneck in multitasking. *Journal of Experimental Psychology: Learning, Memory and Cognition, 36*, 363-382. doi:10.1037/a0018106

Borst, J. P., Taatgen, N. A., & van Rijn, H. (2015). What makes interruptions disruptive? A process-model account of the effects of the problem state bottleneck on task interruption and resumption. In *CHI '15. Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (pp. 2971-2980). New York, NY, USA: ACM Press. doi:10.1145/2702123.2702156

Brickenkamp, R., Schmidt-Atzert, L., & Liepmann, D. (2010). *Test d2-Revision. Aufmerksamkeits- und Konzentrationstest*. Göttingen, Germany: Hogrefe.

Brünken, R., Plass, J. L., & Leutner, D. (2004). Assessment of cognitive load in multimedia learning with dual-task methodology: Auditory load and modality effects. *Instructional Science, 32*, 115-132. doi:10.1023/B:TRUC.0000021812.96911.c5

Brünken, R., Steinbacher, S., Plass, J., & Leutner, D. (2002). Assessment of cognitive load in multimedia learning using dual-task methodology. *Experimental Psychology, 49*, 109-119. doi:10.1027/1618-3169.49.2.109

Byrne, M. D., & Anderson, J. R. (2001). Serial modules in parallel: The psychological refractory period and perfect time-sharing. *Psychological Review*, *108*, 847–869. doi: 10.1037/0033-295X.108.4.847

Chen, F., Zhou, J., Wang, Y., Yu, K., Arshad, S. Z., Khawaji, A., & Conway, D. (2016). *Robust multimodal cognitive load measurement*. Bern, Switzerland: Springer International Publishing.

Chen, F., Zhou, J., & Yu, K. (2018). Multimodal and data-driven cognitive load measurement. In R. Zheng, (Ed.). *Cognitive load measurement and application: A theoretical framework for meaningful research and practice* (pp. 3643-4043). New York, NY, USA: Routledge.

Chen, O., Kalyuga, S., & Sweller, J. (2017). The expertise reversal effect is a variant of the more general element interactivity effect. *Educational Psychology Review, 29*, 393-405. doi:10.1007/s10648-016-9359-1

Cowan, N. (2010). The magical mystery four: How is working memory capacity limited, and why?. *Current Directions in Psychological Science, 19*, 51–57. doi:10.1177/0963721409359277

Csikszentmihalyi, M. (1990). *Flow: The psychology of optimal experience*. New York, NY, USA: Harper & Row.

de Jong, T. (2010). Cognitive load theory, educational research, and instructional design: some food for thought. *Instructional Science, 38*, 105-134. doi:10.1007/s11251-009-9110-0

D'Esposito, M., Onishi, K., Thompson, H., Robinson, K., Armstrong, C., & Grossman, M. (1996). Working memory impairments in multiple sclerosis: Evidence from a dual-task paradigm. *Neuropsychology*, *10*, 51. doi:10.1037/0894-4105.10.1.51

Ebbinghaus, H. (1964). *Memory: A contribution to experimental psychology*. Mineola, NY, USA: Dover Publications.

Ecker, U. K. H., Lewandowsky, S., Oberauer, K., & Chee, A. E. H. (2010). The components of working memory updating: an experimental decomposition and individual differences. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36*, 170-189. doi:10.1037/a0017891

Esmaeili Bijarsari, S., Wirzberger, M., & Rey, G. D. (2017). Lernförderliche Gestaltung computerbasierter Instruktionen zur Roboterkonstruktion. In M. Eibl, & M. Gaedke (Eds.), *INFORMATIK 2017, Lecture Notes in Informatics (LNI)* (pp. 2279-2286). Bonn, Germany: Gesellschaft für Informatik. doi:10.18420/in2017_228

Fitts, P. M., & Posner, M. I. (1967). *Human performance*. Belmont, CA: Brooks/Cole.

Gagné, R. M., & Dick, W. (1983). Instructional psychology. *Annual Review of Psychology, 34*, 261-295. doi:10.1146/annurev.ps.34.020183.001401

Galy, E., Cariou, M., & Mélan, C. (2012). What is the relationship between mental workload factors and cognitive load types?. *International Journal of Psychophysiology*, *83*, 269-275. doi:10.1016/j.ijpsycho.2011.09.023

Gerjets, P., Scheiter, K., & Schorr, T. (2003). Modeling processes of volitional action control in multiple-task performance: How to explain effects of goal competition and task difficulty on processing strategies and performance within ACT-R. *Cognitive Science Quarterly, 3*, 355-400.

Gillie, T., & Broadbent, D. (1989). What makes interruptions disruptive?: A study of length, similarity, and complexity. *Psychological Research, 50*, 243-250. doi:10.1007/BF00309260

Ginns, P. (2006). Integrating information: A meta-analysis of the spatial contiguity and temporal contiguity effects. *Learning and Instruction, 16*, 511-525. doi:10.1016/j.learninstruc.2006.10.001

Ghosh, V. E., & Gilboa, A. (2014). What is a memory schema?: A historical perspective on current neuroscience literature. *Neuropsychologia*, *53*, 104–114. doi:10.1016/j.neuropsychologia. 2013.11.010

Gonzalez, C., Best, B., Healy, A. F., Kole, J. A., & Bourne Jr., L. E. (2011). A cognitive modeling account of simultaneous learning and fatigue effects. *Cognitive Systems Research, 12*, 19-32. doi:10.1016/j.cogsys.2010.06.004

Herms, R. (2018). *Effective speech features for cognitive load assessment: Classification and regression* (Doctoral dissertation). Technische Universität Chemnitz, Germany.

Herms, R., Wirzberger, M., Eibl, M., & Rey, G. D. (2018). CoLoSS: Cognitive load corpus with speech and performance data from a symbol-digit dual-task. In N. Calzolari, K. Choukri, C. Cieri, T. Declerck, S. Goggi, K. Hasida, … T. Tokunaga (Eds.), *Proceedings of the 11th International Language Resources and Evaluation Conference* (LREC 2018) (pp. 4312-4317). Miyazaki, Japan: European Language Resources Association (ELRA).

Hoffman, B., & Schraw, G. (2010). Conceptions of efficiency: Applications in learning and problem solving. *Educational Psychologist, 45*, 1-14. doi:10.1080/00461520903213618.

Kalyuga, S. (2011). Cognitive load theory: How many types of load does it really need? *Educational Psychology Review, 23*, 1-19. doi:10.1007/s10648-010-9150-7

Kalyuga, S., & Chandler, P., & Sweller, J. (1999). Managing split-attention and redundancy in multimedia instruction. *Applied Cognitive Psychology, 13*, 351-371. doi: 10.1002/(SICI)1099-0720(199908)13:4<351::AID-ACP589>3.0.CO;2-6

Kalyuga, S., & Plass, J. L. (2018). Cognitive load as a local characteristic of cognitive processes. In R. Zheng, (Ed.). *Cognitive load measurement and application: A theoretical framework for meaningful research and practice* (pp. 1477-1884). New York, NY, USA: Routledge.

Kalyuga, S., & Singh, A. M. (2016). Rethinking the boundaries of cognitive load theory in complex learning. *Educational Psychology Review, 28*, 831-852. doi:10.1007/s10648-015-9352-0

Kennedy, D., & Scholey, A. (2000). Glucose administration, heart rate and cognitive performance: Effects of increasing mental effort. *Psychopharmacology, 149*, 63-71. doi: 10.1007/s002139900335

Klepsch, M., Schmitz, F., & Seufert, T. (2017) Development and validation of two instruments measuring intrinsic, extraneous, and germane cognitive load. *Frontiers in Psychology, 8*:1997. doi:10.3389/fpsyg.2017.01997

Korbach, A., Brünken, R., & Park, B. (2017). Differentiating different types of cognitive load: A comparison of different measures. *Educational Psychology Review, 30*, 503-529. doi:10.1007/s10648-017-9404-8

Kraiger, K., Ford, J. K., & Salas, E. (1993). Application of cognitive, skill-based, and affective theories of learning outcomes to new methods of training evaluation. *Journal of Applied Psychology, 78*, 311-328. doi:0021-9010/93/S3.00

Krell, M. (2015). Evaluating an instrument to measure mental load and mental effort using Item Response Theory. *Science Education Review Letters, 2015*(1), 1-6 Retrieved from urn:nbn:de:kobv:11-100228733

Langley, P., & Neches, R. (1981). *PRISM user's manual*. Technical report, Department of Psychology, Carnegie-Mellon University and Learning Research and Development Center, University of Pittsburgh, USA.

Lebiere, C. (1999). The dynamics of cognition: An ACT-R model of cognitive arithmetic. *Kognitionswissenschaft, 8*, 5–19. doi: 10.1007/BF03354932

Leppink, J., Paas, F., Van der Vleuten, C. P., Van Gog, T., & Van Merriënboer, J. J. (2013). Development of an instrument for measuring different types of cognitive load. *Behavior Research Methods, 45*, 1058-1072. doi:10.3758/s13428-013-0334-1

Leppink, J., Paas, F. G. W. C., Van Gog, T., Van der Vleuten, C. P. M., & Van Merriënboer, J. J. G. (2014). Effects of pairs of problems and examples on task performance and different types of cognitive load. *Learning and Instruction, 30*, 32–42. doi:10.1016/j.learninstruc.2013.12.001

Leppink, J., & Van Merriënboer, J. J. G. (2015). The beast of aggregating cognitive load measures in technology-based learning. *Educational Technology & Society, 18*, 230-245. Retrieved from http://www.jstor.org/stable/jeductechsoci.18.4.230

Manzey, D. (1998). Psychophysiologie mentaler Beanspruchung. In F. Rösler (Ed.), *Ergebnisse und Anwendungen der Psychophysiologie. Enzyklopädie der Psychologie, Band C/I/5* (pp. 799-864). Göttingen, Germany: Hogrefe.

Martin, S. (2018). A critical analysis of the theoretical construction and empirical measurement of cognitive load. In R. Zheng (Ed.), *Cognitive load measurement and application: A theoretical framework for meaningful research and practice* (pp. 778-1155). New York, NY, USA: Routledge.

Mayer, R. E. (2009). *Multimedia learning* (2nd ed.). New York, NY, USA: Cambridge University Press.

Monk, C. A., Trafton, J. G., & Boehm-Davis, D. A. (2008). The effect of interruption duration and demand on resuming suspended goals. *Journal of Experimental Psychology: Applied, 14*, 299-313. doi:10.1037/a0014402

Müller, C., Großmann-Hutter, B., Jameson, A., Rummer, R., & Wittig, F. (2001). Recognizing time pressure and cognitive load on the basis of speech: An experimental study. In *International Conference on User Modeling* (pp. 24–33). New York, NY, USA: Springer Science + Business Media.

Ngu, B. H., Phan, H. P., Yeung, A. S., & Chung, S. F. (2018). Managing element interactivity in equation solving. *Educational Psychology Review, 30*, 255-272. doi:10.1007/s10648-016-9397-8

Nyamsuren, E., & Taatgen, N. A. (2013). Pre-attentive and attentive vision module. *Cognitive Systems Research, 24*, 62-71. doi:10.1016/j.cogsys.2012.12.010

Paas, F. (1992). Training strategies for attaining transfer of problem-solving skill in statistics: A cognitive-load approach. *Journal of Educational Psychology, 84,* 429–434. doi:10.1037/0022-0663.84.4.429

Paas, F., Tuovinen, J. E., Tabbers, H. und Van Gerven, P. W. M. (2003). Cognitive load measurement as a means to advance cognitive load theory. *Educational Psychologist, 38*, 63-71. doi:10.1207/S15326985EP3801_8

Paas, F., & Van Gog, T. (2006). Optimising worked example instruction: Different ways to increase germane cognitive load. *Learning and Instruction, 16*, 87-91. doi:10.1016/j.learninstruc.2006.02.004

Park, B., & Brünken, R. (2018). Secondary task as a measure of cognitive load. In R. Zheng, (Ed.). *Cognitive load measurement and application: A theoretical framework for meaningful research and practice* (pp. 1901-2317). New York, NY, USA: Routledge.

Park, B., & Brünken, R. (2015). The rhythm method: A new method for measuring cognitive load – an experimental dual-task study. *Applied Cognitive Psychology, 29*, 232-243. doi:10.1002/acp.3100

Redick, T. S., Broadway, J. M., Meier, M. E., Kuriakose, P. S., Unsworth, N., Kane, M. J., & Engle, R. W. (2012). Measuring working memory capacity with automated complex span tasks. *European Journal of Psychological Assessment, 28*, 164-171. doi:10.1027/1015-5759/a000123

Renkl, A., Gruber, H., Weber, S., Lerche, T., & Schweizer, K. (2003). Cognitive Load beim Lernen aus Lösungsbeispielen. *Zeitschrift für Pädagogische Psychologie, 17*, 93-101. doi:10.1024/1010-0652.17.2.93

Rey, G. D., & Buchwald, F. (2011). The expertise reversal effect: cognitive load and motivational explanations. *Journal of Experimental Psychology: Applied*, 17, 33–48. doi:10.1037/a0022243

Rey, G. D., & Fischer, A. (2013). The expertise reversal effect concerning instructional explanations. *Instructional Science, 41*, 407–429. doi:10.1007/s11251-012-9237-2

Salvucci, D. D., & Taatgen, N. A. (2008). Threaded cognition: an integrated theory of concurrent multitasking. *Psychological Review*, *115*, 101–130. doi: 10.1037/0033-295X.115.1.101

Scheiter, K., Gerjets, P., & Heise, E. (2014). Distraction during learning with hypermedia: Difficult tasks help to keep task goals on track. *Frontiers in Psychology, 5*, 268. doi:10.3389/fpsyg.2014.00268

Schnotz, W., & Kürschner, C. (2007). A reconsideration of cognitive load theory. *Educational Psychology Review*, *19*, 469-508. doi:10.1007/s10648-007-9053-4

Schroeder, N. L., & Cenkci, A. T. (2018). Spatial contiguity and spatial split-attention effects in multimedia learning environments: A meta-analysis. *Educational Psychology Review, 30*, 679-701. doi:10.1007/s10648-018-9435-9

Schunn, C. D., & Wallach, D. (2005). Evaluating goodness-of-fit in comparison of models to data. In W. Tack (Ed.), *Psychologie der Kognition: Reden und Vorträge anlässlich der Emeritierung von Werner Tack* (pp. 115-154). Saarbruecken, Germany: University of Saarland Press.

Schweppe, J., & Rummer, R. (2014). Attention, working memory, and long-term memory in multimedia learning: An integrated perspective based on process models of working memory. *Educational Psychology Review, 26*, 285-306. doi:10.1007/s10648-013-9242-2

Seufert, T. (2018). The interplay between self-regulation in learning and cognitive load. *Educational Research Review, 24*, 116-129. doi:10.1016/j.edurev.2018.03.004

Shute, V. (2008). Focus on formative feedback. *Review of Educational Research, 78*, 153-189. doi:10.3102/0034654307313795

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science, 12*, 257-285. doi:10.1207/s15516709cog1202_4

Sweller, J. (2010). Element interactivity and intrinsic, extraneous, and germane cognitive load. *Educational Psychology Review, 22*, 123-138. doi:10.1007/s10648-010-9128-5

Sweller, J. (2018). The role of independent measures of cognitive load in cognitive load theory. In R. Zheng, (Ed.). *Cognitive load measurement and application: A theoretical framework for meaningful research and practice* (pp. 155-267). New York, NY, USA: Routledge.

Sweller, J., Ayres, P., & Kalyuga, S. (2011). *Cognitive load theory.* New York, NY, USA: Springer Science + Business Media.

Sweller, J., & Chandler, P. (1994). Why some material is difficult to learn. *Cognition and Instruction, 12*, 185-233. doi:10.1207/s1532690xci1203_1

Sweller, J., Van Merrienboër, J. J., & Paas, F. G. (1998). Cognitive architecture and instructional design. *Educational Psychology Review*, *10*, 251-296. doi:1040-726X/98/0900-0251S15.00/0

Trafton, J. G., Altmann, E. M., Brock, D. P., & Mintz, F. E. (2003). Preparing to resume an interrupted task: Effects of prospective goal encoding and retrospective rehearsal.

International Journal of Human-Computer Studies, 58, 583–603. doi:10.1016/S1071-5819(03)00023-5

Trick, L. M. (2008). More than superstition: Differential effects of featural heterogeneity and change on subitizing and counting. *Perception & Psychophysics, 70*, 743-760. doi:10.3758/PP.70.5.743

Unsworth, N., Heitz, R. P., Schrock, J. C., & Engle, R. W. (2005). An automated version of the operation span task. *Behavior Research Methods, 37*, 498-505. doi:10.3758/BF03192720

Visser, Raijmakers, & Molenaar (2002). Fitting hidden Markov models to psychological data. *Scientific Programming, 10*, 185-199. doi:10.1155/2002/874560

Vygotski, L. S. (1963). Learning and mental development at school age In B. Simon & J. Simon (Eds.), *Educational psychology in the U.S.S.R.* (pp. 21-34). London, UK: Routledge & Kegan Paul.

Wegener, H. (1967). Die Simulation menschlicher Intelligenz – Leistungen. In W. Kroebel (Ed.) *Fortschritte der Kybernetik* (pp. 375-385). Munich-Vienna, Germany-Austria: R. Oldenbourg Verlag.

Whelan, R. R. (2007). Neuroimaging of cognitive load in instructional multimedia. *Educational Research Review, 2*, 1-12. doi:10.1016/j.edurev.2006.11.001

Wickens, C. D. (2002). Multiple resources and performance prediction. *Theoretical Issues in Ergonomics Science, 3*, 159-177. doi:10.1080/1463922021012380 6

Wickens, C. D. (2008). Multiple resources and mental workload. *Human Factors, 50*, 449-455. doi:10.1518/001872008X288394

Wickens, C. D., Hollands, J. G., Banbury, S., & Parasuraman, R. (2013). *Engineering psychology and human performance* (4th ed.). Upper Saddle River, NJ, USA: Pearson Education.

Wirzberger, M., Beege, M., Schneider, S., Nebel, S., & Rey, G. D. (2016). One for all?! Simultaneous examination of load-inducing factors for advancing media-related instructional research. *Computers & Education, 100*, 18-31. doi:10.1016/j.compedu.2016.04.010

Wirzberger, M., Esmaeili Bijarsari, S., & Rey, G. D. (2017). Embedded interruptions and task complexity influence schema-related cognitive load progression in an abstract learning task. *Acta Psychologica, 179*, 30-41. doi:10.1016/j.actpsy.2017.07.001

Wirzberger, M., Herms, R., Esmaeili Bijarsari, S., Eibl, M., & Rey, G. D. (2018). Schema-related cognitive load influences performance, speech, and physiology in a dual-task setting: A continuous multi-measure approach. *Cognitive Research: Principles and Implications*. doi: 10.1186/s41235-018-0138-z

Wirzberger, M., Rey, G. D., & Krems, J. F. (2017). Modeling cognitive load effects in an interrupted learning task: An ACT-R approach. In G. Gunzelmann, A. Howes, T. Tenbrink, & E. J. Davelaar (Eds.), *Proceedings of the 39th Annual Meeting of the Cognitive Science Society* (pp. 3540-3545). Austin, TX, USA: Cognitive Science Society.

Wirzberger, M., & Rey, G. D. (2018). Attention please! Enhanced attention control abilities compensate for instructional impairments in multimedia learning. *Journal of Computers in Education, 5*, 243-257.doi:10.1007/s40692-018-0106-0

Wirzberger, M., & Russwinkel, N. (2015). Modeling interruption and resumption in a smartphone task: An ACT-R approach. *i-com, 14*, 147-154. doi:10.1515/icom-2015-0033

Yap, T. F. (2012). *Speech production under cognitive load: Effects and classification* (Doctoral dissertation). University of New South Wales, Sydney, Australia.

Zheng, R. (Ed.). *Cognitive load measurement and application: A theoretical framework for meaningful research and practice*. New York, NY, USA: Routledge.

Zheng, R. Z., & Greenberg, K. (2018). The boundary of different approaches in cognitive load measurement. In R. Zheng, (Ed.). *Cognitive load measurement and application: A theoretical framework for meaningful research and practice* (pp. 1179-1472). New York, NY, USA: Routledge.

List of tables

Synopsis

Article 1

Article 2

Article 3

List of figures

Synopsis

Article 1

Article 2

Article 3

Selbstständigkeitserklärung

Vorname, Name: Maria Wirzberger

Geburtsdatum: 15.12.1983

Geburtsort: Lohr am Main, Bayern

Ich erkläre hiermit gegenüber der Technischen Universität Chemnitz, dass ich die vorliegende Dissertation mit dem Titel *Load-inducing factors in instructional design: Process-related advances in theory and assessment* selbstständig und ohne Benutzung anderer als der angegebenen Quellen und Hilfsmittel angefertigt habe.

Die vorliegende Arbeit ist frei von Plagiaten, der elektronischen Überprüfung auf Plagiate stimme ich zu. Alle Ausführungen, die wörtlich oder inhaltlich aus anderen Schriften entnommen sind, habe ich als solche kenntlich gemacht.

Diese Arbeit wurde in gleicher oder ähnlicher Form noch nicht als Prüfungsleistung für den Erhalt eines akademischen Grades eingereicht.

Ort, Datum

Unterschrift

Statement of authorship

I hereby certify to the Technische Universität Chemnitz that this dissertation entitled *Load inducing factors in instructional design – Process-related advances in theory and assessment* is all my own work and uses no external material other than that acknowledged in the text.

This work contains no plagiarism, I agree to an electronical check for plagiarism. All sentences or passages directly quoted from other work or including content derived from such work have been specifically credited to the authors and sources.

This dissertation has not been submitted in the same or similar form for examination related to the award of an academic degree.