Modeling cognitive load effects in an interrupted learning task: An ACT-R approach

Maria Wirzberger (maria.wirzberger@phil.tu-chemnitz.de)
E-Learning and New Media, Chemnitz University of Technology,
Straße der Nationen 12, 09111 Chemnitz, Germany

Günter Daniel Rey (guenter-daniel.rey@phil.tu-chemnitz.de)
E-Learning and New Media, Chemnitz University of Technology,
Straße der Nationen 12, 09111 Chemnitz, Germany

Josef F. Krems (josef.krems@psychologie.tu-chemnitz.de)
Cognitive and Engineering Psychology, Chemnitz University of Technology,
Wilhelm-Raabe-Str. 43, 09120 Chemnitz, Germany

Without the knowledge of human cognitive processes, instructional design is blind. (Sweller, Ayres, & Kalyuga, 2011)

Abstract

Based on the established framework of Cognitive Load Theory, the presented research focuses on the inspection of cognitive load factors in an interrupted learning task. The task itself is inspired from basic cognitive research and demands participants to learn abstract symbol combinations of varying complexity. In addition, they have to deal with interruptions while performing the task. Experimental results indicate the influence of task complexity on how interruptions effect learning performance. However, questions on underlying learner cognition persist, rising the need for a more in-depth way of examination. For this purpose, a cognitive model within the cognitive architecture ACT-R is developed to clarify cognitive processes and mechanisms within different conditions of the task. Preliminary results from a first model for the easy task condition already indicate some fit between human and model data. Modeling work continues with adjusting the current model and implementing a model for the difficult task condition.

Keywords: Cognitive Load; Interruptions; Learning; ACT-R

Introduction

Learning constitutes an omnipresent requirement throughout the entire life, whether practicing to bring out the first words as a toddler, preparing for an exam within a course of study or gaining knowledge in a foreign language in mature age. When approaching learning from a psychological perspective, a variety of cognitive processes related to information capture, storage and retrieval come to the fore. They share the commonality to pose load on learners' limited mental resources, raising the need of well-designed instructional material. Such should support learners' efforts in acquiring the desired knowledge, skills and abilities without overloading their mental capacities.

Theoretical background

A prominent and often quoted theory in the field of instructional design is the Cognitive Load Theory (Sweller, 1988; Sweller, Ayres, & Kalyuga, 2011). It deals with the question how certain aspects of a learning scenario demand learners' cognitive resources. The theory postulates a practically unlimited storage capacity of long-term memory, the mental representation and organization of knowledge via schemata, and a limitation of working memory in terms of duration and capacity. In addition, mental resource demands in learning situations arise from different sources: Schema acquisition and automation build the core focus of each learning process and characterize the facet of *germane load*. Task complexity in relation to learners' previous knowledge constitutes *intrinsic load* and is traditionally defined in terms of related information that has to be processed simultaneously, referred to as element interactivity (Sweller, 2010). *Extraneous load* is increased by inappropriate instructional presentation and situational constraints. The latter comprise, for instance, aspects like performing the learning task in a distracting context with competing goals being present. The activation of such task-irrelevant information detracts cognitive resources needed for the learning task (Gerjets, Scheiter, & Schorr, 2003). In consequence, learners are prone to switch to simpler task-solving strategies that are less demanding, but at the same time less effective.

Cognitive Load Theory assumes that learning performance would be impaired if the sum of load imposed by the outlined factors exceeds the provided capacity of human working memory. However, the assumption of pure additivity has been questioned in more recent research (Park, 2010; Kalyuga, 2011; Wirzberger, Beege, Schneider, Nebel, & Rey, 2016), supporting the need for a theoretical reformulation. A possible time-related extension assumes that intrinsic and extraneous aspects affect performance on a structural and short-term level, while the germane aspect has to be considered on processual and long-term accounts (Wirzberger et al., 2016). In consequence, load induced due to schema acquisition should change over time, while structural load facets should pose a constant level of load. A further essential pre-assumption within the postulated framework comprises the fact that spare cognitive capacity is primarily devoted to foster schema acquisition.

Research focus

The overall project goal comprises to addresses cognitive processes behind the outlined facets of cognitive load. Within the subsection of research introduced in this paper, the particular influence of structural load components over various stages of the task is queried. In more detail, demands posed by increased task complexity and embedded interruptions are assumed to impair performance to different extents, depending on the achieved progress in the process of schema acquisition.

Experimental setting

A basic learning task was used to approach the research focus, facilitating the concise definition and control of experimental factors. Since it required no previous knowledge, potential confounding effects of this relevant predictor could be ruled out.

Methods

The experimental setting comprised 116 student participants (M_{age} = 23.25 years, SD_{age} = 4.34, range: 18-44 years, 80% female) from different courses of study. They were required to figure out and memorize four combinations of arbitrary geometric symbols within 64 trials while being interrupted five times over the task. Interruptions occurred at the same predefined points in time (i.e., after trials 8, 24, 32, 40 and 56) for reasons of comparability across participants. Symbol combinations were either *easy* (two symbols) or *difficult* (three symbols) and split up in input (one or two symbols) and response (always one symbol). Participants were randomly assigned to one of the two combination conditions, resulting in a between-subjects manipulation of task complexity.

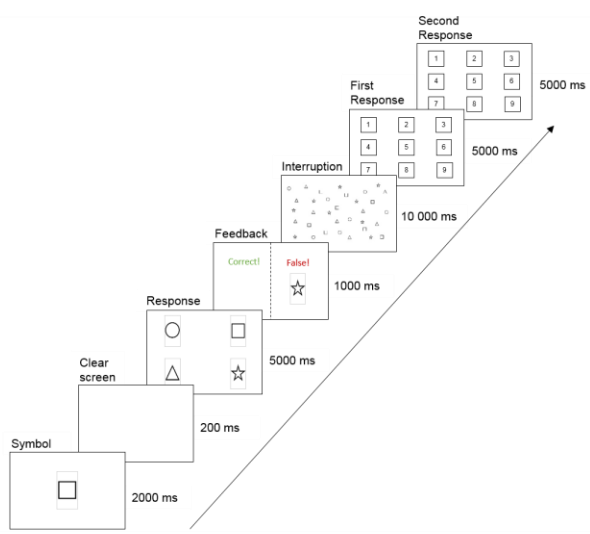

Figure 1: Schematic structure of a learning trial followed by an interruption in the easy task condition.

As depicted in *Figure 1*, in the learning part, symbols were presented one after another at the outset of each trial and participants had to indicate which symbol completed the combination. Responses were provided by selecting the correct symbol from an offered choice on the screen via mouse click. For instance, a square being displayed should result in choosing a star. After indicating their response, participants received feedback, as well as the correct solution in the case of an incorrect response. The target combinations represented the knowledge schemata that should be obtained over the task.

Within the interrupting secondary task, participants had to search, count and indicate two out of four types of geometric symbols from a visual search picture. Inspired by evidence from the subitizing task (Jensen, Reese, & Reese, 1950), seven to nine instances per symbol were displayed, to ensure that equal cognitive mechanisms were used across participants. Performance was recorded continuously during both subtasks via correctness and duration of responses.

Regarding the experimental design, performance efficiency computed as quotient from correct responses and reaction times in seconds (Hoffman & Schraw, 2010), represented the dependent variable. It reflected the amount of mental resources invested to acquire the task-related schema, characterizing the *germane load* component. Both structural load components were considered as independent variables: The number of symbols that defined a combination determined the *intrinsic load* component. Such a priori estimation of task complexity by the number of interacting elements followed Beckmann (2010) and Wirzberger et al. (2016). The interrupting secondary task represented the *extraneous load* component that was addressed in terms of inappropriate situational constraints.

Results

The influence of interruptions on task performance in both conditions was inspected by analyses of variance (ANOVAs) based on linear mixed models with Type III sums of squares and Satterthwaite approximation for degrees of freedom of fixed effects.

Results showed significant main effects of pre- vs. post-interruption performance, $F(1,118.12) = 16.71$, $p < .001$, and time of interruption occurrence over the task, $F(4,152.12) = 11.72$, $p < .001$. Moreover, significant interactions between condition and pre- vs. post-interruption performance, $F(1,118.12) = 16.86$, $p < .001$, and the condition and interruption occurrence, $F(4,152.12) = 11.75$, $p < .001$, were observed. Post-hoc pairwise comparisons with Tukey's HSD supported the pattern depicted in *Figure 2*. They indicated a loss in performance efficiency after facing an interruption, but only in the easy task condition. The entire model achieved a conditional pseudo-R^2 of .44, indicating about 44% of explained variance.

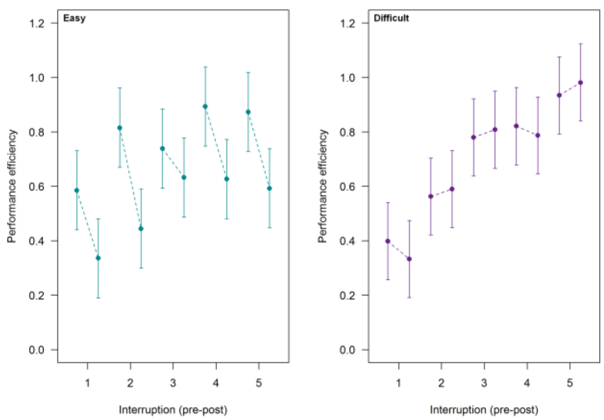

Figure 2: Changes in efficiency due to interruptions. Error bars indicate 95% confidence intervals.

In terms of interruption performance, a significant main effect showed up for interruption occurrence over the task, $F(4{,}464.77) = 12.53$, $p < .001$, while no significant difference between conditions was observable. Such pattern also receives visual support from *Figure 3*. The entire model obtained a conditional pseudo-R^2 of .36, indicating about 36% of explained variance.

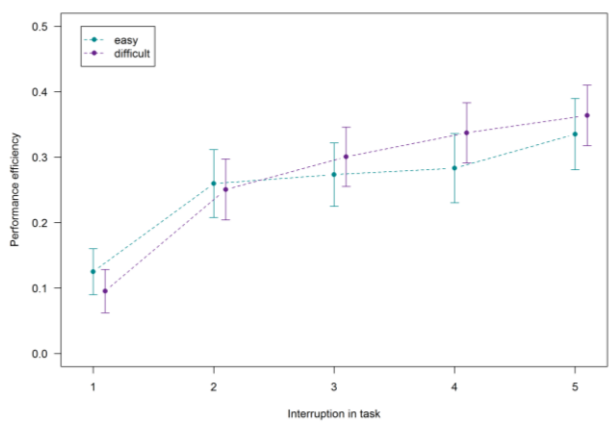

Figure 3: Changes in interruption performance over the task. Error bars indicate 95% confidence intervals.

By contrast, when comparing the amount of totally recalled and correctly recalled symbol combinations in both conditions, participant achieved nearly equal scores that did not differ significantly.

Discussion

Taken together, experimental results support influences of both structural load features on the observed task performance. However, the demand to inspect differences between conditions in more detail on a cognitive level arises. Although experimentally manipulated performance measurement provides a controlled way of assessment, it merely operates on indirect means and therefore lacks accessibility. On that point, the method of cognitive modeling becomes of value, since it offers the opportunity to clarify cognitive processes and mechanisms that underlie observable performance.

Cognitive modeling approach

Implementing a cognitive model structure raises the need to clearly think about each step within a given task and to ensure compatibility with founded psychological theories on human information processing. The cognitive architecture ACT-R (Anderson & Lebiere, 1998; Anderson, 2007) provides an elaborated cognitive modeling approach to establish a relationship between underlying biological structures and emerging patterns of behavior. It operates on a set of modules mapping the structure of the brain, illustrated in *Figure 4*. While the peripheral modules are responsible for handling visual and auditory inputs and motor and vocal outputs, the central modules focus on goal planning, declarative memory, intermediate problem states and action coordination (Anderson, 2007). The predicted BOLD responses in the corresponding brain regions, for instance the basal ganglia in terms of the procedural module, have already been validated by fMRI data (Borst & Anderson, 2015). Although processes in different modules can be executed in parallel, a limitation in capacity to one element at the same time exists, representing known bottlenecks in information processing resources (Borst, Taatgen, & van Rijn, 2010; Nijboer, Borst, van Rijn, & Taatgen, 2016).

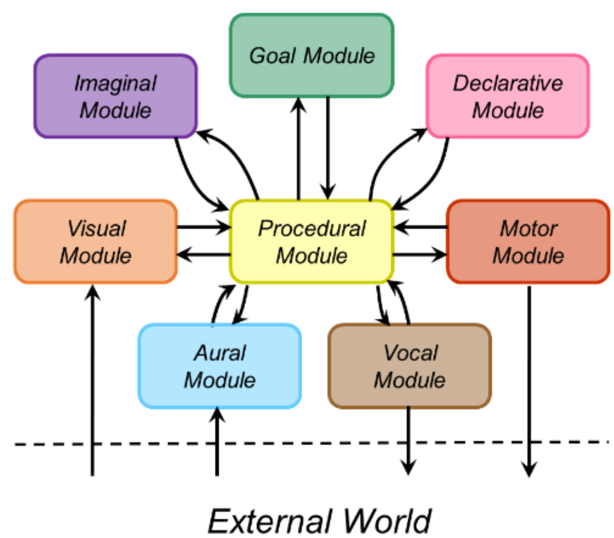

Figure 4: Overview of ACT-R core modules. Adapted from Borst & Anderson (2015) and Anderson (2007).

In contrast to other cognitive modeling approaches, ACT-R is characterized by applying both symbolic and subsymbolic features (Anderson, 2007). Amongst the symbolic aspects, information is stored and processed by chunks. Interaction between modules happens by selection of production rules in the procedural module that scans the content of the buffers and, based on the resulting pattern, chooses a suitable production rule that triggers the related action. If more than one production rule fits, the subsymbolic cost-benefit mechanism of utility decides, which production rule is selected. The level of activation, another important subsymbolic feature, reflects the availability of information in declarative memory and is determined by the context and history of use.

Model concept

A draft of the steps to be performed during the interrupting task and the learning trials are sketched in *Figure 5* and *Figure 6*. If an intended action cannot be finished within the given timeframe, the model can switch to the next logical step instead.

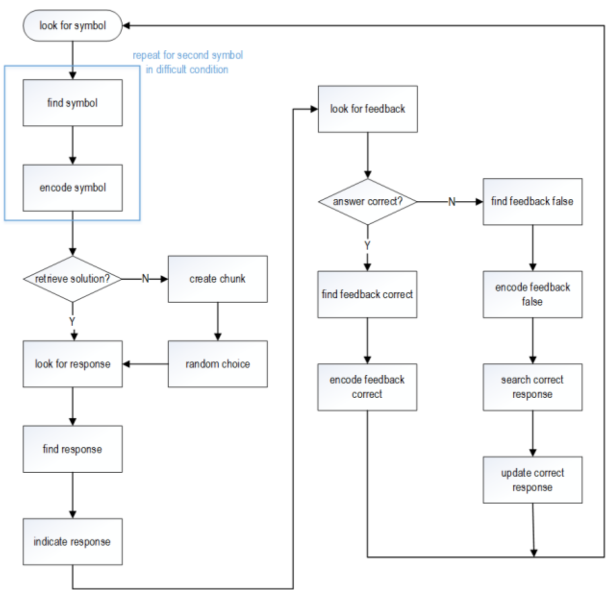

Figure 5: Outline of steps to perform in each learning trial of the task.

The concept of the cognitive model for the actual task setting is inspired by several sources of research. At first, Whelan (2007) framed a potential fMRI based measurement approach of the outlined cognitive load facets. In line with existing evidence from neuroimaging literature, he states that *extraneous load* triggers activity in particular in brain regions corresponding with sensory processing. Such aligns well to the extraneous load induction by a visual search task and is incorporated in the model due to the broad occupation of visual resources. The *intrinsic load* component is proposed to be associated with activity in brain regions responsible for maintaining and manipulating the attentional focus. In more complex tasks, entailing more interrelated elements, higher demands are posed on the corresponding goal and problem state resources. In addition, this provides a toe-hold for subsymbolic mechanisms like spreading activation, directly mapping the concept of activation distribution between related nodes of information.

Regarding the *germane load* facet, Whelan (2007) postulates a correspondence in particular with brain activation patterns representing motivation. This is plausible, since learners need to be motivated to dedicate available cognitive resources exclusively to schema acquisition. Based on that, in the difficult condition, a strategy shift towards a more heuristic encoding approach with increasing task progress is assumed. In detail, participants more and more tend towards retrieving the potential solution right with encoding the first symbol, which compensates for interruption costs and enables faster responses. Due to the resulting reduction in reaction time, they can achieve a better performance efficiency. The model incorporates such behavior by applying the subsymbolic mechanism of utility learning, which rewards each successful strategy adjustment.

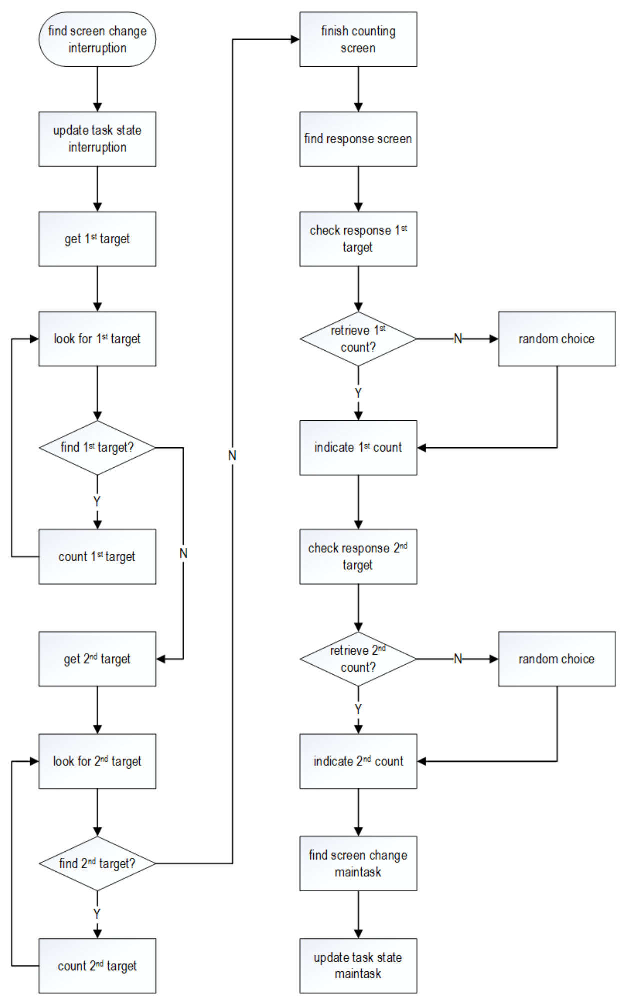

Figure 6: Outline of steps to perform in each occurrence of the interrupting task.

Beyond that, the model bases upon existing modeling work regarding interruption and resumption during task processing (Trafton, Altmann, Brock, & Minz, 2003; Wirzberger & Russwinkel, 2015). In brief, this tradition of research explains the loss in task performance after facing an interruption due to a decay in activation of the task representation. The resulting failure in accessibility of information can be adjusted within the model via subsymbolic chunk-related parameters like retrieval threshold, base-level decay or retrieval latency. On the perceptual level, the cognitive switch between both tasks is triggered bottom-up, at which the change in instruction color represents the salient screen change (Wirzberger & Russwinkel, 2015). On the processing level, due to this salience, the interrupting task receives immediate attention, represented by a high utility of the task switch. In addition, during both stages of the task, more specific actions are regarded as more useful, for instance attending and encoding available stimuli instead of just searching around. Thus, the related productions receive slightly higher utility and can be performed as soon as they match.

Related to the concept of memory activation is the important question, which components constitute working memory in ACT-R models. The current model follows a recently introduced approach by Nijboer et al. (2016), who discuss a multi-component working memory system that can explain memory interference in dual tasking. It involves the problem state as limited short-term resource to hold and manipulate information, the activated content of the declarative memory as well as the mechanism of subvocalized rehearsal as additional support to prevent activation decay. In particular processes of rehearsal are occupied to a greater extent in the difficult condition, potentially explaining the diverging patterns between conditions.

Preliminary results[1]

The currently available preliminary model is able to complete the easy task condition, highly demands visual perception, and already employs some subsymbolic parameters. Besides of an enabled base-level learning parameter, defaulting to the well-established value of 0.5, it operates on increased visual-number finsts, aligning to the available button selection on the screen. Moreover, it induces some instantaneous noise in retrieval-related activation to better account for human variability in memory performance.

Approaching the comparison between human and model data, aside from a graphical inspection, Schunn and Wallach (2005) recommend a combination of numerical goodness-of-fit measures on relative trend magnitude and those assessing deviation from the exact location. In particular, they approve R^2 as a measure of relative magnitude, for it relates directly to the accounted proportion of variance and better evaluates models with strong correlations to human data. In order to assess deviation from the exact location, the RMSSD (root mean squared scaled deviation) constitutes the measure of choice. It scales the deviation between each mean of human and model data by the corresponding standard error of the human data mean. In this vein, the RMSSD provides a scale invariant opportunity to assess model fit in units of the standard error.

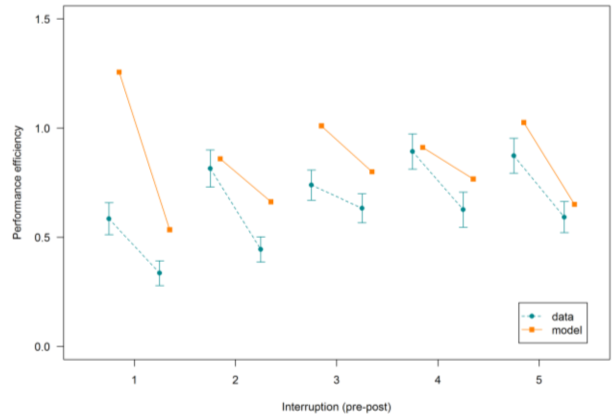

Figure 7: Comparison of performance in trials before and after an interruption.

At first glance, *Figure 7* indicates a reasonable fit in terms of impaired task performance due the induced interruptions. This impression receives support by the quite well RMSSD of 3.73 and an explained proportion of variance of 32 % (R = .32) for the selected pre-post interruption trials. When examining task performance in more detail, although the model can relatively map the given amount of correct responses during the learning trials, it shows a decreased match in terms of reaction time. The model constantly reacts much faster than human participants, which degrades the overall fit in performance efficiency. In addition, the model needs to better map interruption performance, indicated by *Figure 8* as well as the rather high RMSSD of 7.84 and smaller proportion of explained variance of 29 % (R^2 = .29).

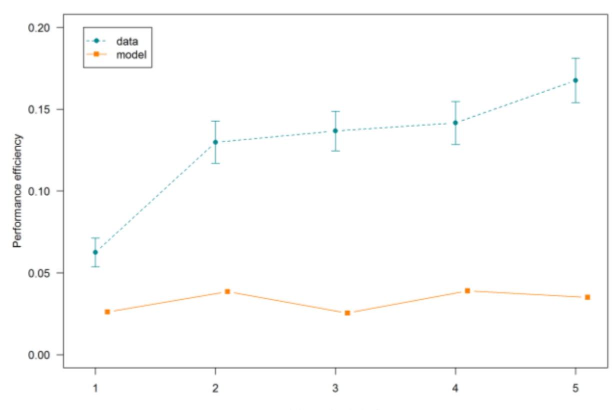

Figure 8: Comparison of performance in interruption task

As a potential solution, the model has to speed up counting during the visual search task, since mostly it is not

[1] Based on n = 55 model runs and n = 55 human participants.

able to successfully finish the second counting run or even end counting earlier.

Further steps

Pending steps within the ongoing modeling project involve the adjustment of outlined weaknesses in model performance. Moreover, due to the theoretical match with the concept of element interactivity, spreading activation has to be included as well. A second stream of work concerns the implementation of the difficult task condition. This involves the inclusion of productions that represent the aforementioned alternating task processing strategies.

Conclusion

Overall, this project constitutes an elaborated contribution to understanding cognitive processes that underlie knowledge acquisition from given instructional content. In doing so, it provides relevant insights into a so far rather vague defined theoretical framework, and additionally contributes to interconnect methodological approaches from different fields of research.

Acknowledgments

The presented research is conducted within the Research Training Group "CrossWorlds – connecting virtual and real social worlds" (GRK 1780/1). The authors gratefully acknowledge funding by the German Research Foundation (DFG).

References

Anderson, J. R., & Lebiere, C. J. (1998). *The atomic components of thought.* New York: Psychology Press.

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

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.

Borst, J. P., & Anderson, J. R. (2015). Using the ACT-R Cognitive Architecture in combination with fMRI data. In B. U. Forstmann, & E.-J. Wagenmakers (Eds.), *An Introduction to Model-Based Cognitive Neuroscience* (pp. 339–352). New York, NY, USA: Springer Science + Business Media.

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.

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.

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

Jensen, E. M., Reese, E. P., & Reese, T. W. (1950). The subitizing and counting of visually presented fields of dots. *Journal of Psychology, 30*, 363–392.

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

Nijboer, M., Borst, J., van Rijn, H., & Taatgen, N. (2016). Contrasting single and multi-component working-memory systems in dual tasking. *Cognitive Psychology, 86*, 1-26.

Park, B. (2010). *Testing the Additivity Hypothesis of Cognitive Load Theory* (Doctoral Dissertation), Saarbruecken, Saarland, Germany: University of Saarland.

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 Vortraege anlaesslich der Emeritierung von Werner Tack* (pp. 115-154). Saarbruecken: University of Saarland Press.

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

Sweller, J. (2010). Element interactivity and intrinsic, extraneous, and germane cognitive load. *Educational Psychology Review, 22*, 123–138.

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

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.

Whelan, R. R. (2007). Neuroimaging of cognitive load in instructional multimedia. *Educational Research Review, 2*, 1-12.

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.

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