PREPRINT OF ACCEPTED MANUSCRIPT

Memory-related cognitive load effects in an interrupted learning task: A model-based explanation

Maria Wirzberger[1], Jelmer P. Borst[2], Josef F. Krems[3], and Günter Daniel Rey[4]

[1]University of Stuttgart, Institute of Educational Science, Department of Teaching and learning with intelligent systems, Stuttgart, Germany

[2]University of Groningen, Bernoulli Institute for Mathematics, Computer Science and Artificial Intelligence, Groningen, The Netherlands

[3]Chemnitz University of Technology, Institute of Psychology, Cognitive and Engineering Psychology, Chemnitz, Germany

[4]Chemnitz University of Technology, Institute for Media Research, Psychology of learning with digital media, Chemnitz, Germany

Corresponding author: Maria Wirzberger
Email: maria.wirzberger@ife.uni-stuttgart.de
Phone: +49 711 / 685 – 81176
University of Stuttgart
Institute of Educational Science
Department of Teaching and learning with intelligent systems
Geschwister-Scholl-Str. 24D
70174 Stuttgart
Germany

Abstract

We use a cognitive model to explain memory-related cognitive strategies and mechanisms in an interrupted learning task. The model was developed in the ACT-R cognitive architecture and incorporates ideas from the well-established Cognitive Load Theory. This theoretical framework deals with the question to what extent aspects of learning situations demand learners' working memory resources. Our model was developed to account for human data of an experiment that required participants to learn easy or difficult combinations of arbitrary geometric symbols while being interrupted by a secondary task. The interrupting task requested participants to search, count, and remember the amount of indicated targets from a screen with geometric symbols. The results showed that participants having to learn more complex symbol combinations learned slower but reached a comparable level of performance in the end. In addition, participants in this condition were less impaired by the induced interruptions. To clarify potential underlying memory-related strategies and mechanisms, the cognitive architecture ACT-R was used to develop four computational models of the task. Our models explore the necessity of including different subsymbolic mechanisms to explain the data. The final model, which included spreading activation and partial matching mechanisms, was used to generate fMRI predictions. These predictions provide a link to the Cognitive Load Theory and contribute to the discussion of underlying mechanisms on a neural level. In conclusion, the presented work emphasizes the influence of task characteristics and supports a process-related view on cognitive load in instructional scenarios.

Keywords: Cognitive load theory; Cognitive modeling; Learning performance; Element interactivity; Task interruption

1 Introduction

Learning is an essential part of human experience over the entire life course, whether trying to ride a bike as a child or learning a new language as an adult. Characteristically, the establishment of the related knowledge structures does not occur all at once but gradually over time. This process is accompanied by changes in performance that correspond to variations in cognitive resource demands. But what cognitive mechanisms are responsible for these changes and how do task-inherent and situational features, such as interruptions, affect the underlying dynamics?

To investigate these issues, we present a cognitive modeling approach that leverages the cognitive architecture ACT-R (Anderson, 2007) to model learner behavior in an interrupted learning task (cf. Fechner et al., 2016). The underlying task setting employed a controlled stimulus-response learning paradigm to put the interplay of load-inducing factors under an experimental microscope. Our basic learning material consisted of four easy or difficult symbol combinations that were presented repeatedly over the task. The included symbols were split up in an input part, that was shown at the beginning of a trial, and a response part, where the corresponding symbol had to be selected. Complexity was determined by the number of symbols in a defined order that formed a combination. Inducing additional situational constraints, the learning process was interrupted several times by an embedded secondary task: a visual search screen with a symbol counting request. As a theoretical framework for model development, we use the well-established Cognitive Load Theory (e.g., Sweller, Van Merrienboër, & Paas, 1998).

To summarize the results: we can connect both learning and interruption effects to activation patterns of memory structures to explain changes in performance. A novel correspondence emerges between the theoretically postulated concept of element interactivity of the Cognitive Load Theory and the spreading activation mechanism in ACT-R. Our simulated fMRI predictions further support a process perspective of decreasing cognitive load with the increasing establishment of stable knowledge structures on a neural level.

We will now first discuss the relevant theoretical and methodological background, outline the experimental task in more detail, and explain the underlying modeling concepts. Subsequently, we summarize the experimentally obtained behavioral patterns from the human data and compare them with the results from our cognitive model. With four alternative versions of our cognitive model, we explore the relevance of defined mechanisms to explain the observed patterns from the human data. We further address the performance of the final model on a neural level by reporting results from simulated fMRI analyses and link these to predictions of the Cognitive Load Theory.

1.1 Cognitive load theory

When instructing someone in a new task, 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). It builds on vested models of memory (Anderson, 1983; Atkinson & Shiffrin, 1971; Baddeley, 1992) that indicate limited working memory resources in terms of both duration (around 20 to 30 s; see e.g., Wickens, Hollands, Banbury, & Parasuraman, 2013) and capacity of stored information (around four; see e.g., Cowan, 2000). In contrast, long-term memory resources provide nearly infinite storage capacity and duration and can be used to establish permanent knowledge structures. These 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. With reference to established models of learning and skill acquisition, higher amounts of resource investment are plausible in earlier process stages when knowledge structures still need to be established (Anderson, 1982; Fitts & Posner, 1967; Ebbinghaus, 1964). The initially declarative knowledge becomes increasingly procedural and automated with task progression and, in consequence, demands less cognitive resources (Tenison & Anderson, 2015; Tenison, Fincham, & Anderson, 2016).

In addition, the theory postulates the distinct facets of extraneous, intrinsic, and germane cognitive load that relate to different sources in learning situations. Extraneous cognitive load refers to harmful effects from task-irrelevant features (Sweller, 1988). It arises from a complex instructional presentation or interfering situational aspects, such as a competing goal from an interrupting task (Gerjets, Scheiter, & Schorr, 2003). Intrinsic cognitive load relates to the complexity of the used learning material in relation to existing previous knowledge (Sweller & Chandler, 1994). In addition, the concept of element interactivity was introduced. Element interactivity in learning material increases with the number of logically related information units, for example, the number of symbols in a defined order, that must be processed simultaneously in working memory (Chen, Kalyuga, & Sweller, 2017; Ngu, Phan, Yeung, & Chung, 2018; Sweller, 2010). In previous research, this aspect has been addressed experimentally with a priori estimates based on the number of interrelated dimensions or elements that a learner had to deal with at the same time (e.g., Beckmann, 2010; Wirzberger, Beege, Schneider, Nebel, & Rey, 2016; Wirzberger, Esmaeili Bijarsari, & Rey, 2017). Characteristically, changes in element interactivity are related to the nature of what is learned, and its amount should be kept at a manageable level for the individual learner to foster optimal learning outcomes (Sweller, 2010). Finally, germane cognitive load emerges from the establishment of knowledge structures itself, which is specified as schema acquisition and automation in the theoretical framework (Paas, Tuivonen, Tabbers, & Van Gerven, 2003; Sweller et al., 1998). An increasing prevalence of schemata that can be applied to solve a task would be evident with improved performance under less invested effort. Related to the current task setting, a learner might achieve accurate reactions on the presented symbols increasingly faster over time. Germane cognitive load operates under the assumption of highly motivated learners that devote all available resources to the establishment of mental knowledge structures.

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). Due to difficulties in empirical assessment and psychometric separation, this assumption of additivity is increasingly queried (e.g., Kalyuga, 2011; Kalyuga & Singh, 2016). Therefore, recent evidence suggests a process-driven reconceptualization of the originally postulated three-component model (Wirzberger, Esmaeili Bijarsari, & Rey, 2017; Wirzberger, Herms, Esmaeili Bijarsari, Eibl, & Rey, 2018). It takes into account that the previously outlined facets represent distinct ontological categories (De Jong, 2010). On the one hand, the intrinsic and extraneous cognitive load are related to static structural aspects of the learning context that can be determined a priori. On the other hand, the germane cognitive load has to be considered on a dynamic process-related level, as it changes with increasing progress in the learning task. Support for separate consideration and dynamic view on changes in cognitive load during learning also arises from research on self-regulated learning. For example, Seufert (2018) emphasizes that both the intrinsic and extraneous cognitive load facet represent task affordances, which are imposed by the learning material, whereas the germane cognitive load refers to learner-based decisions.

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 the initial description of the theory, 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. Amongst the earliest attempts to provide insights into cognitive demands prevalent in a learning situation, retrospectively applied subjective rating scales are a broadly used approach. For example, the unidimensional mental effort scale by Paas (1992) requires participants to rate their perceived mental effort on a nine-point Likert scale ranging from ‘very, very low (1)’ to ‘very, very high (9)’. Cognitive resource demands are further observable in continuous performance-related parameters, such as reaction times, error rates, and accuracy. In instructional cognitive load research, these parameters have been inspected in dual-task paradigms (e.g., Brünken, Plass, & Leutner, 2004). Additional means of continuous assessment employ physiological markers related to heart rate, galvanic skin response, pupil diameter, or electroencephalographic frequency band power (e.g., Scharinger, Kammerer, & Gerjets, 2015). Contrary to these measures, the occurrence of behavioral indices can mostly or entirely be controlled by the learner (Chen et al., 2016). Relevant parameters of inspection arise from speech signals, gaze patterns or mouse movements. For example, existing research shows that higher levels of cognitive load raise planning processes that result in a slower speech tempo and more frequent and longer pauses in the speech flow (e.g., Herms, Wirzberger, Eibl, & Rey, 2018; Müller, Großmann-Hutter, Jameson, Rummer, & Wittig, 2001).

A 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 neural 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 (e.g., Meredith, 2001; Whelan, 2007). 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 (e.g., Banich et al., 2000; Miller & Cohen, 2001; Whelan, 2007). Finally, germane cognitive load is assumed to hold connections to activity in brain regions related to reward, which fosters motivation. Highly motivated learners are more likely to devote available cognitive resources solely to processes and strategies of schema acquisition. Corresponding brain regions involve the superior frontal sulcus and the intraparietal sulcus (e.g., Taylor, Welsh, Wager, Phan, Fitzgerald, & Gehring, 2004; Whelan, 2007). 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. A step towards testing this framework leverages existing approaches in model-based cognitive neuroscience that simulate fMRI data from activity patterns emerging from a cognitive model.

1.2 The cognitive architecture ACT-R

Although the Cognitive Load Theory proposes several important concepts in learning, it does not explain how learning itself works. To connect this conceptual theory directly to human learning, we simulated it in the cognitive architecture ACT-R (Anderson, 2007). This requires a precise formalization of the necessary steps related to solving a given task into underlying cognitive actions.

ACT-R is a prevalent and broadly used production-based modeling approach, that is particularly characterized by its modular brain-inspired structure. The included 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). Buffers serve as an interface between modules and processes in different modules can be executed in parallel. Known bottlenecks in information processing are represented by the limited capacity of a single information element per module at the same time (e.g., Borst, Taatgen, & van Rijn, 2010; Byrne & Anderson, 2001; Salvucci & Taatgen, 2008).

All of ACT-R's modules have been linked to certain brain regions and can be used to predict fMRI data (e.g., Anderson, 2007; for a recent data-driven validation see Borst, Nijboer, Taatgen, van Rijn, & Anderson, 2015). For instance, when a model retrieves information from declarative memory, the increased activity in the *declarative module* corresponds to activity in a region around the inferior frontal gyrus in the prefrontal cortex. This area has proven to be sensitive to both retrieval and storage operations. Efforts made in transforming mental representations, such as the number of targets on a visual search screen during counting, are reflected in activity in the *imaginal module*. It is associated with a region close to the intraparietal sulcus in the posterior parietal cortex, which is sensitive to representational changes. Following a multi-component perspective, both modules are considered part of ACT-R's working memory (Nijboer, Borst, Van Rijn, & Taatgen, 2016).

To decide which cognitive action to perform next, ACT-R applies production rules. Production rules consist of a condition part and an action part and are evaluated by the procedural module with regards 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 previously displayed symbol, the visual module encodes the matching symbol button on the screen, and the motor module is not in use, *then* the mouse can be moved to the button and the mouse click can be initiated. A chosen production rule takes 50 ms to be executed.

Characteristically, ACT-R uses a hybrid approach of both symbolic and subsymbolic mechanisms: chunks of information from declarative memory are retrieved not only on the match of content in their slots but also based on their level of activation. Activation is calculated from the history and context of use of a chunk and has to exceed a defined threshold to be eligible for selection. The full equation for each chunk i involves the components displayed in *Equation 1:*

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

The recency and frequency of use of the chunk i is reflected by the *base-level activation* B_i. As outlined in *Equation 2*, 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. Thus, the more often certain information is experienced, for instance, a 'square – circle' combination, the better it is retained in memory.

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

With the *spreading activation* mechanism, ACT-R accounts for the influence of the current context in retrieving information from memory. That is, related information in ACT-R's buffers will spread information to chunks in memory. For example, if 'square' is encoded in the *imaginal module*, this will make it more likely to retrieve chunks containing 'square'. In some situations, this can make it harder to retrieve the correct information. For example, in the current task, the 'square – circle' and 'circle – square' combinations in the difficult condition contain the same symbols and would receive equal activation upon a retrieval request. In *Equation 1* W_{kj} represents the amount of activation from source j in buffer k, and S_{ji}, the strength of association from source j to chunk i. A source of spreading activation can be any slot of a retrieved chunk, for example, the values 'square' and 'circle' in the slots of a symbol combination chunk. W_{kj} and S_{ji} are summed over all buffers that provide spreading activation, such as the goal and/or imaginal buffer, and all chunks in the slots of the chunks in buffer k.

As humans sometimes retrieve related but ultimately wrong information from memory (so-called 'error of commission'), in the current task a wrong input-response symbol combination, ACT-R further 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). For example, when trying to retrieve 'star – triangle', 'circle – square' would be penalized more than 'circle – triangle'. In *Equation 1*, *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 that is computed at the time of a retrieval request.

The time to retrieve a chunk from memory is also dependent on its activation, as shown in *Equation 3*:

$$Time = Fe^{-A} \ . \tag{3}$$

A represents the activation of the retrieved chunk, while *F* is the latency factor. The higher the activation of a chunk, the faster it can be retrieved. A retrieval failure occurs if no chunk has an activation above the retrieval threshold. To calculate the time until the retrieval failure is reported, the chunk activation is replaced by the retrieval threshold τ, as displayed in *Equation 4*:

$$Time = Fe^{-\tau} \ . \tag{4}$$

Thus, the response time in the current task depends on if and how fast a symbol combination can be retrieved from declarative memory.

2 Methods

2.1 Task

The task setting and the related human data from a sample of 116 student participants were already reported in Wirzberger et al. (2017). As displayed in *Figure 1*, the participants in this study had to learn abstract geometric symbol combinations via trial and error by verifying feedback (Shute, 2008). They completed 64 trials in total that were equally distributed across the four potential symbol combinations. The symbol combinations itself consisted of an input part, which was presented at the beginning of a trial, and a subsequent response part. Following the between-subjects manipulation of intrinsic cognitive load, participants were randomly assigned to either the easy or difficult condition. In the easy condition, the input part included one symbol, while in the difficult condition two symbols were presented one after another. They always differed from each other, i.e., there were no combinations such as 'square – square'. Corresponding to the concept of increased element interactivity, in the difficult condition the order of the displayed input symbols mattered as well. For example, 'square – circle' resulted in a different result compared to 'circle – square'. In the response part, the participants had to choose the symbol that completed the combination from the four potential response options by clicking on the matching symbol button. Afterward, they received feedback on their response, that also included the correct response in case of an error. All participants in the same condition received the same symbol combinations.

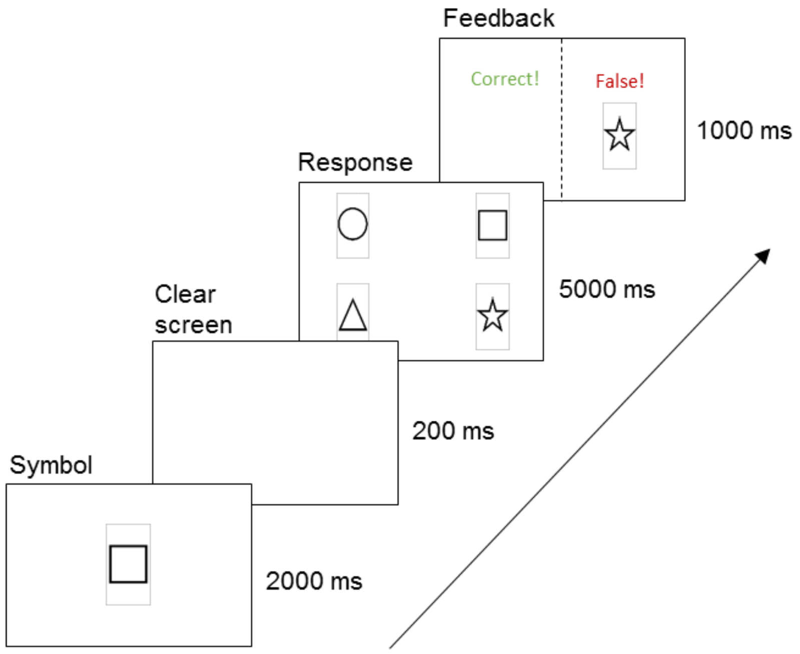

Figure 1. Sample learning trial in the easy condition (adapted from Wirzberger et al., 2017). In the difficult condition, the second symbol was shown accordingly after the first symbol for 2000 ms, separated by an additional clear screen for 200 ms.

All participants further received an interrupting task at the same five defined moments in the task to induce extraneous cognitive load. During the task, the participants were shown a visual search screen with four types of abstract geometric symbols, as shown in *Figure 2*. The symbols were similar to the learning task and for each symbol, seven to nine items were present on the screen. Similarity and an appropriate task duration should ensure the interrupting potential (Gillie & Broadbent, 1989; Monk, Trafton, & Boehm-Davis, 2008; Trick, 2008). The visual search screen was accompanied by the instruction to count two of the displayed types of symbols (targets), for example, all stars and squares. After participants had finished counting or after the maximum time span of 10 s had passed (whatever came first), they received separate response screens for each symbol one after another. Again, the instruction on the target symbol was displayed above the response buttons. The participants had to choose their answer by mouse click from the presented options. Including this task followed the underlying assumption that the prevalence of distracting tasks with competing goals represents a common situational constraint in computer-based learning environments (Gerjets et al., 2003). According to Wickens (2002), a resource-demanding perceptual task can cause substantial interference with a cognitive task that involves the storage and/or transformation processes in working memory.

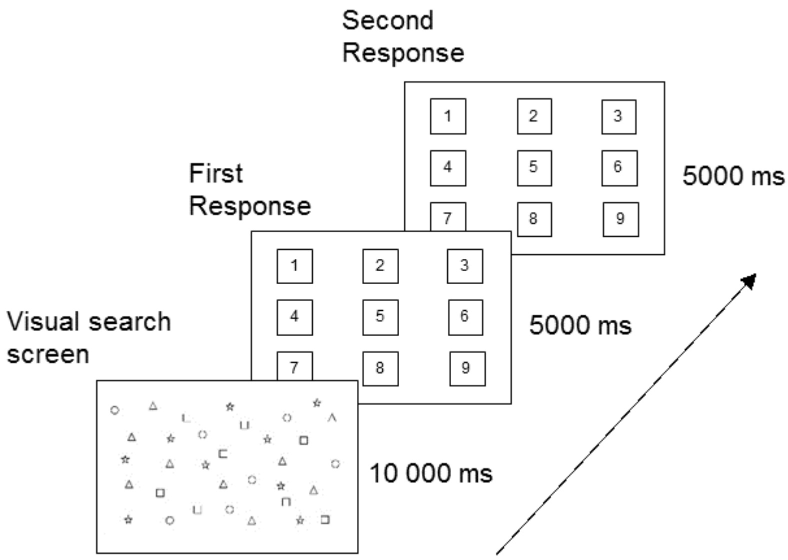

Figure 2. Sample interruption trial (adapted from Wirzberger et al., 2017). Participants received an instruction on the target symbols above the visual search screen (not displayed). Both response screens were separated by a clear screen for 200 ms.

2.2 Model concept

The main idea of our model was to learn the previously introduced stimulus-response combinations in declarative memory, while coping with the induced interruptions. To achieve that, we presented the same 64 learning trials with the same duration and contents (see *Figure 1*) and the same five interruptions (see *Figure 2*) to the model as to the human participants. As we aimed to determine which underlying mechanisms were necessary to explain the human data, we compared four versions of our model that included different activation-related parameters: a model that operated on base-level activation only (Base), a model that additionally used spreading activation (SA-only), a model that additionally used partial matching (PM-only), and a model that additionally used a combination of both spreading activation and partial matching (SA-PM).

2.2.1 Symbol learning task

Each model run starts with an initial setting of the task goal, which was assumed to result from the previously read instruction. Next, each learning trial builds upon three task-related steps, as displayed in *Figure 3*. At first, the presented symbol is visually 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, our 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 encoding the feedback on the given response and, in the case of a false response, an update of the existing intermediate representation. This final information contains both the input and the correct response parts of the symbol combinations, such as 'square – circle – square'. This information is then stored in declarative memory.

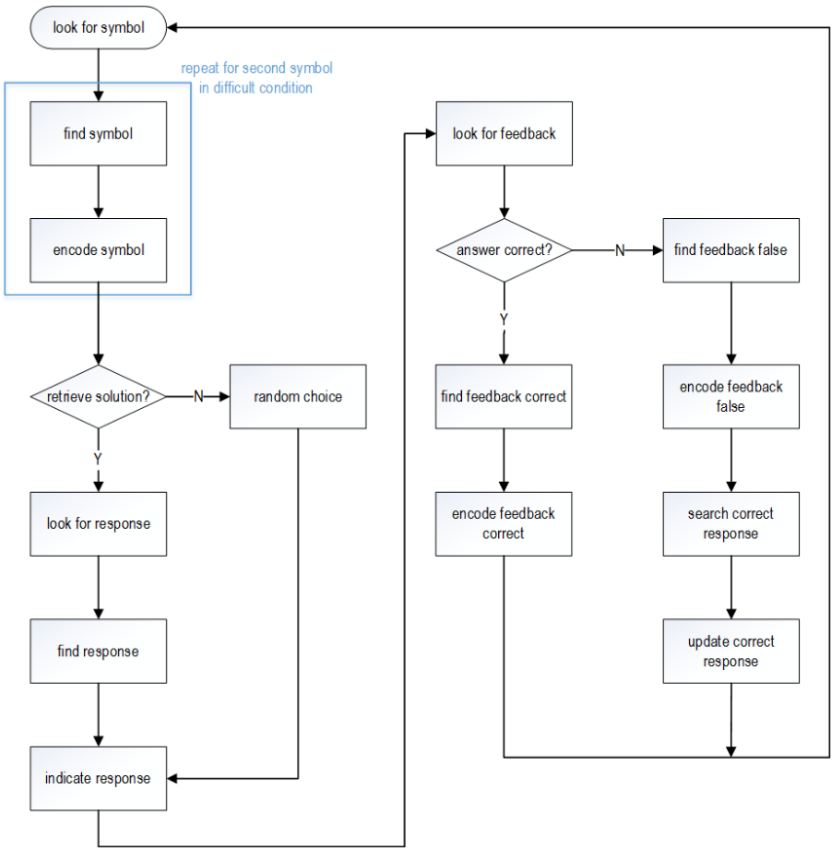

Figure 3. Outline of steps to perform in each the learning trial of the task. The core structure stayed the same across all versions of the model. The inclusion of spreading activation and/or partial matching affected the *retrieval of the solution* after encoding the presented input symbol/s.

2.2.2 Base-level activation, spreading activation, and partial matching

In the first trials, there is no sufficiently matching content or no content at all to retrieve, resulting in slower and less accurate responses. Corresponding to the underlying base-level activation (see *Equation 2*), after being presented the input symbols several times and retrieving related content from declarative memory, our model performs increasingly faster and more accurate due to increasing chunk activation, as we will show in the results section in detail.

Increased element interactivity related to intrinsic cognitive load (Sweller & Chandler, 1994; Sweller, 2010) is reflected in the spreading activation mechanism. Spreading activation arises in our model from the intermediate representation of the shown input symbol/s in the imaginal buffer. In line with the theory, symbol combinations in the difficult condition contain an additional symbol that increases element interactivity due to the related order information. On this account, the spreading activation mechanism particularly affects the difficult condition: Symbol combinations including the same input symbols, such as 'square – circle' and 'circle – square', receive equal activation, independent of the correct symbol order. In consequence, if input order is not taken into account, the model retrieves a wrong input-response association in half of the cases.

Regarding the partial matching mechanism, if the model retrieves a wrong stimulus-response combination, the activation of the mismatching combination is decreased due to the mismatch penalty. In the easy condition, the match or mismatch relates to one symbol slot, while in the difficult condition, there are two slots in the retrieved chunk that either match or mismatch, one for each symbol. In consequence, the related chunk activation gets penalized double in the difficult condition as none of the slots matches. We will see this difference reflected in the results.

2.2.3 Interrupting visual search task

As displayed in *Figure 4*, following a goal change due to the bottom-up triggered saliency of the interrupting task, the model encodes, counts, and responds to the indicated target symbols. The task switch represents the immediate attention to the related screen change and followed the model implemented by Wirzberger and Russwinkel (2015). Tying in with evidence on pre-attentive and attentive processes in the visual module of ACT-R (Nyamsuren & Taatgen, 2013), the second and subsequent steps in the visual search are enhanced by additional information related to stimulus characteristics. In more detail, after finding the first symbol on the screen, it is compared with the target symbol. If there is a match, the stimulus color is added to the next search request. Subsequently, the pre-attentive search for unattended stimuli focuses only on matching colors and only these will receive full attention. Hence, the search process is sped up. In addition, we assumed counting to be a highly trained behavior that occurs almost automatically, thus a simple counting function was applied instead of intermediate retrievals after each counting step (cf. Nijboer et al., 2016). After finishing the counting part, on each of the two response screens the model encodes the requested symbol and retrieves the potential answer.

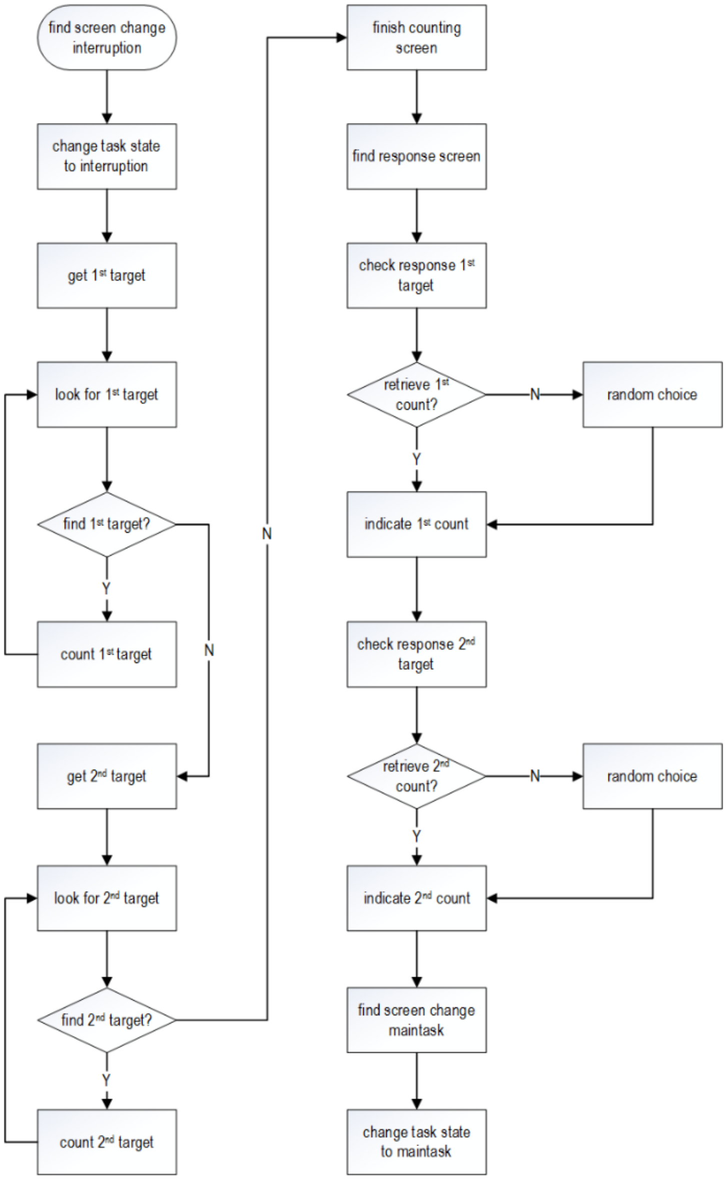

Figure 4. Outline of steps to perform in the interrupting task. The task structure was the same for both conditions and all occurrences of the interruption.

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. 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).

2.3 Parameter fitting

The behavior of ACT-R models can be changed within limits by fitting a number of parameters. Parameter settings in the reported models are outlined in *Table 1* and are in line with the range of reported standard values (e.g., Anderson, Bothell, Lebiere, &, Matessa, 1998). The base-level learning (bll) parameter was fixed for the different models, the other parameters in *Table 1* were adapted to improve the fit to the data. In addition, we assigned the goal chunk for the learning task an initial base-level of 70 (relates to n in *Equation 2*) to account for the fact that participants had received a comprehensive instruction on this task. For the partial matching mechanism, we set the similarities between non-matching symbols to -1 (relates to M in *Equation 1*), and for the spreading activation mechanism, we defined the content of the imaginal buffer as a source for spreading activation upon each retrieval from declarative memory with a total activation of 1 (relates to W_k in *Equation 1*).

Table 1

Parameter settings related to chunk activation and retrieval time

Parameter	Description	Link to Equations	Base model	SA-only model	PM-only model	SA-PM model
mas	*Maximum associative strength*	S in Equation 1	/	4.2	/	1.7
mp	*Mismatch penalty*	P in Equation 1	/	/	1.0	0.401
ans	*Instantaneous noise*	ε in Equation 1	0.6	0.6	0.6	0.2
bll	*Base-level decay*	d in Equation 2	0.5	0.5	0.5	0.5
lf	*Latency factor*	F in Equation 3/4	0.7	1.4	1.4	3.8
rt	*Retrieval threshold*	τ in Equation 4	-1.3	-0.99	-0.85	0.11

3 Results

All inspected model data based on $n = 100$ model runs in each condition since it was not our goal to create an exactly mapping model run for each human participant ($n_{\text{easy}} = 55$, $n_{\text{difficult}} = 58$). Rather, we aimed to obtain robust conclusions from the average model performance. We further aimed at achieving the best possible model fit related to both accuracy and reaction time in either condition. *Table 2* summarizes the goodness-of-fit indices across the compared models. Compared to the base model, the overall pattern demonstrates the benefit of including both the spreading activation and partial matching mechanisms to explain the human data. On closer inspection, the root mean squared scaled deviation (RMSSD) drops by almost one standard error in the SA-PM model. The proportion of explained variance (R^2) stays at the same level but shows a more balanced pattern across measures and conditions.

Table 2

Goodness-of-fit indices across compared models separated by task conditions

	Base model		**SA-only model**		**PM-only model**		**SA-PM model**	
	RMSSD	R^2	RMSSD	R^2	RMSSD	R^2	RMSSD	R^2
Accuracy								
Easy	3.25	0.76	2.02	0.62	5.19	0.18	1.51	0.69
Difficult	4.29	0.56	5.17	0.08	1.87	0.68	2.07	0.57
Reaction time								
Easy	1.59	0.64	1.75	0.41	1.75	0.42	1.67	0.52
Difficult	2.08	0.39	1.98	0.19	2.33	0.50	2.16	0.58
Overall	2.80	0.59	2.73	0.33	2.79	0.45	1.85	0.59

Note. RMSSD = root mean squared scaled deviation, indicates deviation from exact location in units of standard errors; R^2 = proportion of explained variance, indicates fit in relative trend magnitude.

3.1 Human data

To enable the comparison with model performance, human and model performance are displayed in *Figure 5* to *Figure 8.* The grey dots represent the mean across all participants for each of the 64 trials in the learning task, with standard errors indicated by error bars. The red dashed lines mark the first trial after an interruption (i.e., resumption period).

To summarize, the accuracy data indicate that the human participants in the difficult condition learn slower. In the end, both conditions reach an almost comparable level. Overall, the emerging interruption effects are more prevalent in the easy condition.

With respect to reaction times, the human participants speed up with increasing task progress in both conditions. Since our comparisons focus only on correctly solved trials, standard errors decrease over the trials as the number of correct reactions increase. Interruption effects are more distinctive in the easy condition, although they are visible in the difficult condition as well.

3.2 Base model

At first, we created a base model that operates only on base-level activation, without considering spreading activation or partial matching. The obtained results are shown in *Figure 5*. The model data indicate an almost comparable pattern of accuracy in both conditions. On closer look, the model learns slightly slower in the difficult condition, but outperforms the human participants in both conditions.

With respect to reaction time, our model is faster than the human participants at the beginning in both conditions. Over the subsequent trials, it can map the human performance quite well in the easy condition. In the difficult condition, our model reacts constantly slower than the human participants. Interruption effects are visible in both conditions, but especially in the easy condition, they are less distinct compared to the human data.

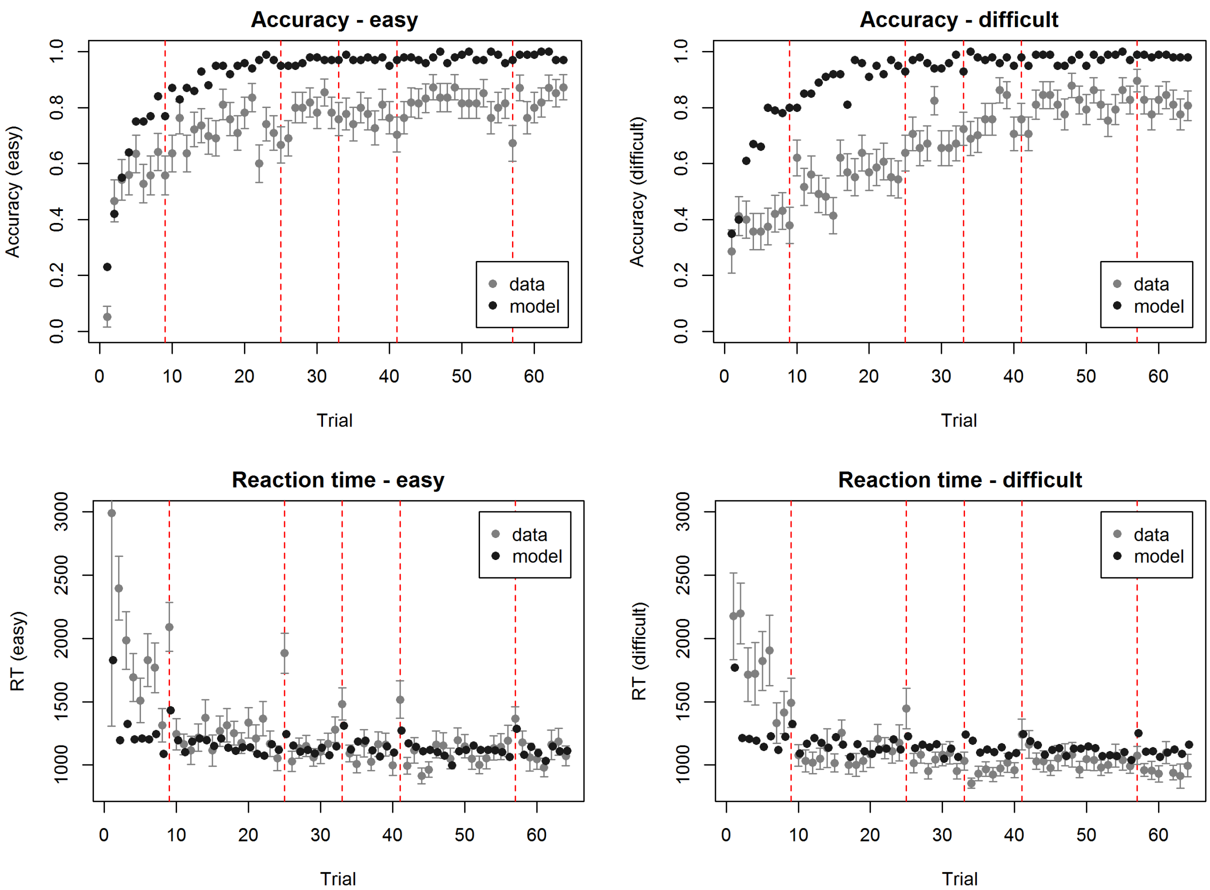

Figure 5. Accuracy (upper part) and reaction times (lower part) for human and model data for the base model in the easy and difficult condition. Reaction times are based only on correct trials. Error bars indicate standard errors for human data. Red dashed lines indicate trials after an interruption.

3.3 SA-only model

Our second model adds spreading activation to the base-level model. *Figure 6* shows the results. It shows quite distinct patterns for accuracy in both conditions. For the easy condition, the model achieves a good mapping particularly in the first half of the trials. The difficult condition reveals that the SA-only model cannot fully map the learning effect but stays at around 50% accuracy until the end (with a 25% chance level).

Approaching reaction time, after the first ten trials our model shows an almost comparable speed of reactions to the human participants in both conditions. The mapping is slightly better in the easy condition. In general, the interruption effects are prevalent in both conditions but less visible compared to the human data.

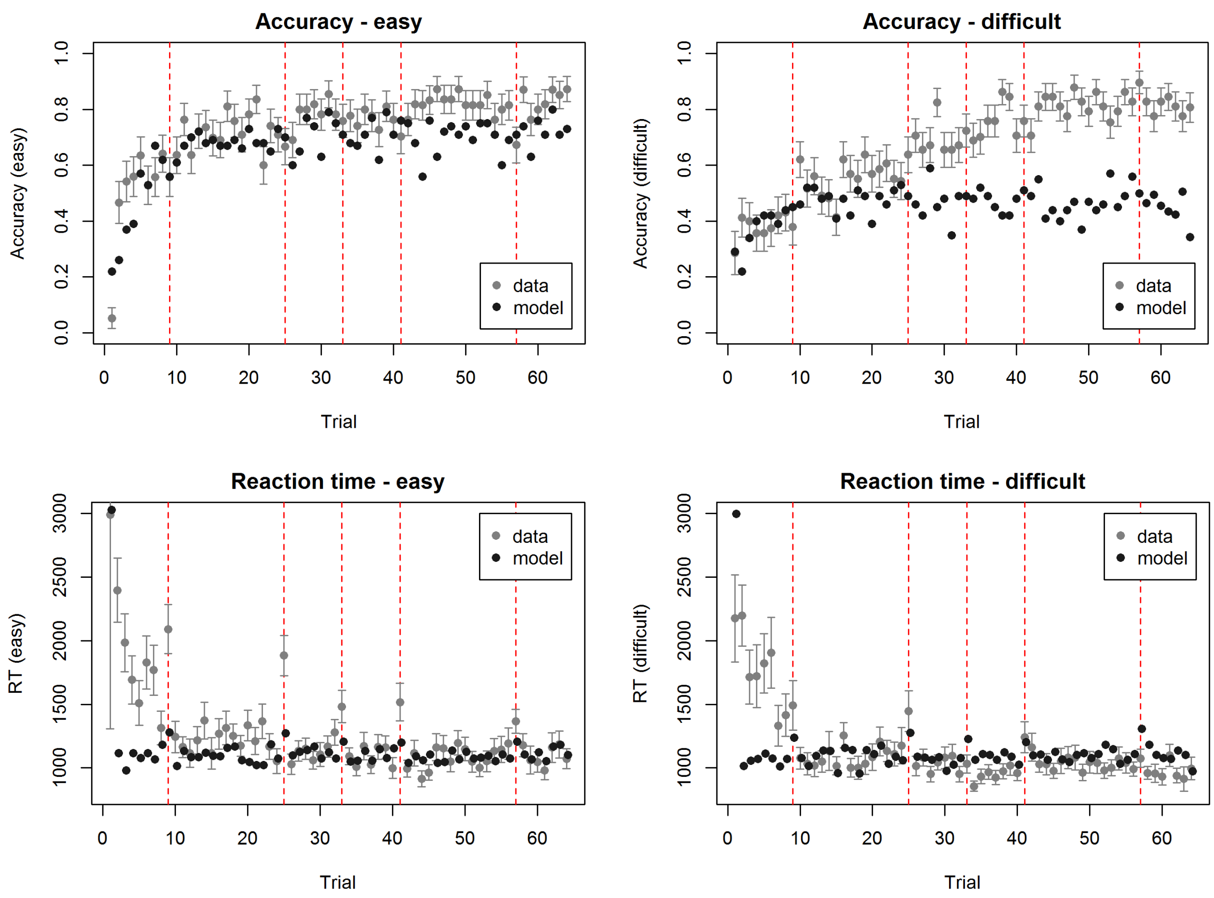

Figure 6. Accuracy (upper part) and reaction times (lower part) for human and model data for the SA-only model in the easy and difficult condition. Reaction times are based only on correct trials. Error bars indicate standard errors for human data. Red dashed lines indicate trials after an interruption.

Compared to our base model, the additional activation that spreads from the content in the imaginal buffer increases the activation of chunks with matching slots. As we had to remove the ordered retrieval to map this effect, in the difficult condition, this applies to chunks regardless of the symbol order. For example, both the chunks 'square – circle' and 'circle – square' receive equal activation, which results in retrieving the wrong symbol combination in 50% of the cases. This is reflected in the decreased level of accuracy in the difficult condition compared to the base model.

3.4 PM-only model

Our third model adds partial matching instead of spreading activation. The obtained data are shown in *Figure 7*. Again, we observe distinct patterns for accuracy in both conditions. In this case, the model achieves a better fit in the difficult condition, particularly after the first 25 trials, and maps the learning effect in the human data. For the easy condition, following an initial increase, the accuracy stays around 50%. This condition-related pattern reverses the pattern observed for the SA-only model.

Inspecting the visual impression for reaction time reveals an increased mapping in the difficult condition, although the model reacts slower than the human participants after the first ten trials. In the easy condition, the model reacts faster than the human participants in the first ten trials but resides around the same level in subsequent trials. Compared to the human data, the interruption effects are less visible in either condition.

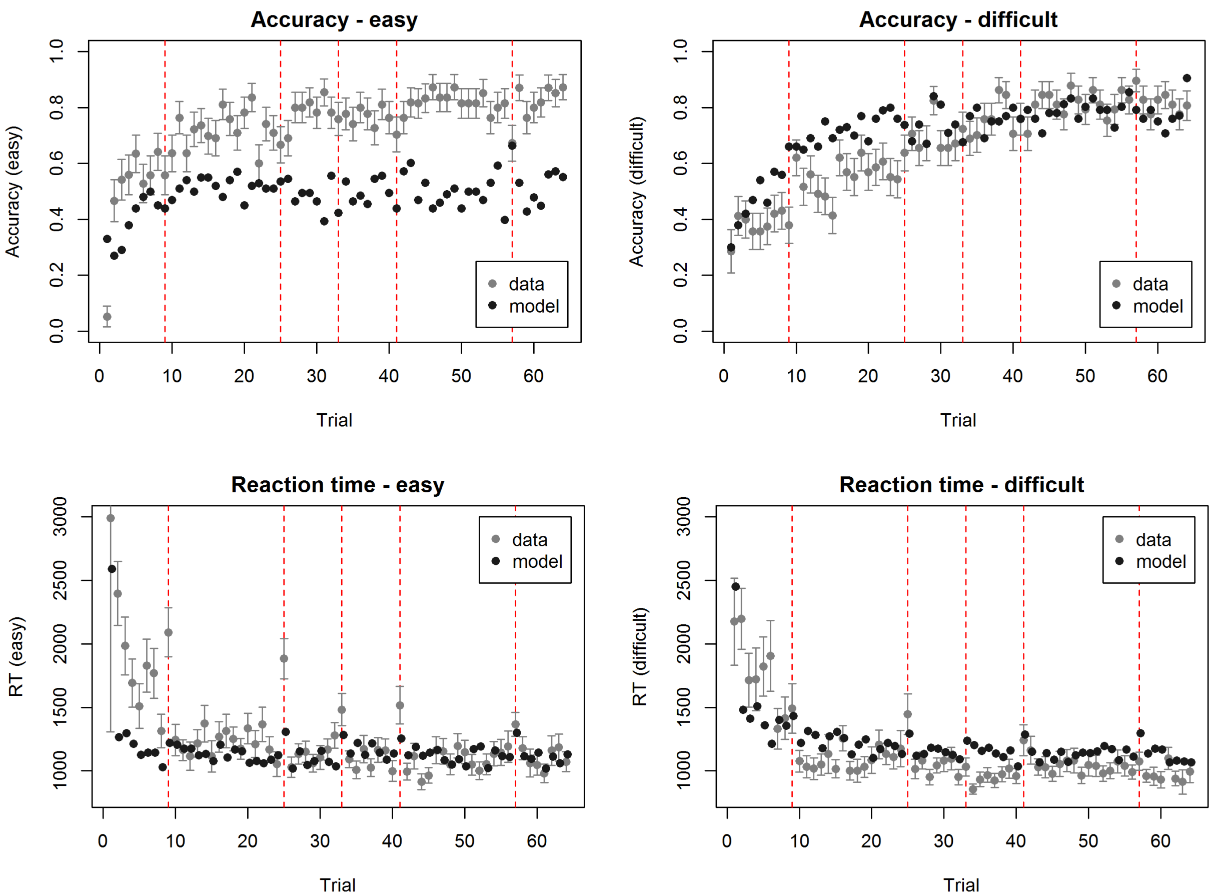

Figure 7. Accuracy (upper part) and reaction times (lower part) for human and model data for the PM-only model in the easy and difficult condition. Reaction times are based only on correct trials. Error bars indicate standard errors for human data. Red dashed lines indicate trials after an interruption.

Referring back to our base model, the PM-only model also accounts for errors of commission, as wrong information can be retrieved from memory as well. When retrieving wrong information, the resulting mismatch gets penalized and reduces chunk activation. This also relates to longer chunk retrieval times, as observable in the reaction time for the difficult condition. In contrast to the easy condition, chunks in the difficult condition contain additional information that can potentially result in a mismatch, thus retrieving a wrong chunk in this condition is penalized double. This results in the displayed pattern of a better fit in the difficult condition, but only around 50% correct reactions in the easy condition with the same parameter setting.

3.5 SA-PM model

To overcome the shortcomings of both the SA-only and PM-only model, our final model combines both spreading activation and partial matching. *Figure 8* shows the average model performance across the trials for both accuracy and reaction time.

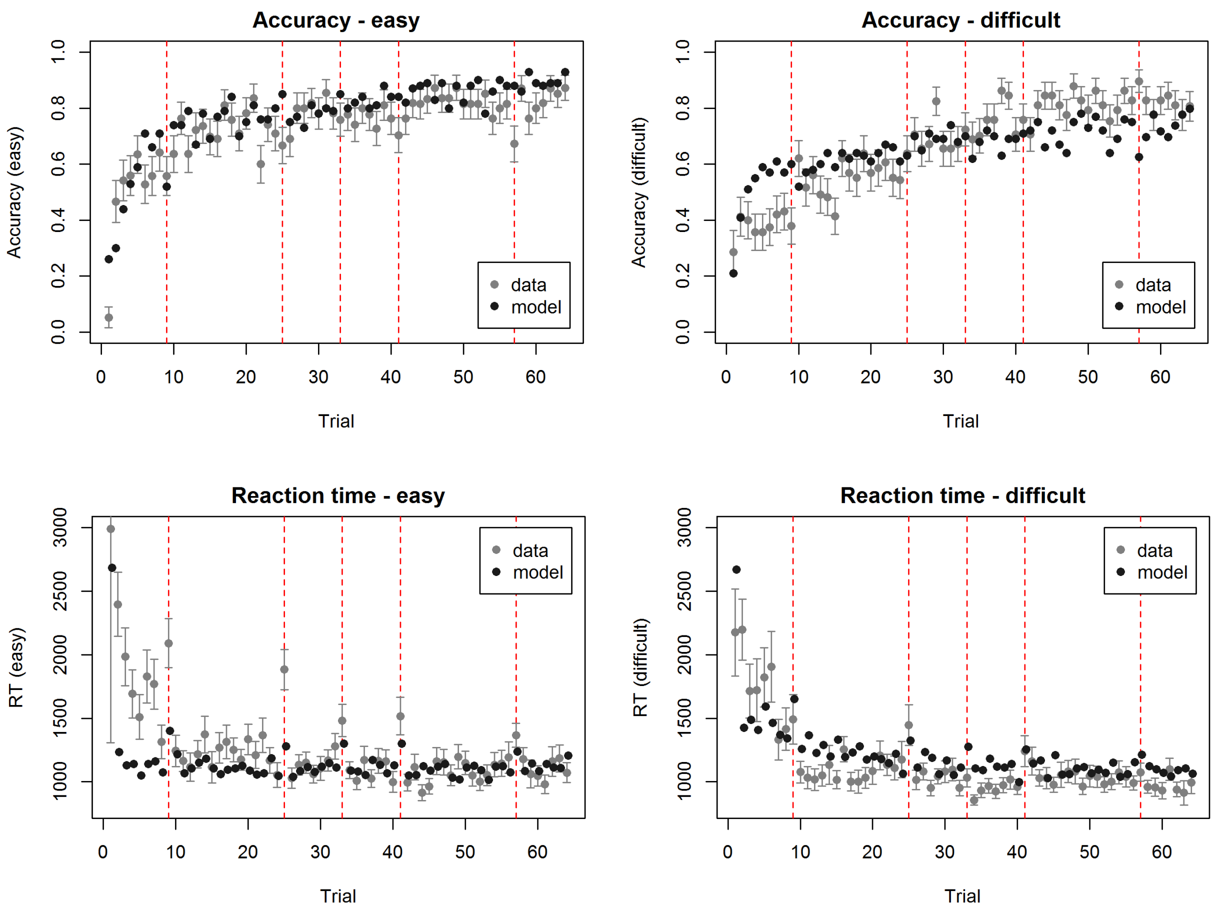

Figure 8. Accuracy (upper part) and reaction times (lower part) for human and model data for the SA-PM model in the easy and difficult condition. Reaction times are based only on correct trials. Error bars indicate standard errors for human data. Red dashed lines indicate trials after an interruption.

For accuracy, the data indicate that the SA-PM model can map the development in human performance quite well in the easy condition, although it achieves a higher performance in the end and shows a subtler reflection of interruption effects. In the difficult condition, our model learns slower compared to the easy condition, but still faster than the human participants.

For reaction times, the visual inspection indicates that the model data can map the initial decrease in the difficult condition. In general, it performs slightly slower than human participants during most of the trials. In the easy condition, our model shows a subtler decrease and faster reactions in the beginning, although, for later trials, its mapping fits quite well. Interruption effects are observable in both conditions.

3.6 Simulated fMRI data

For the SA-PM model, we further generated predefined ROI-predictions (Borst & Anderson, 2017) to inspect underlying cognitive resource demands in more detail and link our model results to predictions from Whelan (2007). The learners' motivation was not addressed explicitly in the underlying task design, which according to Whelan (2007) determines brain areas related to germane cognitive load. Hence, our inspection focused only on the declarative and imaginal module, which are connected to working memory (Nijboer et al., 2016). We considered them to reflect effects of task complexity as indicator of intrinsic cognitive load. Related to resumption trials, they further reflect effects of the embedded interruptions, which should induce extraneous cognitive load.

The ROI-predictions based on the previously outlined mapping of activity in ACT-R modules on defined brain regions (see Anderson, 2007; Borst, Nijboer et al., 2015). 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 neural activity. The activity of each inspected module is represented as 0-1 demand function and convolved with a hemodynamic response function (e.g., Anderson, 2007; Borst, Taatgen, Stocco, & Van Rijn, 2010; Stocco & Anderson, 2008). For example, lower levels of chunk activation in declarative memory would result in longer retrieval times, which would then be reflected in increased activity in the dorsolateral prefrontal cortex. They would be observable by higher peaks in the simulated fMRI signal.

Changes in module activity that originate in learning processes are displayed in *Figure 9.* The curves indicate a decrease in cognitive activity in intermediate task stages (displayed in red) and later task stages (displayed in black) compared to early task stages (displayed in blue) in both conditions in the declarative module (prefrontal cortex). As both input symbols are presented one after another, the later retrieval onset in the difficult condition relates to the encoding of the second input symbol. Furthermore, the difficult condition shows a higher level of activity across all stages, with a particularly distinctive peak across early task stages. Activity in the imaginal module (posterior parietal cortex) exerts a longer duration and shows an increased level in the difficult condition across all stages. On closer look, we can relate this pattern to the update of the intermediate representation of a symbol combination after the second symbol has been shown.

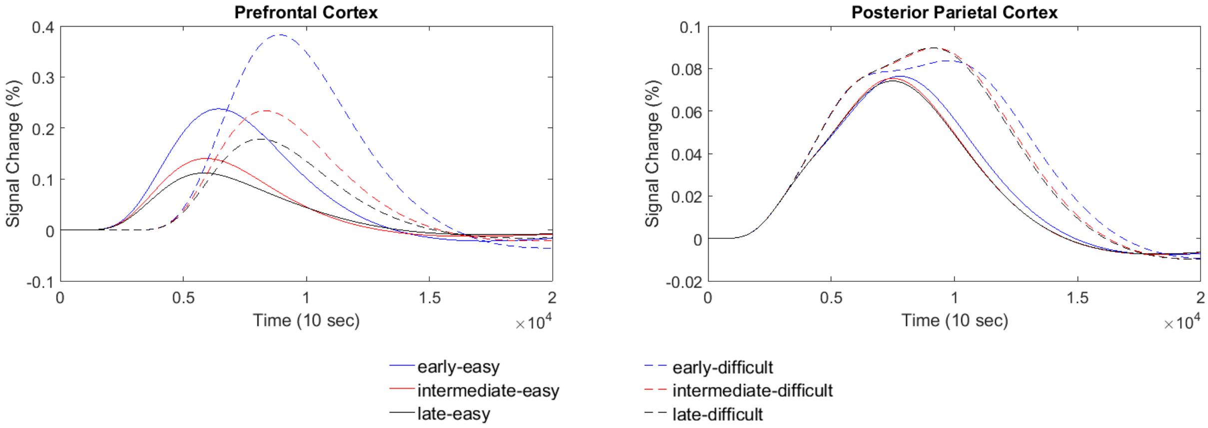

Figure 9. Predicted activity in the prefrontal cortex (declarative module) and posterior parietal cortex (imaginal module) in early, intermediate, and late temporal stages of the symbol learning task (excluding resumption trials). Solid lines represent the easy condition, dashed lines represent the difficult 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).

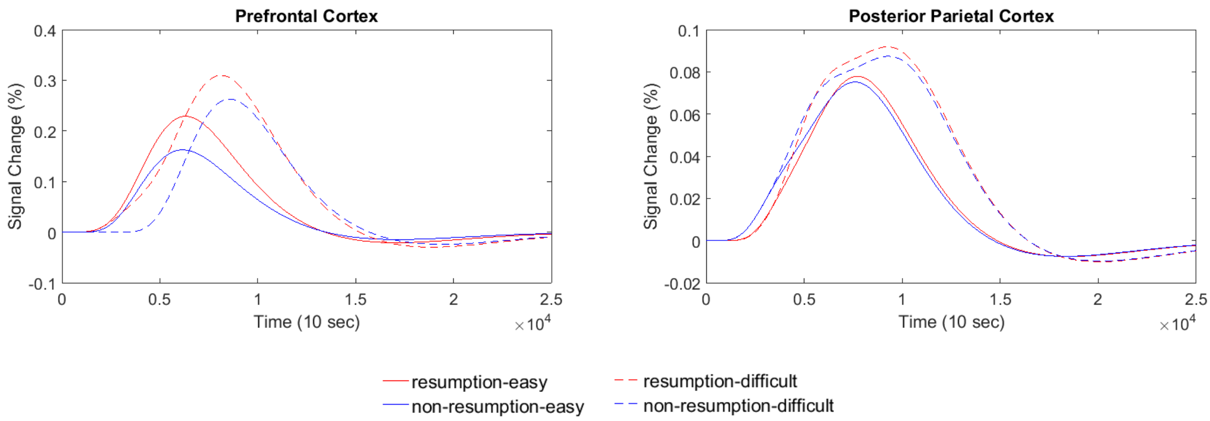

Figure 10. Predicted activity in the prefrontal cortex (declarative module) and posterior parietal cortex (imaginal module) in resumption and non-resumption trials of the learning task. Solid lines represent the easy condition, dashed lines represent the difficult condition. Red lines represent resumption trials (n = 5), and black lines represent the non-resumption trials (n = 59).

Comparisons shown in *Figure 10* include a separate visualization of the resumption phase, defined as the first trial that immediately follows the interrupting task. For the declarative module (prefrontal cortex), a higher activity across resumption trials (displayed in red) compared to the remainder of trials in the learning task (displayed in blue) results for both conditions. On closer look, this pattern reflects the activation decay in the symbol combination chunks during the interruption. In addition, we can observe an earlier onset of retrieval activity in the resumption trials in the difficult condition, which relates to the process of restoring the previous task goal after being interrupted. The imaginal module (posterior parietal cortex) displays a slightly later onset and slightly higher activity across resumption trials, which also reflect the outlined task switching costs.

4 Discussion

The current model explores cognitive processes and mechanisms that underlie changes in performance due to the inserted interruptions and task-related learning processes. Comparing model performance across both conditions on a behavioral level, the results obtained from our final SA-PM model indicate a reasonable fit in terms of reaction times and accuracy. This model can map both the learning-related increase in performance and the decrease in performance due to experiencing an interruption. To model the learning effects, we take into account that information can be retrieved increasingly faster and more accurate the more often it is used (base-level activation). We further consider that people retrieve the wrong symbol combination sometimes (partial matching), particularly at the beginning of the learning task. Theoretically related to the concept of element interactivity, we also build on the fact that activation is distributed across related information elements (spreading activation). We approach the interruption effects by including the need to restore the task representation after being interrupted. This delays the model's response to the subsequent learning trial.

On a neural level, the smaller level of invested cognitive activity with increasing task progress emphasizes the prevalence of learning effects in both conditions. Existing content in the declarative memory receives increasingly higher activation and can be retrieved faster and more accurately. In the difficult condition, invested declarative resources are constantly higher across all stages, which by closer inspection relates to the increased influence of partial matching. The partial matching mechanism, which penalizes chunk activation and extends retrieval times, exerts higher influence as the chunks contain more potentially mismatching information. For interruption effects, increased levels of activity in the resumption period arise from restoring the previous task representation due to the task switch. It further relates to the occurrence of an activation decay in chunks related to the acquired symbol combinations (Altmann & Trafton, 2002).

4.1 Implications for CLT

As obvious from the ROI-analysis, our final model needs to invest a higher amount of cognitive resources for each retrieval in the early task stage due to the lack of suitable chunks and lower levels of chunk activation. The less effortful accessibility of established knowledge structures (schemata) with increasing progress in the learning task corresponds to the process-driven reconceptualization of the Cognitive Load Theory (De Jong, 2010; Seufert, 2018; Wirzberger et al., 2018).

The application of the spreading activation mechanism maps the theoretically introduced concept of element interactivity (Chen et al., 2017; Ngu et al., 2018; Sweller & Chandler, 1994; Sweller, 2010). It offers the potential for deconstructing and formalizing effects of increased task complexity on a cognitive level. For intrinsic cognitive load, Whelan (2007) postulated a connection to the prefrontal cortex. Activity in this brain area corresponds well to the observed increased activity in the declarative module.

In addition, Whelan (2007) discussed a relation between extraneous cognitive load and the posterior parietal cortex. In the current task design, the embedded interruption task should induce the same amount of extraneous cognitive load in both task conditions. However, due to the correspondence in the underlying brain region, we might connect the imaginal module to extraneous cognitive load. The observed increase in activity in this module in the difficult condition suggests that this condition also induced more extraneous cognitive load. It might result from presenting the symbols one after another, which raises the necessity of integrating the perceptual representations of both input symbols. In line with the split-attention effect (Schroeder & Cenkci, 2018), the arising demand to distribute the attention across task-relevant sources of information impairs learning. This effect occurs not only related to distributed spatial locations but also related to a temporally distributed presentation, which is described in the temporal contiguity effect (Ginns, 2006).

4.2 Future directions

For the difficult condition, the performance that we observed in our models 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. This directly maps the approach of task execution in the easy condition. Explaining a potential 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 a potential for future work, first by broadening the scope of the existing model and second by validating this model with new tasks.

5 Conclusion

Taken together, the obtained results emphasize the importance of considering activation-related dynamics when approaching changes in cognitive resource demands underneath performance in learning situations. Our outlined cognitive modeling approach inspects the influence of both internal and external factors in these contexts. On the one hand, we modeled learning-related change over the task that relates to increasingly faster and less erroneous information retrieval. On the other hand, we included the requirement to recreate the previous task representation after being interrupted. Since our approach extends beyond human experiments and model-based behavior on a neural level, it provides a more detailed understanding, which is crucial to exploring related patterns of cognitive resource investment. It can be taken as promising step for developing adequate instructional support and minimizing harmful effects in learning scenarios.

6 Acknowledgments

The authors gratefully acknowledge funding by the German Research Foundation (DFG), GRK 1780/1 and the Saxon State Ministry for Higher Education, Research and the Arts (SMWK).

7 References

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

Anderson, J. R. (1982). Acquisition of cognitive skill. *Psychological Review, 89*, 369–406. https://doi.org/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. https://doi.org/10.1006/jmla.1997.2553

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

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

Banich, M,, Milham, M., Atchley, R. A,, Cohen, N. A., Webb, A., Wszalek, T., et al. (2000). Prefrontal regions play a predominant role in imposing an attentional 'set': Evidence from fMRI. *Cognitive Brain Research, 10,* 1-9. https://doi.org/10.1016/S0926-6410(00)00015-X

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. https://doi.org/10.1016/j.learninstruc.2009.02.024

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. https://doi.org/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. https://doi.org/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. https://doi.org/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. https://doi.org/10.1145/2702123.2702156

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. https://doi.org/10.1023/B:TRUC.0000021812.96911.c5

Byrne, M. D., & Anderson, J. R. (2001). Serial modules in parallel: The psychological refractory period and perfect time-sharing. *Psychological Review*, *108*, 847–869. https://doi.org/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, 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. https://doi.org/10.1007/s10648-016-9359-1

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *The Behavioral and Brain Sciences, 24*, 87–114. https://doi.org/10.1017/S0140525X01003922

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

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

Fechner, H. B., Pachur, T., Schooler, L. J., Mehlhorn, K., Battal, C., Volz, K. G., & Borst, J. P. (2016). Strategies for memory-based decision making: Modeling behavioral and neural signatures within a cognitive architecture. *Cognition, 157*, 77–99. https://doi.org/10.1016/j.cognition.2016.08.011

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. https://doi.org/10.1146/annurev.ps.34.020183.001401

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. https://doi.org/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. https://doi.org/10.1016/j.learninstruc.2006.10.001

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).

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

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. https://doi.org/10.1007/s10648-015-9352-0

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

Meredith, M. A. (2002). On the neuronal basis for multisensory convergence: A brief overview. *Cognitive Brain Research, 14.* 31-40. https://doi.org/10.1016/S0926-6410(02)00059-9

Miller, E. K., & Cohen. J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience*, *24*, 167-202. https://doi.org/10.1146/annurev.neuro.24.1.167

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. https://doi.org/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. https://doi.org/10.1007/s10648-016-9397-8

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. https://doi.org/10.1016/j.cogpsych.2016.01.003

Nyamsuren, E., & Taatgen, N. A. (2013). Pre-attentive and attentive vision module. *Cognitive Systems Research, 24*, 62–71. https://doi.org/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. https://doi.org/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. https://doi.org/10.1207/S15326985EP3801_8

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

Scharinger, C., Kammerer, Y., & Gerjets, P. (2015). Pupil dilation and EEG alpha frequency band power reveal load on executive functions for link-selection processes during text reading. *PloS one*, *10*(6), e0130608. https://doi.org/10.1371/journal.pone.0130608

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. https://doi.org/10.1007/s10648-018-9435-9

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

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

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

Sweller, J. (2010). Element interactivity and intrinsic, extraneous, and germane cognitive load. *Educational Psychology Review, 22*, 123–138. https://doi.org/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. https://doi.org/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. https://doi.org/1040-726X/98/0900-0251S15.00/0

Taylor, S., Welsh, R., Wager, T., Phan, K., Fitzgerald, K. D., & Gehring, W. J. (2004). A functional neuroimaging study of motivation and executive function. *Neurolmage, 21*, 1045-1054. https://doi.org/10.1016/j.neuroimage.2003.10.032

Tenison, C., & Anderson, J. R. (2015). Modeling the distinct phases of skill acquisition. *Journal of Experimental Psychology: Learning Memory and Cognition, 42*, 749–767. https://doi.org/10.1037/xlm0000204

Tenison, C., Fincham, J. M., & Anderson, J. R. (2016). Phases of learning: How skill acquisition impacts cognitive processing. *Cognitive Psychology, 87*, 1–28. https://doi.org/10.1016/j.cogpsych.2016.03.001

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. https://doi.org/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. https://doi.org/10.3758/PP.70.5.743

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

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. https://doi.org/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. https://doi.org/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,* 3:46. https://doi.org/10.1186/s41235-018-0138-z

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

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.