Research article

Maria Wirzberger*, Nele Russwinkel

Modeling Interruption and Resumption in a Smartphone Task: An ACT-R Approach

DOI 10.1515/icom-2015-0033

Abstract: This research aims to inspect human cognition when being interrupted while performing a smartphone task with varying levels of mental demand. Due to its benefits especially in the early stages of interface development, a cognitive modeling approach is used. It applies the cognitive architecture ACT-R to shed light on task-related cognitive processing. The inspected task setting involves a shopping scenario, manipulating interruption via product advertisements and mental demands by the respective number of people shopping is done for. Model predictions are validated through a corresponding experimental setting with 62 human participants. Comparing model and human data in a defined set of performance-related parameters displays mixed results that indicate an acceptable fit – at least in some cases. Potential explanations for the observed differences are discussed at the end.

Keywords: ACT-R, Cognitive Modeling, Interruption, Mobile Interaction

1 Introduction

According to statistical information, about 45 million people in Germany [18] and nearly two billion people worldwide [19] use a smartphone. Despite their great convenience in daily use, interruption is a frequently occurring phenomenon when interacting with such devices. Moreover, due to the mobile design, smartphone use is typically embedded into various situational contexts. They might put additional limits on the available cognitive capacity if a user is interrupted during periods with already increased mental demands. To avoid or at least moderate the resulting impairment, considering task-related cognition already when developing and designing such interfaces might be of great value. Contributing to this intent, the existing research inspects how a certain kind of interruption affects human cognition while performing a smartphone task. Due to the mentioned use scenario, various levels of mental demand are taken into account.

1.1 Matter of Interruption

According to [7], interruption as a certain kind of human experience is usually neither planned nor expected and represents a cognitive break with the task performed up to that time. It can be induced by internal or external sources, resides within a given situational context and indicates a delay in finishing the previous activity. The main goal after facing an interruption comprises successfully returning the mental resources to the actual focus of attention, commonly denoted as resumption. Interruptions are known to impair the main task performance due to a set of disruptive aspects, including those which have a great similarity to the main task [10], an immediate occurrence [20] or the lack of opportunity to refuse or delay the interruption [16].

Cognitive processes in the face of external interruptions are described by Altmann, Trafton and colleagues [20]. They postulate a time course model of interruption and resumption: After starting a main task and performing it for some time, an alert appears announcing the interruption before it actually occurs. The time span between the alert and the upcoming interruption is called *interruption lag* and is supposed to prepare for effectively returning to the main task. The *resumption lag* as the interval between ending the interruption and successfully resuming the main task comprises an authentic measure for the extent of disruptiveness. In theoretical accounts, this model is based on the memory-for-goals theory described by [1], assuming an activation decay of the cognitive representation of the main task in aid

***Corresponding author: Maria Wirzberger,** DFG Research Training Group "CrossWorlds", TU Chemnitz, Chemnitz, Germany, e-mail: maria.wirzberger@phil.tu-chemnitz.de
Nele Russwinkel: Cognitive Modeling in dynamic Human-Machine Systems, TU Berlin, Berlin, Germany

of the cognitive representation of the interruption. Such decay might be reduced by either rehearsing information related to the main task or linking environmental cues to certain aspects of the main task.

1.2 Resource Limitation

When dealing with interruptions, one important restricted resource in information processing exists due to working memory limitations, both in terms of duration and capacity [21]. The first aspect refers to the fact that information in working memory decays after a certain time. In order to extend such periods of availability, people can rehearse relevant information. In contrast, the matter of capacity indicates that just a defined amount of information can be held active at the same time. According to [14], it should be between five and nine items, although more recent research proposes smaller numbers [8, 9]. Again, rehearsing information depicts a way to increase this span. In general, when performing a memory-related task, memory load has to be maintained by working memory [4]. Increasing such load might affect task performance and result in difficulties to retrieve necessary information.

2 Methods

2.1 Cognitive Modeling Approach

In the field of human-computer interaction, applying cognitive user models is gaining more and more attention. Above all, this method aims to understand task-related human behavior by inspecting underlying cognitive processes. Especially cognitive architectures have proven of value within this context since they provide a theoretical framework to explain basic and constant mechanisms of human cognition behind a variety of tasks [11]. Moreover, they offer a computational platform for model execution and in this vein give the opportunity to derive predictions on task-related behavior direct.

So far, the Lisp-based cognitive architecture ACT-R (Adaptive Control of Thought – Rational) developed by John R. Anderson and colleagues [3] has been used actively within a vibrant and growing research community to address a variety of basic and applied subjects. It comprises the assumption of different modules that serve as the foundation of any human behavior, occupy defined responsibilities and interact in certain ways to create cognitive processing [2]. *Figure 1* gives an overview of the standard modules contained in the current ACT-R distribution. These modules have been validated by human experimental data and hold a vested biological background as well since fMRI studies [2, 5] indicate the association of each module with a brain region relevant to the respective cognizance. Each module holds a buffer, serving as an interface to enable communication among the modules. Information is processed within the outlined structure via chunks, i. e. small units encoding relevant elements of knowledge affiliated with a certain category *(chunk-type)* and containing specific attributes *(slots)*.

Although processes in different modules can be executed in parallel, each buffer can hold just one chunk at the same time, representing the existing limitations in information processing resources. Interaction between modules happens by means of production rules, consisting of a condition and an action part and depicting the main duty of the procedural module. It scans the buffer's contents and, based on the resulting pattern, selects a suitable production rule that initiates the related action. In case more than one production rule fits, subsymbolic mechanisms apply and a cost-benefit function *(utility)* decides which production rule is selected. If and how fast a chunk can be retrieved from declarative memory depends on another subsymbolic mechanism called *activation*. It reflects the availability of information and is determined by the respective context and history of use.

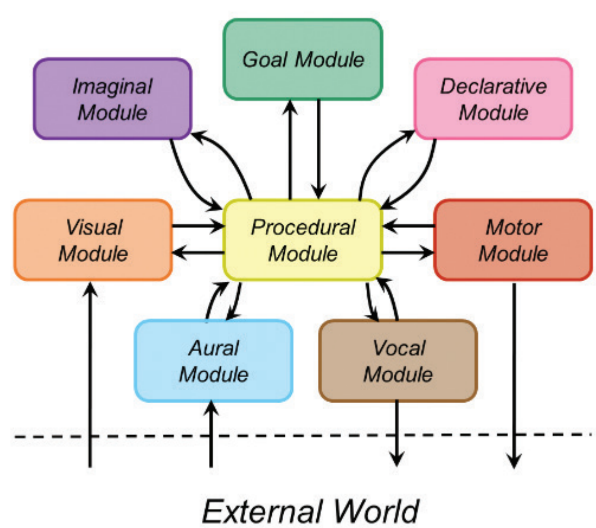

Figure 1: Overview of modules contained in ACT-R 6.0. Adapted from [5] and [2].

2.2 Inspected Task Setting

To examine cognitive processes in terms of an interrupted task, a shopping scenario was employed within this project. It operated on a simple smartphone shopping list application [15], illustrated in *Figure 2*, and consisted of encoding, remembering, searching for and selecting a set of 12 predefined products, equally distributed within three runs. Products appeared in a fixed sequence listed on the screen at the beginning of each run. At the end of the task, all products still remembered from the whole selection part had to be recalled.

During two of the three runs, interruptions in terms of product advertisements occurred with varying frequency. They were always related to the previously selected shop, announced a special offer, and were triggered by the successful selection of a certain amount of products within the respective run. In order to end the interruption and return to the product selection, a decision for or against the offered product was forced. In conditions with enhanced mental demands, shopping was done for three different people. In consequence, additional information about whom the product should be bought for had to be remembered and recalled throughout the task as well.

Task-related cognition was assessed in terms of certain behavioral performance parameters. *Product selection time* was computed as the time difference between successfully selecting a product and returning to the related shop menu. For the *number of selected products*, correctly selected products per run were counted. *Resumption time* comprised the difference between the offset of the interruption and the transition back to the shop menu. The *final recall performance* was assessed by summing up correctly recalled products after finishing the selection part.

2.3 Creating the Model

Based on the outlined assumptions from relevant literature on similarity [10], immediacy [20] the forced occurrence of interruptions [16] and the role of working memory demands [21, 4], an ACT-R model was established. Key features of the described shopping list application were implemented in the ACT-R experimental GUI in a simplified way. In contrast to the model of [20], no alert was included, resulting in the absence of an interruption lag. Without this time span, there was no opportunity to explicitly create environmental cues or apply rehearsal before turning to the interrupting task. In consequence, naturally existing cues from memory or environment had to be used for resumption, e. g. the memorized selection content or visible selection marks. The modeled task always started with reading the written product list and remembering its content. Products were then selected within a navigation and selection procedure spanning the three menus illustrated in *Figure 2*.

In case an interruption occurred, its message was read and a random decision for or against the offered product was made. By performing either a memory retrieval of the previously selected product or looking for a relevant environmental cue, i. e. a selection mark linked to the lastly selected product, resumption started. Next a retrieval of the selection history, i. e. products already selected within the run and previously stored in the intermediate memory, was attempted. This is reminiscent of the memory-for-problem-states theory recently reported by [6]. Afterwards, the next product was then searched for or the run ended when the reconstruction of the previous selection failed. A run ended through either successfully selecting all four products or failing

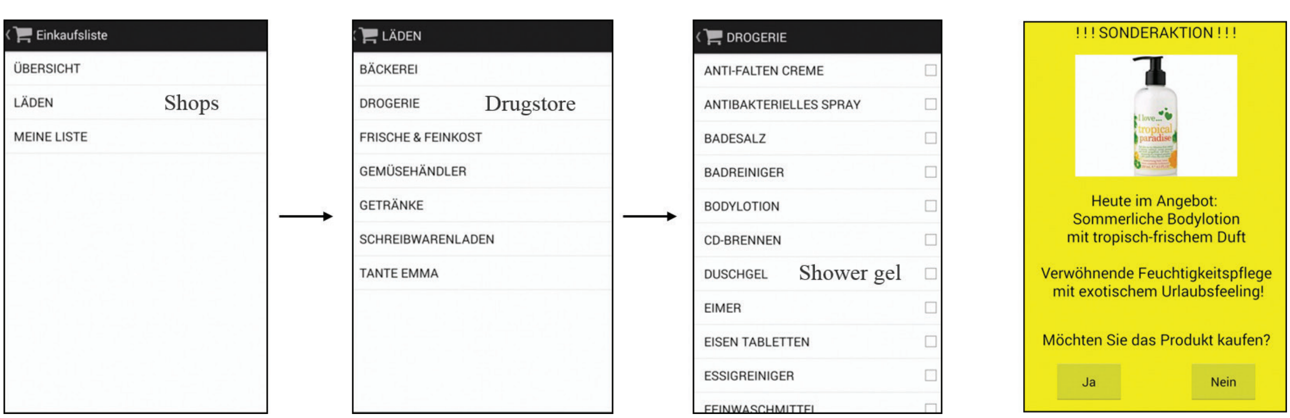

Figure 2: Main menu, store menu, example product menu for drugstore and example product advertisement for “body lotion” within the shopping list application.

to retrieve the next product. After completing the third and last run, the final product recall occurred. Under conditions of enhanced memory demands, the product-related person had to be remembered throughout the task and recalled at the end as well.

2.4 Experimental Validation

To assess the adequacy of the cognitive user model in terms of human behavior, an experimental validation was conducted. 62 participants aged 20 to 49 years ($M = 28.53$, $SD = 7.16$ years) performed the described task using a LG Google Nexus 4 smartphone with a screen size of 4.7", a display resolution of 1280 x 768 pxl, a pixel density of 319 ppi and Android 4.4.2 (KitKat) serving as operating system.

A shopping scenario was created to foster the participant's involvement in the task. It asked them to imagine being a virtual character that does shopping with aid of the app during a usual day in life. Under conditions of enhanced memory demands two additional characters were introduced, an old neighbor and a sick friend. Participants were provided with information on attributes and habits of both people as well as their relationship to the main character. Since the latter was distinguished as caring and social person, shopping assistance for needy fellows was regarded as natural behavior and suitable for enhancing the external validity of the task.

3 Results

3.1 Model Behavior

Looking at the resulting model behavior, when there were interruptions, fewer recalled products were shown and product selection took slightly longer. As can be seen in *Table 1*, such effects showed up especially with increasing frequency of interruption. Moreover, without enhanced memory demands, the model performed better across all performance-related parameters.

3.2 Model Comparison

Based on the inspected model behavior, a decline in task performance with increasing interruption and mental demands was expected in human data as well. Descriptive values are displayed in *Table 2*, statistical effects concerning differences in terms of interruption and mental demands were inspected by computing either analyses of variance (ANOVAs) in the case of product selection time, selected products and resumption time or a χ^2-test in the case of final recall performance. In summary, none of the ANOVAs achieved significant results, neither for interruption nor mental demands. In contrast, a significant difference between high and low mental demands in the case of final recall showed up, $\chi^2(15, N = 62) = 25.397$, $p = .045$, supporting the assumption of worse final recall performance when mental demands are enhanced.

Table 1: Descriptive values of the model behavior regarding the inspected performance parameters.

	Mental demands	No ad		Low ad		High ad		Overall	
		M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Product selection time (in sec)	H	7.42	0.64	7.31	0.73	7.58	0.77	7.41	0.28
	L	6.09	0.64	6.09	0.69	6.10	0.72	6.05	0.20
	–	6.76	0.93	6.70	0.94	6.84	1.05	6.73	0.73
Selected products (sum)	H	3.67	0.48	3.03	0.18	3.00	0.00	9.70	0.47
	L	4.00	0.00	3.77	0.43	3.30	0.47	11.07	0.52
	–	3.83	0.38	3.40	0.49	3.15	0.36	10.38	0.85
Resumption time (in sec)	H			3.08	0.68	3.95	0.36	3.66	0.29
	L			2.65	0.27	2.72	0.26	2.69	0.20
	–			2.86	0.56	3.33	0.69	3.18	0.54
Final recall (in %)	H							60.56	9.01
	L							84.17	6.32

Note. H: mental demands enhanced (data based on $n = 30$ model runs), L: mental demands not enhanced (data based on $n = 30$ model runs), – : no separation by mental demands (data based on $N = 60$ model runs).

Table 2: Descriptive values of the human data regarding the inspected performance parameters.

	Mental demands	No ad			Low ad			High ad			Overall		
		N	*M*	*SD*	*N*	*M*	*SD*	*N*	*M*	*SD*	*N*	*M*	*SD*
Product selection time (in sec)	H	31	9.32	5.72	31	9.58	5.89	31	10.07	7.44	31	9.61	4.63
	L	31	9.28	4.35	31	8.55	3.79	31	9.71	8.68	31	9.21	4.39
	–	62	9.30	5.04	62	9.06	4.94	62	9.89	8.02	62	9.41	4.48
Selected products (sum)	H	31	3.81	0.48	31	3.81	0.48	31	3.77	0.50	31	11.39	0.88
	L	31	3.90	0.30	31	3.77	0.50	31	3.81	0.48	31	11.48	0.85
	–	62	3.85	0.40	62	3.79	0.48	62	3.79	0.48	62	11.44	0.86
Resumption time (in sec)	H				26	3.45	2.63	26	4.26	2.35	29	3.88	1.83
	L				28	4.41	3.20	28	4.30	3.18	31	4.47	2.79
	–				54	3.95	2.95	54	4.28	2.79	60	4.19	2.37
Final recall (in %)	H										31	51.01	25.01
	L										31	73.59	14.50

Note. H: mental demands enhanced, L: mental demands not enhanced, –: no separation by mental demands. Differences in reported subsample sizes result due to missing values.

According to [17], for purposes of model comparison, human and model data points for high respective low mental demands, as displayed in *Table 1* and *Table 2*, were compared by visual as well as numerical means.

The graphic comparison is shown in *Figure 3* and *Figure 4*. Obviously, apart from the amount of selected products, human data remain on a continuously higher level for all displayed parameters, indicating the model predominantly performed better than the human sample. For *product selection time*, in the case of enhanced mental demands, model and human data point towards a similar direction, marginally increasing with enhanced frequency of interruption. In conditions with low mental demands, model data form a nearly straight line but human data show a considerable difference with increasing interruption frequency. Regarding the *number of selected products*, human data stay almost at the same level across interruption frequencies for both kinds of mental demands but model data perform differently. Whereas there is a noticeable decrease with increasing interruption frequency for high mental demands, a decrease towards highly interrupted trials occurs for low mental demands. Considering *resumption time*, the visual comparison indicates a high similarity between model and human data for both levels of mental demand, although deviation between both datasets is slightly higher under conditions with low mental demands.

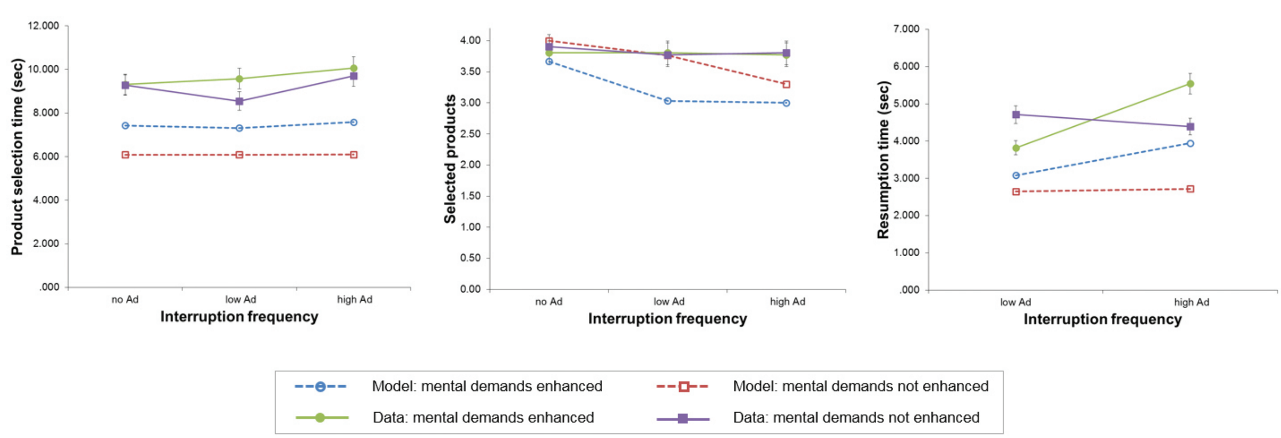

Figure 3: Comparison of model and human data concerning time, products and resumption. Error bars represent 95 % confidence intervals on human data.

In terms of *final recall performance*, the model performs considerably better in both conditions but model and human data exhibit an equivalent trend with a significantly higher amount of recalled products under low mental demands.

For numerical comparisons, the commonly applied *root mean squared deviation (RMSD)* was used. It assesses the mean deviation between model and human data points regarding their exact location in units of the respective scale. Taking into account the scale's extremities, smaller values indicate less distance between both datasets, i. e. point to a better fit. Results of the comparison, separated by the level of mental demands, are shown in *Table 3*. As displayed, e. g. in terms of resumption time under high mental demands, model and human data are located quite close together, whereas in terms of final recall, deviation is quite substantial.

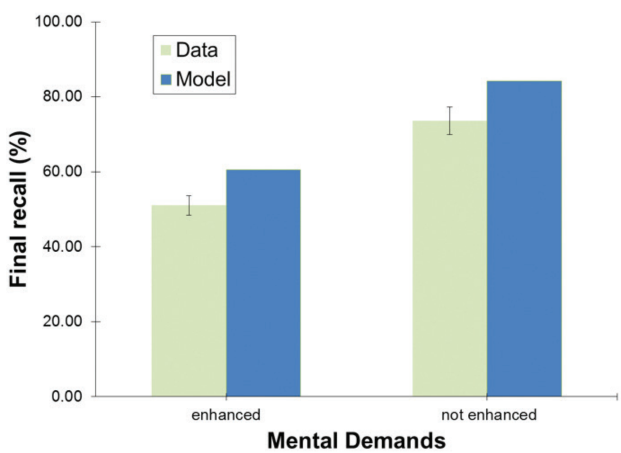

Figure 4: Comparison of model and human data concerning final recall. Error bars represent 95 % confidence intervals on human data.

Table 3: Differences in exact locations between model and human data points.

	RMSD	
	High mental demands	**Low mental demands**
Product selection time (maximum: 10.07 sec)	2.230 sec	3.125 sec
Selected products (range: 0–4 products)	0.637 products	0.298 products
Resumption time (maximum: 4.41 sec)	0.343 sec	1.675 sec
Final recall (range: 0–100 %)	10.076 %	

4 Discussion

This research aimed to inspect the effects of interruption and increased mental demands on human cognition in a smartphone task by applying a cognitive user model. The expected decrease included in the model could be confirmed experimentally in particular for the amount of finally recalled products. Taking a closer look, there might be a number of explanations for the observed differences between model and experimental data.

On one hand, the used interruption could have lacked disruptiveness due to its short duration and familiar content. Indeed, more than 80 % of the tested participants reported being familiar with smartphone use and for this reason may deal with interrupting advertisements on a regular basis. Because of its shortness, the interruption may have not been able to prevent people from rehearsing the content of the product list during its appearance. Moreover, participants conducted the task at their own pace, potentially resulting in the performance of short cognitive breaks to create selective mental cues before actually reading the advertisement [13]. To achieve stronger disruption effects, it could be possible to increase disruptiveness by extending the duration or the amount of cognitive demands needed to deal with the interruption. In the given context, for instance, both aspects could be realized by adding prices that have to be compared or subtracted from a given budget. In addition, in some cases participants seemed to have negated instructions on increased mental demands, resulting in a highly similar performance between both conditions. Solving this issue might require broad changes in instruction, task setting or both to ensure that participants are motivated to behave in the desired manner.

On the other hand, computational models like the used one implicate focusing on selected aspects of the inspected task. Of course, this goes along with limitations in expressiveness and the scope of the model, always raising the issue of the cost and benefit of a given level of model complexity. There are various opportunities to extend the model in order to enhance proximity to the obtained human data. A next step might comprise adjusting and / or including parameters that affect chunk activation and retrieval to achieve a more lifelike memory performance. Additionally, in the longer run the inclusion of further elaborated features like individual differences in working memory [4] or strategies and heuristics of decision-making [12] when facing an interruption might depict valuable extensions as well.

In conclusion, the obtained results can definitely be taken as a cue for the benefit of using such an approach to predict and explain task-related cognitive processing in the given context. In particular due to their value in being applied at an early stage, they provide valuable input for developers and designers in creating interfaces able to actively support users when being interrupted.

References

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

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

[3] Anderson, J. R., & Lebiere, C. (1998). *The atomic components of thought*. Mahwah, NJ: Erlbaum.

[4] Anderson, J. R., Reder, L. M., & Lebiere, C. (1996). Working memory: Activation limitations on retrieval, *Cognitive Psychology, 30*, 221–256.

[5] 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: Springer Science + Business Media.

[6] 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 *Proceedings of the CHI 2015, April 18–23 2015*. Seoul, Republic of Korea: ACM Press.

[7] Brixey, J. J., Robinson, D. J., Johnson, C. W., Johnson, T. R., Turley, J. P., & Zhang, J. (2007). A concept analysis of the phenomenon interruption. *Advances in Nursing Science, 30*(1), E26–E42.

[8] Cowan, N. (2010). The magical mystery four: How is working memory capacity limited, and why?. *Current Directions in Psychological Science, 19*(1), 51–57.

[9] Cowan, N., Morey, C. C., & Chen, Z. (in press). The legend of the magical number seven. In S. Della Sala (Ed.), *Tall tales about the brain: Things we think we know about the mind, but ain't so*. Oxford University Press.

[10] Gillie, T., & Broadbent, D. E. (1989). What makes interruptions disruptive? A study of length, similarity, and complexity, *Psychological Research, 50*, 243–406.

[11] Gray, W. D., Young, R. M., & Kirschenbaum, S. S. (1997). Introduction to this special issue on cognitive architectures and human-computer interaction. *Human-Computer Interaction, 12*, 301–309.

[12] Marewski, J. N., & Mehlhorn, K. (2011). Using the ACT-R architecture to specify 39 quantitative process models of decision making. *Judgment and Decision Making, 6*, 439–519.

[13] McFarlane, D. C., & Latorella, K. A. (2002). The scope and importance of human interruption in human-computer interaction design. *Human-Computer Interaction, 17*, 1–61.

[14] Miller, G. A. (1956). The magical number seven plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 63*, 81–97.

[15] Prezenski, S., & Russwinkel, N. (2014). Combining cognitive ACT-R models with usability testing reveals users mental model while shopping with a smartphone application. *International Journal on Advances in Intelligent Systems, 7*(3–4), 700–715.

[16] Salvucci, D. D., & Taatgen, N. A. (2010). *The multitasking mind.* New York: Oxford University Press.

[17] 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.

[18] Statista (2015a). *Anzahl der Smartphone-Nutzer in Deutschland in den Jahren 2009 bis 2015 (in Millionen)* [Amount of smartphone users in Germany from 2009 to 2015 in millions]. Retrieved from http://de.statista.com/statistik/daten/studie/198959/umfrage/anzahl-der-smartphone-nutzer-in-deutschland-seit-2010/ at June 17th, 2015.

[19] Statista (2015b). *Prognose zur Anzahl der Smartphone-Nutzer weltweit von 2012 bis 2018 (in Milliarden)* [Predicted amount of smartphone users worldwide from 2012 to 2018 in billions]. Retrieved from http://de.statista.com/statistik/daten/studie/309656/umfrage/prognose-zur-anzahl-der-smartphone-nutzer-weltweit/at June 17th, 2015.

[20] 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.

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

Bionotes

Maria Wirzberger
TU Chemnitz, Germany
maria.wirzberger@phil.tu-chemnitz.de

Maria Wirzberger, M. Sc. works as research assistant within the interdisciplinary DFG Research Training Group "CrossWorlds" at the TU Chemnitz. In 2012, she received a B. Sc. in Psychology from the University of Hagen, and completed the master's program Human Factors at the TU Berlin in 2014. Her current PhD project focusses on connecting cognitive modeling and instructional design by exploring the construct of cognitive load with the cognitive architecture ACT-R.

Prof. Dr.-Ing. Nele Russwinkel
TU Berlin, Germany
nele.russwinkel@tu-berlin.de

Nele Russwinkel holds a junior professorship for Cognitive Modeling in dynamic Human-Machine Systems at the TU Berlin. Based on a B. Sc. and M. Sc. in Cognitive Science from the University of Osnabrueck and some experience in industry, she held a scholarship within the DFG Research Training Group "ProMeTeI" at the TU Berlin, where she obtained her PhD in 2009. Besides others, her research is concerned with cognitive modeling of human-computer interaction on purposes of usability, embodied spatial cognition and time estimation.