Cognitive Model Data Analysis for the Evaluation of Human Computer Interaction

Jeronimo Dzaack[1] and Leon Urbas[2]

[1] Berlin University of Technology, Center of Human-Machine Systems
Franklinstraße 28-29,
Sekretariat FR 2-7/2, D-10587 Berlin
Jeronimo.Dzaack@zmms.tu-berlin.de
[2] Dresden University of Technology, Institute of Automation
D-01062 Dresden
Leon.Urbas@tu-dresden.de

Abstract. In industry and consumer electronic, more and more operative tasks are changing to supervisory control and management tasks. This leads to more complex and dynamic user interfaces (e.g. integrated control interfaces, infotainment systems in cars). Because of the integrated functionality and the complex data structures, these interfaces require more cognitive information processing. Usability of such interfaces can be evaluated by using cognitive modeling to investigate cognitive processes and their underlying structures. So far the explanatory power of cognitive models is limited due to the lack of fine-grained simulation data analysis. Having realized this drawback we developed SimTrA (*Sim*ulation *Tr*ace *A*nalyzer) to simplify the analysis of cognitive models. The tool automatically processes and analyzes data from cognitive models and allows the comparison of simulated data with empirical eye movement data. This paper describes the approach and its implementation. The practicability of SimTrA is demonstrated with an example in the domain of process control.

Keywords: Cognitive Architectures, Eye Movement Data, Analysis, Human Computer Interaction.

1 Introduction

Recent introductions of new information technologies in the field of dynamic human-machine systems (e.g. process control systems in the chemical industry or airplane cockpits) have led to increasing cognitive requirements caused by a shift from manual process operation to the management of complex automated processes. This calls for highly complex dynamic user interfaces. Because of the integrated functionality and the complex data structures, these interfaces require more cognitive information processing. In a practical application cognitive models can be used to evaluate the usability of prototypes. This helps to detect errors in the interaction design of interfaces and gives indications about the cognitive demands of the future user. Cognitive models extend classical usability methods (e.g. questionnaires, observing) and expand the

D. Harris (Ed.): Engin. Psychol. and Cog. Ergonomics, HCII 2007, LNAI 4562, pp. 477–486, 2007.

repertoire by cognitive aspects. Consider three groups each developing an interface for a new human-machine system with the same functionality. The three interfaces can differ in appearance, and the choice and arrangement of displayed information. Empirical tests can be conducted to detect the interface that best fits human cognitive demands, or a cognitive model can operate the machine via the interface. The quantitative and normative cognitive data output (i.e. error rates, times, decisions) helps to evaluate the interfaces and to detect the best one for the given purpose. This could shorten the development process and reduce costs and required resources. The potential of cognitive architectures and cognitive models is known in research and development related to usability. However, this method is rarely employed in usability research and development because of a lack of support tools for cognitive modeling. Focus of this research is the analysis of cognitive model data.

2 Cognitive Modeling as Method for the Evaluation of HCI

Cognitive architectures incorporate psychological theories (e.g. visual information processing, decision making) and empirically based representations about aspects of human cognition. There is general agreement that these cognitive aspects are relatively constant over time and relatively task-independent [8]. Therefore cognitive architectures present these aspects in a software framework to explain and predict human behavior in a detailed manner. In this context, a cognitive model can be seen as an application of a cognitive architecture to a specific problem domain with a particular knowledge set. Building a cognitive model, the modeler must describe cognitive mechanisms in a highly-detailed and human-like way. Two levels of cognitive architectures can be differentiated [14]. High-level architectures (e.g. GOMS [4]) describe behavior at a basic level and define interactions as a static sequence of human actions. Low-level architectures (e.g. ACT-R, SOAR or EPIC, see [3] for an overview) describe human behavior at an atomic level. They allow a more detailed insight into cognitive processes than high-level architectures. Most low-level architectures use production systems to simulate human processing and cognition. The use of independent production rules allows cognitive models to react on external stimuli (bottom-up processes) and to model interruption and resumption of cognitive processes in contrast to high-level architectures which are usually controlled top-down. The research presented in this paper uses the cognitive architecture ACT-R [1].

To use cognitive models for the evaluation of Human Computer Interaction, the analysis of the cognitive model data is an important aspect. Most low-level simulation experiments use global information to analyze the model and its fit to empirical data (e.g. errors, times). There are two problems. First, in order to ensure that a cognitive model acts like a human and that this behavior can be explained theoretically, the results with a cognitive model and the human and the internal computations to achieve these results have to be the same [15]. Second, different cognitive models can have the same overall information but can differ in underlying sub-processes. That means that in order to analyze the model's performance on the basis of the integrated psychological theories (e.g. visual perception and processing) it is necessary to observe the global structure and also the microstructure level to identify the underlying processes [6]. The microstructure can be characterized by the sub-processes of the cognitive model, i.e. repeated short sequences of action such as control-loops. This makes it possible to analyze the kind of computations that lead to a result and to determine the level of correctness of the cognitive model with respect to the underlying psychological theories.

However, using cognitive models for usability evaluation reveals some problems. Cognitive architectures and models are incomplete and do not describe all the processes that are responsible for human cognition [10]. This is due to the partial knowledge of internal cognitive processes and also the fact that it is not necessary to implement cognitive aspects such as esthetics, boredom, fun, or personal preferences in order to explain an effect [3]. Analyzing and comparing cognitive model data with empirical data requires taking these differences into account.

So far, no tools exist for the extraction of fine-grained information from model data for the evaluation of user interfaces.

3 Analyzing Cognitive Model Data

The research project aims to support the analysis of cognitive model data for the evaluation of human computer interaction. For this reason the *Sim*ulation *Tr*ace *A*nalyzer (SimTrA) was developed. The implemented process of SimTrA is divided into three components: (1) preprocessing of the raw data, (2) analysis of the preprocessed data, and (3) comparison with further models or empirical data (see Figure 1). This allows the independent development of each module and the easy extension of its functionality in the future. SimTrA enables the user to apply basal algorithms regarding the process of analyzing and comparing. After each step the processed data is stored in a general-purpose format and can be processed by external tools (e.g. MatLab, R, SPSS). It is possible to import cognitive model data as well as empirical data in the tool. The results are plotted as in classical usability-evaluation methods.

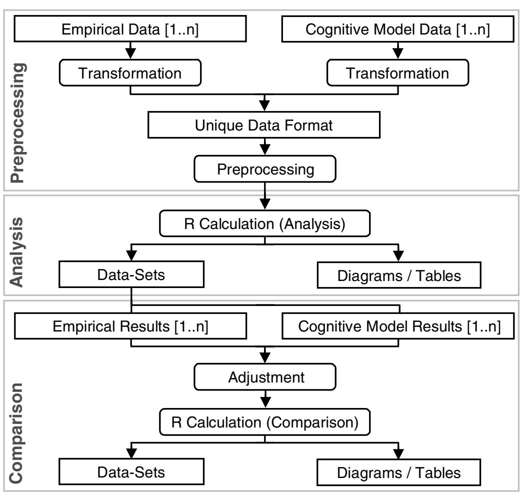


Fig. 1. Overview of the implementation of SimTrA

Eye movement studies in psychological and cognitive research give insights into human behavior and underlying cognitive processes [9, 12, 13]. Cognitive models with ACT-R can process visual information with the integrated visual module, providing spatial and temporal information of the simulated eye movement as found in empirical studies. The extraction of this information provides a way to analyze and compare cognitive model data with empirical data at a global level (e.g. overall number of fixations in a trial and their durations) and a microstructure level (e.g. scan-paths). In a prior study we showed the possibility to extract simulated eye movement data from cognitive models and to compare it with human eye movement data [5]. That is the reason for the implementation of SimTrA for eye movement data.

3.1 Preprocessing

For the preprocessing, raw data of empirical experiments or cognitive model experiments can be imported into SimTrA and transferred into a general-purpose format (see Figure 1) which contains all information on the eye movement. A plug-in has to be integrated into the cognitive model which allows the storage of the current location of focus, timestamp, action and state of the visicon of ACT-R (i.e. state of the simulated interface) for each access of the visual module by the cognitive model (e.g. shift of attention, encoding) in an Extended Markup Language file (XML-File). For the empirical data, the raw data output of the eye movement tracking software is reduced to the current location of focus and timestamp (minimum demands of eye movement studies). The transformation allows processing of data of different quality and origin (e.g. different cognitive architectures, empirical studies) with the same algorithms.

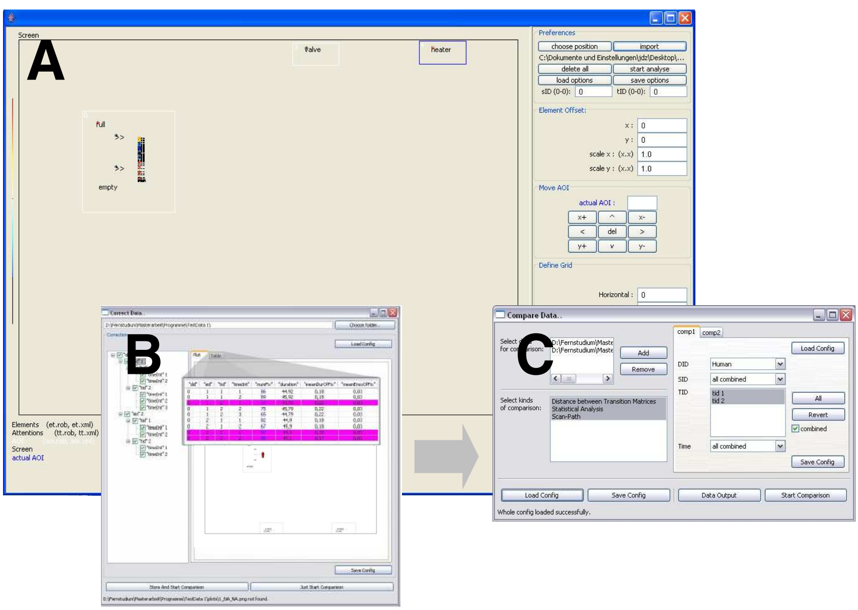


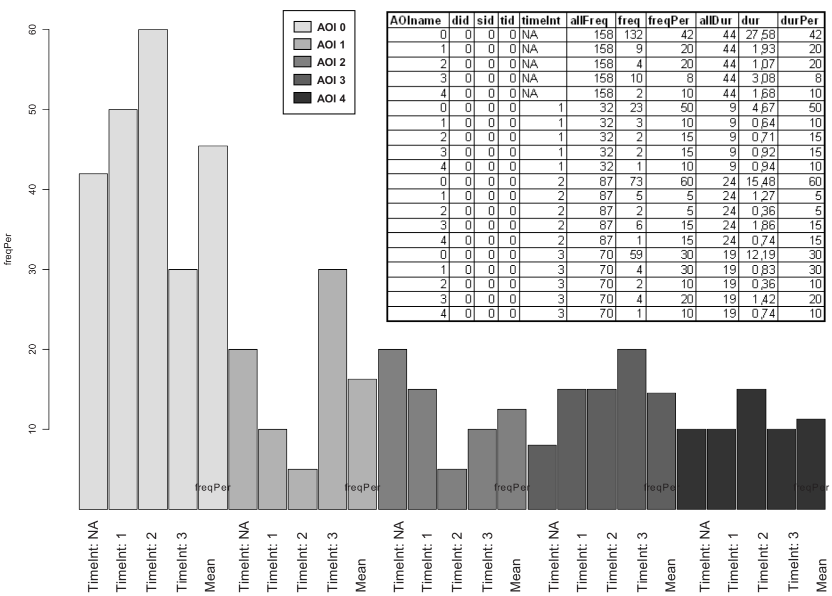
Fig. 2. Screen views of SimTrA. A: Preprocessing. B: Adjustment of data. C: Comparison of data.

While preprocessing each dataset, the stored data is augmented by information on the origin of the data (e.g. cognitive model or empirical data), the subject and the trial identification (i.e. data-Id, subject-Id and trial-Id). Each kind of information can be altered by the user in the preprocessing component of SimTrA.

In this component, the data of different origins can be scaled to a similar resolution by the user to ensure that the data is comparable. To apply eye movement algorithms, the user has to enrich the current data with additional information that is needed for the analysis (i.e. areas of interest, time-intervals) and has to set the output folder. This is done in the preprocessing interface (see Figure 2) by simply clicking in the plotted data-points to set an area of interest (AOI) or by using the tools for the time-intervals and other additional information. The preprocessing of the raw data is finished by choosing the desired analysis methods. All information is stored in XML-files in the output folder, and is available for reuse (e.g. AOIs, time-intervals, and datasets), making it possible to skip the import of the raw data in some subsequent analyses. For this the user can use the *load options* command in the interface.

3.2 Analysis

This component enables the analysis of the preprocessed data (see Figure 1). The algorithms for the analysis are implemented in R, a free tool for statistical computing [11]. The results are saved in tables and as graphical plots (see Figure 3). In this step algorithms are implemented to analyze the transition frequencies, the fixations on AOIs, the spatial density, the statistical dependency in visual scanning, the local scanpaths and a mean for each of these algorithms (for an overview see [13]). In the result



AOIname	did	sid	tid	timeInt	allFreq	freq	freqPer	allDur	dur	durPer
0	0	0	0	NA	158	132	42	44	27,58	42
1	0	0	0	NA	158	9	20	44	1,93	20
2	0	0	0	NA	158	4	20	44	1,07	20
3	0	0	0	NA	158	10	8	44	3,08	8
4	0	0	0	NA	158	2	10	44	1,68	10
0	0	0	0	1	32	23	50	9	4,67	50
1	0	0	0	1	32	3	10	9	0,64	10
2	0	0	0	1	32	2	15	9	0,71	15
3	0	0	0	1	32	2	15	9	0,92	15
4	0	0	0	1	32	1	10	9	0,94	10
0	0	0	0	2	87	73	60	24	15,48	60
1	0	0	0	2	87	5	5	24	1,27	5
2	0	0	0	2	87	2	5	24	0,36	5
3	0	0	0	2	87	6	15	24	1,86	15
4	0	0	0	2	87	1	15	24	0,74	15
0	0	0	0	3	70	59	30	19	12,19	30
1	0	0	0	3	70	4	30	19	0,83	30
2	0	0	0	3	70	2	10	19	0,36	10
3	0	0	0	3	70	4	20	19	1,42	20
4	0	0	0	3	70	1	10	19	0,74	10

Fig. 3. Example output of SimTrA

tables the data is distinguished by the data-Id, subject-Id and the trial-Id. The absolute and the cumulative number of fixations, fixation durations and general information are stored in a general table. Each algorithm accesses the general table and processes the information needed for the calculation of the respective parameter.

To support the user, plots are generated for each analysis. All input-files (i.e. empirical and cognitive model data), results, general information and the distribution of eye movement data are stored in either tables or plots.

3.3 Comparison

This component enables the comparison of the analyzed data from cognitive models or empirical studies (see Figure 1) and the reapplication of analyses of data files with adjusted data. The comparison allows the rating of the data (e.g. cognitive model vs. empirical data, cognitive model 1 vs. cognitive model 2) with respect to the given context (e.g. empirical findings, psychological theories). The reapplication and the adjustment allows the user to detect missing data and to delete the corresponding dataset and finally to reanalyze the data file.

In both cases the data has to be loaded into the comparison-component. In a first step each data file can be adjusted by the user (e.g. missing data, insufficient data points). In this step all datasets of one data file are displayed and the user can select or deselect the datasets by clicking on them. To support the user, the upper and lower boundary values of the given data for one column of the displayed general table (upper and lower percentile, [2]) are calculated and displayed and a plot of the distribution of the eye movement is shown for each dataset (see Figure 2). The percentile thresholds and the column used for the calculation can be altered by the user. All adjustments are stored in a file in the input folder and can be loaded into the tool if this step is repeated.

In the next step, the type of algorithms and the output folder are chosen by the user and the comparison or analysis is started. The algorithms for the comparison are also implemented in the software R. The comparison is done by descriptive methods because cognitive model data do not show a high variance and statistical interference methods are not applicable in this case. The input, the analyzed data and the compared data are stored in tables and plotted in files.

4 Practical Application: Process Control System

We conducted an study to show the practicability of SimTrA. Therefore, we processed and compared simulated eye movement data of two different cognitive models with empirical data. Eye movements of 12 participants were recorded during their interaction with three complex dynamic interfaces of a process control system microworld (randomized within design, independent variable: version of interface; dependent variables: eye movement data and performance). The task lasted 45 seconds for each interface. The aim was to stabilize the level of liquid in a container, which is moderated by inflow, outflow and evaporation, between two boundary values. The parameters can be regulated by a valve and a heater connected with the container. The three interfaces differ in the position of the interaction and feedback elements of the interface (see Figure 4).

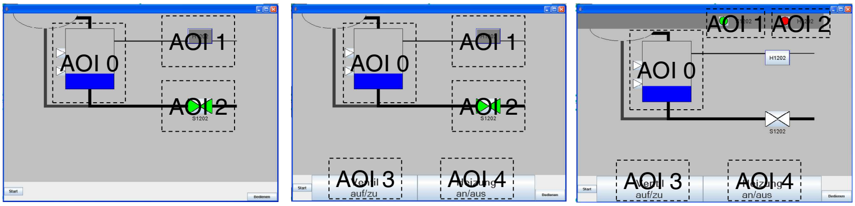


Fig. 4. Three interfaces for one task. A: Combined information and interaction elements. B and C: Distinct information and interaction elements.

For the model based investigation we developed two ACT-R 6.0 models representing the internal knowledge of an experienced user familiar with the technical process (as found in a survey of experts). One model is implemented in the minimal control principle style suggested by Taatgen [16]. That allows the cognitive model to react to external stimuli (bottom-up control). The other model is implemented with a fixed sequence of actions moderated by states (top-down control).

In both the empirical experiment and the simulation, x- and y-coordinates of gaze targets with timestamp and the overall performance of task were recorded. SimTrA was used to analyze the microstructure behavior and to compare the empirical data with the cognitive model data.

4.1 Model Data and Empirical Data

For the evaluation of the practicability of SimTrA, we compared the interaction of the humans (H), cognitive model 1 (bottom-up control, CM 1) and cognitive model 2 (top-down control, CM 2) with the three different interfaces (3 x 3 design). We investigated the overall data and microstructure data.

For overall data analysis, we determined the percentage of the time where the level is between the two boundary values (i.e. in control) and the amount of interactions. The performance of the three groups for all 3 interfaces is nearly the same (H: 90 % in control / Heater: 1 press / Valve: 8 presses, CM 1 & CM 2: 85 % in control / Heater: 1 press / Valve: 9 presses).

For the microstructure analysis we defined three or five AOIs (see Figure 4) and applied all integrated algorithms of SimTrA. Because of the large numbers of tables and plots we discuss an extract of the analyzed data. Ten datasets could be used for the analysis after the adjustment. Table 1 shows the data for the comparison of the local scanpaths (consistent patterns of consecutive fixations) of actual perceptions (closed-loop control) [7]. All theoretical triple of AOIs are determined whereby any sequence of fixations falling into the same AOI is treated as a single gaze fixation and the number of occurrence in the whole sequence of AOIs is assigned to them (e.g. AOIs: 1, 2 - Sequence: 2122112 - triple: 121: 1, 212: 2). Ordering these by frequency shows the most important local scanpaths. The local scanpaths make it possible to determine important visual scanpaths and help to detect faults and obstacles of the interface design that can influence the information encoding and processing [13].

Table 1. Comparison of the frequencies of local scanpaths greater 3 % as found in the empirical study in percent for the three interfaces (H: Human, CM 1: bottom-up model, CM 2: top-down model)

Interface 1				**Interface 2**				**Interface 3**			
Path	**H**	**CM 1**	**CM 2**	**Path**	**H**	**CM 1**	**CM 2**	**Path**	**H**	**CM 1**	**CM 2**
0>2>0	26.1	36.4	9.9	0>2>0	11.6	3.0	37.8	0>3>0	14.6	11.9	9.5
2>0>2	21.7	33.3		0>3>0	11.6	24.2	9.5	0>1>0	6.4	4.8	
1>0>2	8.1	9.1	9.9	0>1>0	7.5	15.2		2>0>3	6.4	2.4	
2>0>1	8.1		9.9	2>0>2	5.5		29.7	3>0>1	6.4	4.8	
0>1>0	6.8	3.0	39.4	3>0>1	5.5	12.1		3>0>3	5.3	7.1	
2>1>0	6.8	9.1	1.4	1>0>3	4.8	12.1		0>4>0	4.7	2.4	1.4
				3>0>2	4.8		8.1	0>1>2	4.1		
				2>0>3	3.4	3.0	9.5	1>2>0	4.1	2,4	1.4
								1>0>1	3.5		
								3>0>2	3.5		8.1
Sum	**83.8**	**90.9**	**70.4**	**Sum**	**54.8**	**69.7**	**94.6**	**Sum**	**59.1**	**35.7**	**20.3**

4.2 Discussion

The empirical study and the simulation show that the three groups do not differ essentially in the overall behavior in the interaction with the three interfaces. Table 1 shows the comparison of the local scanpaths of the three groups (i.e. human, bottom-up model and top-down-model). Scanpaths with a frequency < 3% were excluded and the remaining important scanpaths were used for the comparisons. The table shows that the frequencies and the rank order of the predicted local scanpaths is different between all groups. Nevertheless the bottom-up model (CM 1) seems to have a better fit to the empirical data than the top-down model (CM 2) based on the rank order and the reached cumulative percents of predicted local scanpaths. This allows the conclusion that the cognitive processes that moderate human eye movement and the simulated cognitive processes are different. We infer that the cognitive models and its architecture need to be improved in several respects (e.g. update circles and uncontrolled eye movements). Taking into account all processed data and the extracts shown in this paper, we conclude that SimTrA is able to extract and process cognitive model as well as human eye movement data. SimTrA helps to detect differences and similarities between different cognitive models for the same task as well as between simulated and human eye movement data.

5 Conclusion

We have developed SimTrA to analyze cognitive model data at a microstructure level. In this paper we show the internal approaches and the practicability of SimTrA for the analysis of simulated and human eye movement data. Cognitive models allow a more detailed investigation of simulated cognitive processes (e.g. decision-making, arousal) than the eye movement data. But it has not been established which cognitive processes are executed during human information processing. It is therefore better to rely on observable parameters than on uncertain internal processes.

SimTrA enables the detection of diverse microstructure behavior between different cognitive models that seem to behave equally from an overall point of view. There seem to be three possible areas for the application of SimTrA. First, it allows the verification of distinctive psychological theories integrated in cognitive models. The analysis and comparison of cognitive models implemented with different theoretical groundings can help to identify varieties and similarities in these theories. Second, the analysis and comparison of empirical and simulated data helps to improve the cognitive model and its internal structure. Part of this research is the improvement of cognitive models. Our results accentuate the need to adjust cognitive models at a microstructure level to make human behavior and cognitive models more comparable. And finally, the cognitive model data analysis for the evaluation of human computer interaction is enabled. Altogether, this leads to more detailed model data and allows a better prediction of human behavior for usability questions.

Further work is needed on the extension of SimTrA. It is necessary to evaluate which additional analysis and comparison algorithms can be integrated.

To summarize, our tool enables the analysis and the comparison of cognitive model behavior with human behavior at a microstructure level. But the cognitive model data analysis must be enhanced by a greater variety of algorithms. This could lead to an increased application of cognitive models in the usability evaluation of computer interfaces.

Acknowledgment. This project is funded by Deutsche Forschungsgemeinschaft (Research Training Group 1013 prometei).

References

1. Anderson, J.R., Bothell, D., Byrne, M.D., Douglass, S., Lebiere, C., Qin, Y.: An integrated theory of the mind. Psychological Review 111, 1036–1060 (2004)
2. Bortz, J.: Statistik für Sozialwissenschaftler. Springer-Verlag, Heidelberg (1999)
3. Byrne, M.D.: Cognitive Architectures. In: Jacko, J., Sears, A. (eds.) Handbook of Human-Computer Interaction, pp. 97–117. Lawrence Erlbaum Associates, Hillsdale, NJ (2003)
4. Card, S.K., Moran, T.P., Newell, A.: The Psychology of Human-Computer Interaction. Lawrence Erlbaum Associates, Hillsdale, NJ (1983)
5. Dzaack, J., Urbas, L.: Kognitive Modelle zur Evaluation der Gebrauchstauglichkeit von Mensch-Maschine Systemen. In: Grandt, M., Bauch, A. (eds.) Cognitive Systems Engineering in der Fahrzeug- und Prozessführung. Deutsche Gesellschaft für Luft- und Raumfahrt, Bonn, pp. 305–318 (2006)
6. Gray, W.D., Sims, C.R., Schoelles, M.J.: Musings on Models of Integrated Cognition: What Are They? What Do They Tell Us that Simpler Approaches Cannot? How Do We Evaluate Them? In: Proceedings of the 7th ICCM, pp. 12–13. Lawrence Erlbaum, Hillsdale, NJ (2006)
7. Groner, R., Walder, F., Groner, M.: Looking at faces: local and global aspects of scanpaths. In: Gale, A.G., Johnson, F. (eds.) Theoretical and Applied Aspects of Eye Movement Research, pp. 523–533. Elsevier, Amsterdam (1984)

8. Howes, A., Young, R.M.: The role of cognitive architecture in modeling the user: Soar's learning mechanism. Human-Computer Interaction 12(4), 311–343 (1997)
9. Just, M.A., Carpenter, P.A.: Using eye fixations to study reading comprehension: Individual differences in working memory. Psychological Review 99, 122–149 (1984)
10. Opwis, K.: Kognitive Modellierung: Zur Verwendung wissensbasierter Systeme in der psychologischen Theoriebildung. Huber, Bern Göttingen Toronto Seattle (1992)
11. R Development Core Team: R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria (2006)
12. Rayner, K.: Do eye movements reflect higher order processes in reading? In: Groner, R., d'Ydewalle, G., Parham, R. (eds.) From Eye to Mind: Information Acquisition in Perception, Search, and Reading, Elsevier, New York (1999)
13. Rötting, M.: Parametersystematik der Augen- und Blickbewegungen für arbeitswissenschaftliche Untersuchungen. Shaker, Aachen (2001)
14. Salvucci, D.D., Lee, F.J.: Simple Cognitive Modeling in a Complex Cognitive Architecture. In: Proceedings of the CHI '03, pp. 265–272. ACM Press, New York (2003)
15. Strube, G.: Cognitive Modeling: Research Logic in Cognitive Science. In: Smelser, N.J., Baltes, P.B. (eds.) International encyclopedia of the social and behavioral sciences, pp. 2124–2128. Elsevier, Oxford (2001)
16. Taatgen, N.A.: The Minimal Control Principle. In: Gray, W. (ed.) Integrated Models of Cognitive Systems. Oxford University Press, Oxford (in press)