How to model different strategies in dynamic task environments

Juergen Kiefer (juergen.kiefer@ zmms.tu-berlin.de)
User Modeling in Dynamic Systems,
Technische Universität Berlin
Jebensstraße 1, J2-21, 10623 Berlin, Germany

Leon Urbas (leon.urbas@zmms.tu-berlin.de)
User Modeling in Dynamic Systems,
Technische Universität Berlin
Jebensstraße 1, J2-21, 10623 Berlin, Germany

Abstract

In a study about interaction with in-vehicle information systems, participants were put in a dual-tasking scenario. They had to operate a basic research test while driving in a driving simulator. Individual differences were found both in single tasking as well as in multitasking condition. It turned out that during multitasking participants seem to change their interaction strategies. In an immediate follow up single task setting, the participants retain these newly learned strategies. This paper introduces the experimental background of the study, presents corresponding ACT-R models that reflect the participants` behavior before and after the dual tasking treatment.

Individual differences

Modeling of individual differences in the context of user-centered design is of particular interest. From a cognitive point of view, psychology provides two perspectives on how to explain individual differences: a first approach focuses on different aptitudes or abilities of people. Differential psychology – one core discipline on individual differences – is interested in concepts such as intelligence or creativity. Ackerman (2005), for instance, is a prominent supporter of this approach. In his view, individual differences are based on ability determinants. He proposes to consider individual differences in respect of skill acquisition. The clarifying example of driving illustrates this position: at the beginning, we are novices altogether. Practice makes us more or less experts in driving. This first direction is denoted as parameter-based approach. In the context of cognitive modeling, prominent parameters that have already been considered are working memory (capacity) or processing speed (Schunn & Reder, 2001). For instance, Daily, Lovett & Reder (2001) propose an ACT-R model on individual differences that focuses on working memory performance. Lovett (2001) gives a clear distinction between architectural differences, such as processing speed and working memory capacity, and knowledge-based differences (Lovett, 2005).

A second direction accounts for individual differences on the basis of individual, different strategies. Different strategies, for example, can be attributed to different groups of people (e.g., younger vs. older children). Facts about the world, experience and strategies belong to this approach. Schunn & Reder (1998) propose to combine the parameter-based with the strategy-based approach and to pay more attention to the aspect of strategies selection. The strategy that is most appropriate for a concrete situation is selected among a set of possible strategies. One major advantage of this perspective on individual differences is its explanatory power in many domains. A popular domain in the field of human-machine interaction is driving. The focus of this paper is on individual strategies in a multitasking scenario in a driving simulator. We argue that in our study, not all participants use the same strategy. Strategies, moreover, seem to be caused by adaptive behavior in the context of multitasking.

Overview

To investigate individual differences in human-machine-interaction, we investigated a dual tasking scenario in which participants were asked to perform a task while driving in a simulator. We believe that users in human-machine-interaction on the one side feature different abilities. But these differences do not solely explain the broad range in performance, rather different strategies account for the diversity found in the data, especially those developed while dual tasking. In addition, individual strategies, in some cases, can lead to the same output. But they can as well produce differences in performance. In what follows, a dual tasking study will be introduced. We present three computational models: the first model is an ACT-R 6 update of the ACT-R/PM model developed by Dzaack, Kiefer & Urbas (2005). This model serves as baseline for the secondary task and reflects the behavior of most of the participants in the pretest. A second model illustrates the behavior of another group of people in the pretest. The third model explains the observed increase in participants' performance in the posttest. The article closes with a critical discussion and an outlook on future work.

Experimental study

Design

In a multitasking scenario with compound continuous tasks (for a specification of different task types, see Salvucci, 2005), 36 students of the TU Berlin were asked to perform a test of attention (referred to as D2-Drive-test) while driving in a simulator.

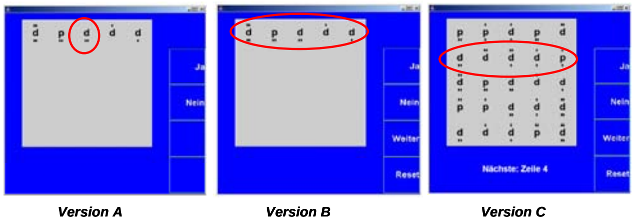

Figure 1: D2-Drive-test (Version A, B, C)

The study is an extension of the study by Dzaack, Kiefer & Urbas (2005). The complete sequence contained training (participants get used to the D2-Drive-test), two pretests, 8 real-tests (dual tasking condition) and two posttests. Post driving, the experiment ended with a structured interview on participants` applied strategies in the D2-Drive-test, both while single tasking as well as dual tasking. Unlike other studies in this area (Gunzelmann & Anderson, 2002), participants were *not* trained specific strategies, the strategies rather emerge without instruction. Sequence and number of trials were unknown to the participants to prevent motivational effects. As in the previous study, we used three versions of the D2-Drive-test (Fig. 1). Test-version (A vs. B vs. C) was assigned to three groups of participants. In version A, participants have to decide whether the pattern in the middle of the line contains the letter "d" and two strokes. After the response, the next screen appears and the procedure restarts. In version B, the decision has to be made for all of the patterns in the entire row, after the fifth pattern a new screen appears. Version C is an extension of B inasmuch as the number of the subsequent row has to be memorized. Version B and C are of special interest in the dual tasking condition: due to their structure (performance of a complete line instead of only one single pattern in the middle), they both can be interrupted on various positions: this structure releases individual, self-defined strategies of coping with timing, coordination and resource allocation in multiple task environments. Dependent variables correspond to driving behavior (deviation from the middle line), eye movement (gazes at lane) and performance in the D2-Drive-test (reaction time, number of correct and incorrect responses). In a structured interview after the posttest, we took verbal responses of the participants on experienced performance, used strategies and overall questions on the management and requirements of the complete study.

The aspect of practice

In the study by Dzaack, Kiefer & Urbas (2005), the D2-Drive-test was presented on the in-vehicle display. To get accustomed to the D2-Drive, participants were trained until performance was error-free without driving (but in the car). Error-free means a (not further defined) sequence of patterns was responded correctly. Of course, this is no guaranty for participants not doing any error at all furthermore. After 50 patterns training, the error-free level was reached for all participants. They then were asked to drive at constant speed (130km/h) and to perform the D2-Drive-test for 60 seconds while driving. We chose to repeat the D2-Drive-test three times. While the performance while driving suffered significantly, the performance in the posttest increased compared to the pretest. As Newell & Rosenbloom (1981) point out, "practice brings improvement, and more practice brings more improvement". To investigate the limits of learning by practice for this task, in this study the number of dual tasking trials was enhanced to 8 trials. For the current study, we expected a sharper increase of performance (number of executed patterns) in the subsequent posttests. Sufficient practice includes a change from a cognitive to an associative to an autonomous stage (Taatgen, 2005). Consequences of this processing are a change from error-prone to error-free and from slow to fast performance. For the current study, we therefore expect shorter reaction times from trial to trial.

Results

For the purpose of this article, we concentrate on the performance in the D2-Drive-test. For single tasking, performance in the D2-Drive-test for version A and B turned out to be almost identical. The number of performed patterns (74 patterns per trial on average). Version C produces a lower number of executed patterns as well as the highest error rate (on average, 3-4 errors per trial, compared to 0-1 error per trial for A and B). The subsequent modeling approach describes a general model of interaction with the D2-Drive-test, the model for explaining the pre-post test difference however focuses on version B only.

Table 1: Average number of patterns (#) / *min* and average reaction time (rt) in *msec* per pattern, both in pre- and in post-test

	# pre	# post	rt pre	rt post
D2-A	74	82	815	733
D2-B	74	92	826	661
D2-C	56	65	1131	945

Performance increase as a function of practice

Table 1 represents number of responses and reaction times for each version of the D2-Drive-test. Throughout all three versions, we observe a permanent increase in performance: from pre- to posttest, participants perform more patterns (for version A, F = -2.4, $p < .05$, for version B, F = -3.7, $p < .01$). Interestingly, version B gains the highest benefit: in this version, performance increase is more pronounced than in A and C (See table 1). We hypothesize that one reason for this is the development of individual strategies. In the current study, strategies are build on the basis of practice and experience. The verbal protocols show that participants indeed apply different strategies (based on the processing of a pattern reported by the participants) and demonstrate that the participants do not necessarily have to be conscious of these applied strategies.

Modeling different strategies

ACT-R is a cognitive architecture and theory for simulating and understanding human cognition. It contains modules, buffers, and a pattern matcher. ACT-R is a framework for different tasks with perception, motor, memory and information processing activities that can be described by means of a rational analysis. Because interaction with the D2-Drive-test includes all of these elements, we chose ACT-R to test our assumptions on a fine granular quantitative level.

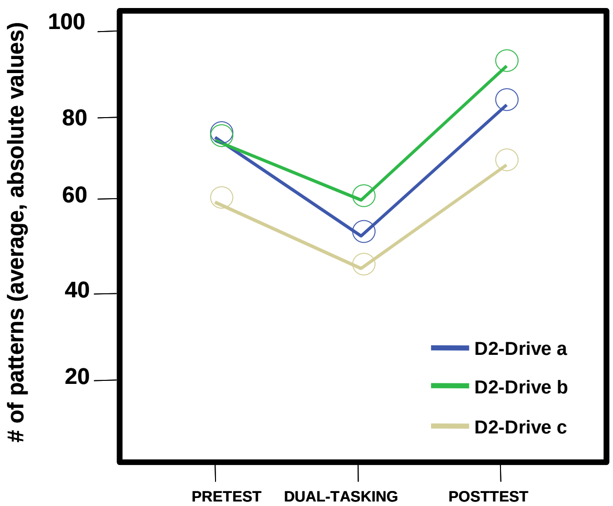

Figure 2: Performance in the D2-Drive-test

Based on the empirical results of a previous study (Dzaack, Kiefer & Urbas, 2005), an ACT-R/PM model of the D2-Drive-test was build. This model describes reading of a pattern as a sequence of encoding of the distinct parts of the pattern (center d or p, upper strokes, lower strokes), followed by a manual response (pressing the keys 'y' or 'n'). In Table 2, this basic structure is referred to as *S-I*. In detail, a rational analysis would suggest, that patterns containing 'd' or 'p' are handled differently. For a *P-Pattern*, encoding the letter would be sufficient to come to a decision. Our empirical results support this modeling hypothesis with

(1) shorter reaction times for *P-Patterns*
(2) verbal responses of the participants

Encoding a 'p' leads immediately to a no-response, whereas the detection of a 'd' alone is not sufficient to judge the pattern as a D2 – both the upper and the lower part have to be considered to come to a decision. The core part of the *S-I*-model of the D2-Drive-test contains the steps *move eyes to pattern*, *encode pattern*, *store result* and *respond*. This strategy leads to response times that can be predicted by the type of pattern. An Analysis of the empirical data confirms that this is the dominant strategy for all participants in version A, and for most (but not all) participants in version B during the pretest-phase.

These other subjects seem to encode a block of two or more patterns first and then respond the memorized items in a sequence of rapid key-presses. We call this alternative strategy "blocking" to illustrate the corresponding response blocks. Please note that the two strategies for processing a pattern (Dzaack, Kiefer, & Urbas, 2005) occur in all three conditions. The "blocking"-strategy implicitly uses the structure of version B of the D2-Drive-test, i.e. it is applied in B but not in A (where it is neither possible nor plausible). Table 2 illustrates how the emerging of blocking is implemented. In contrast to the generic structure in *S-I*, for the blocking strategy it is necessary to consider which pattern is performed. The applied blocking strategy contains two elements. The sequence is as follows: attention is put on the 1st pattern (*move eye*), the 1st pattern is encoded and the result is appended to ***isD2***, the 2nd pattern is encoded and the result is stored in ***isD2***. That is, ***isD2***, the is a chunk that contains a sequence of two answers. Both results are entered (two productions) and attention it put to the next pattern (another production).

Table 2: Pseudo code of strategies for handling a single screen of the D2-Drive-test at different levels of experience

S-I	Basic model **FOR EACH PATTERN IN PATTERNS** **MOVE EYES TO PATTERN** **ENCODE PATTERN** **STORE RESULT IN isD2** **RESPOND isD2***
S-II	Disjunctive blocking (2 elements) **FOR EACH BLOCK IN PATTERNS** **FOR EACH PATTERN IN BLOCK** **MOVE EYES TO PATTERN** **ENCODE PATTERN** **APPEND RESULT TO isD2** **FOR EACH response IN isD2*** **RESPOND response**
S-III	Aggregated blocking (2 elements) **FOR EACH BLOCK IN PATTERNS** **FOR EACH PATTERN IN BLOCK** **MOVE EYES TO PATTERN**** **ENCODE PATTERN** **APPEND RESULT TO isD2** **FOR EACH response IN ALL-BUT-LAST(isD2)*** **RESPOND response** **RESPOND isD2* | MOVE-EYES TO 1.st PATTERN of NEXT BLOCK**
	*** This operations removes the element(s) from *isD2*.** **** This operation takes no time if focus is already at the position.**

S-II is a model that reflects blocking in an early phase.

However, empirical data of responses and fixations give evidence that while entering the 2nd of the two responses, people learn to coordinate manual action and visual attention to the subsequent pattern. For this reason, a third model (*S-III*) was build. The core mechanism of *S-III* resembles *S-II*. Participants` experience is integrated by assuming one single production rule for responding the 2nd blocking element and moving eyes one step further. The last pattern of a row (the 5th element) cannot be blocked. Both *S-II* and *S-III* perform this single pattern without blocking. The only difference is that *S-II* requires a further production rule (for move eyes) after the response whereas in *S-III*, manual response (key-press) and moving eyes one step further is implemented in only one single production rule.

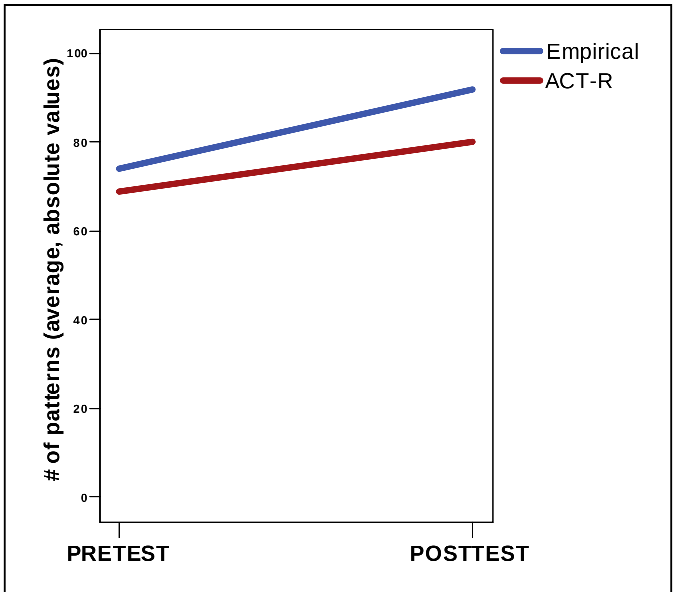

Figure 3: Empirical behavior vs. ACT-R predictions in version B of the D2-Drive-test (we refer to *S-I* in the pretest and to *S-II* in the posttest)

For the pretest, we argue that the first model (*S-I*) is a adequate candidate: *S-I* predicts on average 69 patterns per minute, compared to 74 patterns on average in the empirical data. We further assume that a blocking strategy is developed. First, the general mechanism of blocking is learned (*S-II*) and afterwards, in the posttest, *S-II* updated to *S-III* . *S-III* does a good job in predicting a similar increase (on average, it performs 80 patterns) as found in the data (with an average value of 92 patterns in the posttest).

Discussion

In this paper, three models were introduced: the first model turned out to be a good predictor of the behavior in the pretest. With the help of practice and training, blocking strategies emerged as confirmed by participants` verbal responses after the study. We implemented a blocking strategy that allows to block two elements. The implementation contained two steps: first the general mechanism for blocking was realized (*S-II*). In a next step (*S-III*), we considered the aspect of practice by merging actions together. In this article, blocking of more than two elements is not integrated. However, empirical data show that some subjects retain three and more patterns. This is part of future work on individual strategies.

Acknowledgments

We would like to thank all at MODYS, especially Dirk Schulze-Kissing, Marcus Heinath, Jeronimo Dzaack, Robert Lischke, Sandra Trösterer and Nele Pape. This work was funded by grants of VolkswagenStiftung (research group User Modeling in Dynamic Systems) and Deutsche Forschungsgemeinschaft (research training group GRK 1013 prometei).

References

Ackerman, P. L. (2005). Ability determinants of individual differences in skilled performance. To appear in Sternberg, R. J., & Pretz, J. (Eds.) Cognition and Intelligence: Identifying the Mechanisms of the Mind (pp. 142-159). NY: Cambridge University Press.

Byrne, M. D., & Anderson, J. R. (2001). Serial modules in parallel: The psychological refractory period and perfect time-sharing. Psychological Review, 108, 847-869.

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. Behavioral and Brain Sciences, 24, 87-185.

Daily, L. Z., Lovett, M. C., & Reder, L. M. (2001). Modeling individual differences in working memory performance: A source activation account in ACT-R. Cognitive Science 25, 315-353.

Dzaack, J., Kiefer, J. & Urbas, L. (2005). An approach towards multitasking in ACT-R/PM. In Proceedings of the 12th Annual ACT-R Workshop, Trieste, Italy.

Gunzelmann, G., & Anderson, J. R. (2002). Strategic differences in the coordination of different views of space. In W. D. Gray and C. D. Schunn (Eds.), Proceedings of the Twenty-Fourth Annual Conference of the Cognitive Science Society (pp. 387-392). Mahwah, NJ: Lawrence Erlbaum Associates.

Lovett, M. C. (2001). Modeling individual differences. Invited talk at the Post-graduate summer school, Coolfont, WV, July, 2001.

Lovett, M. C. (2005). A strategy-based interpretation of Stroop. Cognitive Science, 29, 493-524.

Newell, A. & Rosenbloom, P. (1981). Mechanisms of skill acquisition and the law of practice. In J. R. Anderson (Ed.) Cognitive skills and their acquisition. Hillsdale, NJ: Lawrence Erlbaum.

Salvucci, D. D. (2005). A multitasking general executive for compound continuous tasks. Cognitive Science, 29, 457-492.

Schunn, C. D., & Reder, L. M. (1998). Individual differences in strategy adaptivity. In D. L. Medin (Ed.), The psychology of learning and motivation.

Schunn, C. D., & Reder, L. M. (2001). Another source of individual differences: Strategy adaptivity to changing rates of success. Journal of Experimental Psychology: General, 130(1), 59-76.

Taatgen, N.A. (2005). Modeling parallelization and speed improvement in skill acquisition: from dual tasks to complex dynamic skills. Cognitive Science, 29, 421-455.