Pape, N., Dzaack, J., Leuchter, S., & Urbas, L. (2005). How to integrate time-duration estimation in ACT-R/PM. In *Proceedings of the 12th Annual ACT-R Workshop, Trieste, Italy.*
http://www.safety-critical.de/doc/actr2005.pdf

How to integrate time-duration estimation in ACT-R/PM

Nele Pape, Jeronimo Dzaack, Sandro Leuchter[1], Leon Urbas[2]
Research Group User Modeling in Dynamic Human-Machine-Systems
Technische Universität Berlin
Jebensstr. 1, J2-2
10623 Berlin, Germany
+49 (30) 314-72007

ABSTRACT

We conducted an experimental study to provide more detailed experimental data to derive a model of retrospective time-duration estimation for the ACT-R/PM architecture. The experimental setting, the implementation of the timer module and preliminary results are discussed.

INTRODUCTION

The estimation of time-duration in dynamic human-machine-systems is an essential requirement for system control (Schulze-Kissing et al., 2004). Time-duration estimations help us to stay tuned to the sequential occurrence of events in a complex environment and should be considered as important aspect in developing human-machine interaction. In some situations the processing of temporal information is the only means to gain information, for instance, operators of chemical plants differentiate between a feedback delay caused by a system innate latency, and an expanded feedback delay that is caused by abnormal system performance. That is, if the duration of a feedback delay exceeds the expected normal latency duration, an operator should suspect a malfunction.

Since James (1890, c.f. Hicks et al. 1976), psychological research distinguishes between prospective and retrospective duration estimation. In a prospective setting, the subject knows in advance that duration is important and presumably would like to put attention on time during the course of the task. For retrospective duration estimation experiments the subject is not informed that the duration of some part of the action might be of interest.

Relatively few studies have used a retrospective paradigm despite the fact that both human duration estimation methods (and possibly mixtures of them) play an important role in human-machine interaction (HMI), in particular in high workload conditions. A concise theory and computational implementation of retrospective duration estimation would be an important building block for the application of cognitive architectures like ACT-R/PM for HMI Engineering. Therefore, we derived a new approach of retrospective time-duration estimation from literature data and conducted an experiment to provide more empirical evidence. The newly developed algorithm has some promising properties and probably allows the integration of different theoretical accounts that try to explain human time-duration estimation. Our experimental implementation of a timer buffer extension for ACT-R/PM 5.0 efficiently encapsulates the complexity of the calculation from the modeler and thus eases the use of time estimation in human performance models for HMI Engineering.

THEORETICAL BACKGROUND

Several theoretical approaches are proposed to explain human time judgment characteristics. Different variables and processes are claimed to influence duration judgments. This results in a vast body of studies on the prospective

---

[1] Fraunhofer IITB, Fraunhoferstr.1, 76131 Karlsruhe, Germany.

[2] Corresponding author urbas@zmms.tu-berlin.de

paradigm and to a lower extent also for retrospective estimation of time intervals over the last century. The meta-analysis of Block and Zakay (1997) on prospective and retrospective duration judgments reveals two variables that influence the retrospective paradigm only - duration length and stimulus complexity - while others do not differentiate between the paradigms. However, the collected data does not give a consistent picture, probably because the stimuli and cognitive aspects of the tasks were as different as the theoretical accounts of the authors.

Table 1: Different theoretical accounts on retrospective time-duration estimation

Author	Duration judgement increases as a function of
Fraisse 1963	Number of Remembered Changes
Ornstein 1969	Storage Size
Block & Reed 1978, Zackay 1996	Number of encoded and available contextual changes
Poynter 1983	Degree of segmentation of events during the time period

Fraisse (1963) reviewed and summarized many investigations on time experience and relates both short and long duration experience to the ‘number of changes’ which occur in a given interval (c.f. Ornstein, 1969, p 38). Ornstein (1969) assumes that the duration of retrospective estimation is constructed from the contents of storage. His findings show that retrospective duration judgment increases as a function of the amount of stored and retrieved information or storage size allocated during the interval. Hicks and co-workers (Hicks, Miller, & Kinsbourne, 1976) ruminate that subjective time-duration estimation is assumed to increase as a function of subject’s attention to time. This attention results in the storage of subjective temporal units. In the retrospective paradigm, subjective temporal units are presumably not stored.

In the contextual-change model proposed by Block and Zakay (1996) retrospective time-duration estimation depends on retrieval of contextual information which is encoded in association with event information. The estimated time-duration is dependent on the amount of contextual changes stored in memory until a point of request.

ACTIVE AND IDLE TIME INTERVALLS

From a cognitive modeling perspective these different theoretical accounts might nicely converge if analyzed on an atomic production-level in a cognitive architecture like ACT-R/PM. The proposed variables like number of changes, complexity of processed stimuli, number of responses and contextual changes should find their counterpart in the number of meaningful productions that are fired during an interval.

This assumption leads directly to the following question: do all productions that are fired during the estimation interval provide equal ‘time cues’? What about time intervals, when users are waiting for a system response? What if they ‘idle’ because they know that the evolution of an anticipated system state takes some probably unknown time? An illustrating daily life example would be ringing the door bell and waiting for someone to open the door.

Unfortunately, many experiments that include idle periods utilize this variable to vary the previously mentioned variables. Thus, idle time and information processing demands are unintentionally varied at the same time and the effects of each of the variable cannot be separated post hoc. McClain (1983) for instance, conducted an experiment where subjects had to judge a time interval either in prospective or retrospective. The subjects had to encode wordlists presented in several intervals. In a fixed total task time of 120 seconds the subjects had to encode 15, 30 or 45 words in three different information-processing conditions. She found, that retrospective duration judgment increases with the amount of words encoded (see diamonds in figure 3). There is no data about response-times reported, however a simple analysis of the ‘shallow’ processing condition (attend, read word, decide if first letter is between A-M or N-Z, respond) reveals an average active period of about 1 second. That is, her subjects experience ‘idle’ times of 1.67 up to 7 seconds per item depending on the experimental condition.

Most of the data reported in retrospective duration judgment experiments had to be rejected for the analysis of idle versus active time. In most cases there is little or no information that allows a proper analysis of the subjects’ micro-level information processing. To address the question of idle time more concisely, we’ve decided to conduct an explorative study with clearly defined idle and active time intervals.

NEW EXPERIMENTAL DATA

A good candidate as a task for contextual changes and information processing demands is the D2-Drive test (Urbas et al. 2005a). The test was intended to measure residual resources of car drivers (Urbas et al. 2005b). The tasks consist of a series of uniform unit tasks that can be described easily within the ACT-R/PM architecture. Next, the experimental device gives rich information on subjects’ information processing without the need of gaze-equipment. Finally the complexity of the task is scalable but subjects still learn error free interaction with the task fast and easily.

Subjects had to fulfill this task with an active and an idle time interval. Thirty-one participants (twelve female, nineteen male) took part in the study. The sample covers students and graduates of the TU Berlin. The participants were naïve to the purpose of the experiment and the relevance to time. One participant was excluded because of the comparable large error rate.

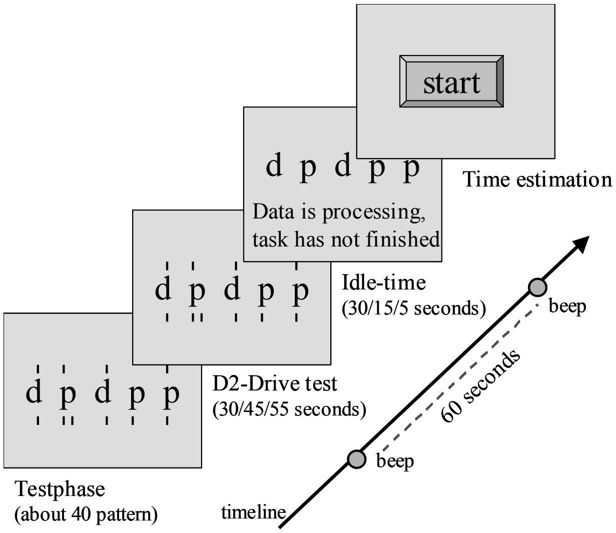

Figure 1: Experimental design

The course of the experiment is illustrated in figure 1. After a training-phase, the subjects were instructed to execute version A of the D2-Drive. In this version, the subjects have to decide, if the middle element of a five-pattern row is a "d" with two strokes and respond with a y/n-key-press. A screenshot is given in figure 2. They were instructed to do the task as fast and correct as possible. The D2-Drive test started with an acoustic signal. After 30, 45 or 55 seconds of active time (AT) the D2-Drive task stopped and the subjects were informed on screen that data is processing and that the task had not finished. This idle time period (IT) was chosen to lead to a total task time (TT) of 60 seconds and finished with the second acoustic signal.

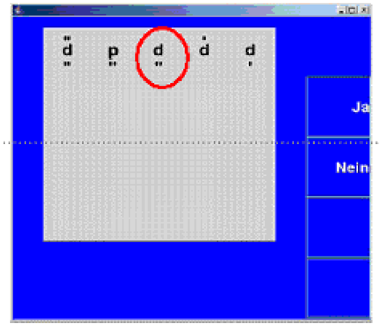

Figure 2: D2-Drive version A

Subsequently, the subjects were asked to reproduce the passed time between the two acoustic signals (see McKay, 1977 for a detailed discussion of the time reproduction method vs. verbal judgments and other measures of duration judgment). They were instructed to press a button to indicate the start of their estimation and then wait until they thought that the same amount of time had passed by. The end of the duration was indicated by a second press on the same button.

Results

For all three conditions, the duration estimations of the subjects are shorter than the real time task duration of 60 seconds. The mean of the duration estimates of all subjects is approximately 42 seconds with a standard deviation (SD) of 21 seconds (see table 2). The first condition with 30 seconds D2-Drive task and 30 seconds idle time was rated with a mean time-duration estimation of 45.6 seconds (SD 22.7). The condition with 45 seconds D2-Drive and 15 seconds idle time was estimated 49.5 seconds (SD 23.7). The average number of responses of the D2-Drive test was 39.1 (SD = 4.8) respectively 59.7 (SD = 7.8). Due to the large individual differences no significant difference between the group's duration estimation can be found. Furthermore, there is no evidence for a correlation between responses and duration estimation of different subjects.

This picture completely changes for the third condition (55 seconds D2-Drive and 5 seconds idle time). The mean duration estimate DE of 28 seconds is remarkably lower than in the 30 and 15 seconds idle-time conditions ($\chi^2 = 5.8$, $p < 0.055$). SD of judgments is reduced to 6.6. Furthermore we can see a clear correlation between the number of responses and the estimated time-duration ($R^2 = 0.72$).

Table 2: Results of the empirical study

Condition	Mean DE (duration estimation)	SD DE	Mean responses	SD responses
Average	41.6 sec	21	55.5	16.2
30/30	45.6 sec	22.7	39.1	4.8
45/15	49.5 sec	23.7	59.7	7.8
55/05	28 sec	6.6	70.9	14

INTEGRATIVE DISCUSSION OF EMPIRICAL BODY

The results are consistent with other findings concerning duration and active information processing (e.g. Block & Zakay, 1996)

1. Duration is almost always underestimated.

2. When idle time is low, the ratio between estimated and elapsed time gets better with the amount of information processing (measured as number of responses).

The idle time interval (30 or 15 seconds) seems to induce considerable variance on time-duration estimation and finally wipes out almost any correlation between information processing activity and retrospective duration estimation. Unfortunately this account can not be covered by McClain's (1983) data, because she does not provide any data about the dispersion of her retrospective judgments.

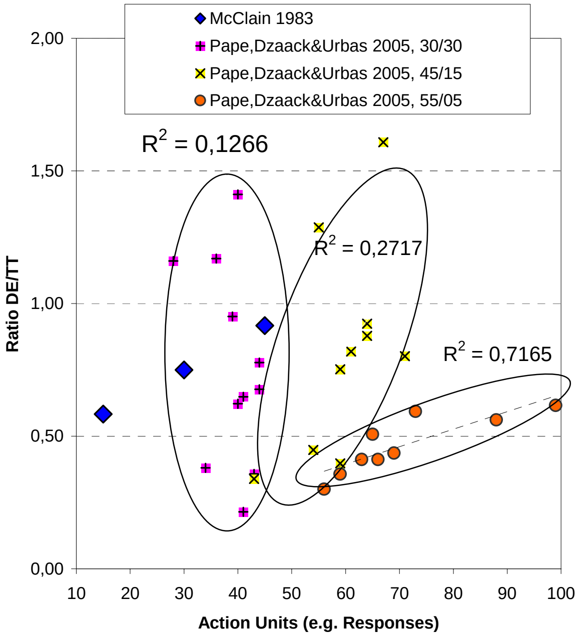

Figure 3 Influence of idle time (IT) and action units on duration estimation (DE)

THE ACT-R/PM TIMER-MODULE

Due to the importance of time-duration estimation for HMI, we developed a retrospective timer-module for ACT-R/PM (Anderson & Lebiere 1998, Anderson et al. 2004).

Table 3: Commands for the use of the timer-buffer in ACT-R/PM

Set a timer-reference for retrospective time-duration estimation

```
RHS:
+timer>
    isa   timer-reference
    mode  retro
    id     =id
```

Prepare an estimation to be available in the timer-buffer

```
RHS:
+timer>
    isa   timer-duration
    id    =id
```

Access an estimation that was made available,

```
LHS:
=timer>
    isa   timer-duration
    id    =id
    duration =duration
```

timer-failure if retrieval failed

```
=timer>
    isa     timer-failure
```

Delete a timer-reference (only for technical reasons)

```
RHS:
-timer>
    isa   timer-reference
    id     =id
```

The derived estimation algorithm is implemented as an ACT-R/PM buffer as described in (Leuchter, in Prep.). The buffer design pattern allows the modeler a uniform access to different processes and storages. By encapsulating most of the implementation details, this approach aims to enable the modeler to concentrate on the high-level aspects of the modeled task. Table 3 summarizes the three statements that are necessary to use the timer-buffer. To start with, a modeler needs to set at least one reference-point of type timer-duration. A reference-point might be considered as a pointer to an element of an episodic memory store. The retrospective time-duration estimation will be started by asking the timer buffer to retrieve a timer-duration. The estimated duration then will be made available as a time-duration chunk in the timer-buffer. For technical reasons (debugging and programming) the buffer provides a means to delete a timer-reference.

THE THEORY OF THE TIMER-MODULE

In the following we describe the theory of the retrospective time-duration estimation module. Roughly speaking, the module can be compared with a highly abstracted episodic memory store. New episodic reference points can be set to split up the passed time. In the retrospective approach this is made by the explicit setting of distinctive reference-points according to reality (i.e., distinctive actions will be hold in memory as landmarks and help to navigate through the passed time). The time between these reference-points and the time of retrieval is estimated by the timer-module algorithm (see figure 4).

$$DE = \begin{cases} (A + B \times (AT/TT)^{-0.5}) \quad \times AU \times TT & ;AT/TT < 0{,}9 \\ \\ C \quad \times AU \times TT & else \end{cases}$$

Figure 4: Duration estimation algorithm (DE: duration estimation, AU: action unit, AT: active time, TT: total time, A,B,C: parameters fitted to empiric evidence)

The timer-buffer operates on time cues that are provided by action units (AU). These action units may be interpreted as an integrative approach to represent contextual changes, information processing or storage size and complexity. In the current implementation of the buffer, any execution of a production instantiation that is not marked as 'idle'[3] is considered to be an AU and provides a time cue for the retrospective duration estimation.

By the means of *(firing-hook-fn)* whenever a production fires two basic calculations take place:

First, the active time (AT) to total time (TT) ratio is calculated. TT is measured individually for each timer-reference as the time elapsed between instantiation of the timer-reference and the firing time of the current production instance. If the measured real-time of a production fired is 3/2 times greater than its default action-time (i.e., the time a production needs to be completed) the calculated time is balanced with an idle-factor to adapt the according idle system waiting-time. Thus problems running a simulation in real-time environments can be adopted.

---

[3] This is accomplished by adding „idle-" as a prefix to the name of productions that are executed during idle time intervals.

Second, if the current production is member of the active unit set, than AU is increased. At this point of time any production contributes just its execution time.

The retrieval of the timer-duration uses the formula given in figure 4. We suggest to consider two different cases depending on the value of AT/TT-ratio. If the idle time contributes no more than 10% to TT, that is AT/TT > 0.9, than the duration estimation is calculated as a product of AU, TT and a fixed empirical factor C. Otherwise, the factor C is to be replaced by a more complex term that was derived by fitting to current empirical data. The first constant of this sum term (A) represents the influence of the time cues given by the AUs just the same way as C does for the time-duration estimation without idle-time. The second term of the sum, the constant B multiplied by the AT/TT-ratio to the power of -0.5, represents the influence of the corrective processes that are triggered by the experience of idle time intervals. This term can be interpreted as the probability of a positive correction of the time-duration estimation as a power function of AT/TT. Because AT/TT = (1 – IT/TT) this model predicts that the probability of the necessity of a positive correction of an information processing based time-duration estimation increases as a power function of idle time.

APPLICATION

To verify the implementation details, we added the logic for the experimental setting described earlier in this paper to an ACT-R/PM model of the interaction of humans with the D2-Drive (Dzaack, Kiefer, & Urbas 2005). This model now runs the D2-Drive test for a given time and then changes to an idle mode that emulates waiting. As in the experimental study, the total time was 60 seconds with different active and idle times (active/idle: 30/30, 45/15 and 55/5).

Results

As intended, the ACT-R/PM model with and without the integrated timer-module does not show any differences in performing the task. This was measured by the given responses of identifying the pattern (i.e., while doing the active task).

Running the ACT-R/PM model with the timer-module shows the retrospective estimation of time-duration as anticipated: the duration estimated by the ACT-R/PM model grows with the number of responses as well as the idle-time. The case with 55 seconds active time shows the shortest estimated time of the model. The case with 30 seconds idle-time shows the longest estimated time of the model. Thus we conclude that the timer-module is a suitable instrument for prospective time-duration estimation that efficiently allows to consider effects of retrospective time-duration estimation as found and described in the literature.

DISCUSSION & OUTLOOK

The empirical evidence for the proposed timer algorithm is rather shallow at this point of time. Follow-up experiments are already planed as well as some modifications of the underlying ACT-R/PM model to reflect individual differences in processing of the D2-Drive task.

Our model would predict that a long idle-time at the end of a trial or smaller idle-time periods that are spread all over the task make no difference. To prove this account we currently prepare a first study with an interleaved idle-time during the task. The active time/ to idle time ratio of the groups will be chosen to be identical to those of McClains experiments.

To investigate retrospective time-duration estimation and explicit setting of reference-points by humans we plan to utilize version B of the D2-Drive in a second study. In this study subjects have to fulfill five patterns arranged in a row, after completing this task the window changes to another view with a new pattern, and the task starts again (Dzaack, Kiefer, & Urbas 2005). We are curious, if this window-change can be seen as a distinctive event that provides its' own time cue for a retrospective duration estimation.

Furthermore we want to question the concept of action units. At the current state the timer-module each production a action unit are formed through the productions fired. An interesting approach would be to subsume successional combinations of productions to a single action unit.

Another issue of our research is the integration of prospective time-duration estimation and its experimental prove. Our investigations on prospective time-duration estimation (Schulze-Kissing et al. 2004) provided a huge body of empirical data that calls for being modeled. We believe that the combination of retrospective and prospective time-duration estimation might be a promising approach. Both methods should be integrated in cognitive architectures. This opens a wide range of new applications in the field of designing dynamic human-machine-systems.

ACKNOWLEDGEMENTS

This work was funded by grants of VolkswagenStiftung (research group user modeling), DFG (Research training group GRK 1013 prometei) and IBB PROFit (HMI Engineering for networked driving). Special thanks go to Dirk Schulze-Kissing for his valuable remarks on duration estimation.

REFERENCES

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

Anderson, J., Bothel, D., Byrne, M.D., Douglass, S., Lebiere, C., and Qin, Y (2004) An Integrated Theory of the Mind. Retrieved from http://act-r.psy.cmu.edu/-papers/403/IntegratedTheory.pdf.

Block, R. A., & Reed, M. A. (1978). Remembered duration: Evidence for a contextual-change hypothesis. *Journal-of-Experimental-Psychology*.

Block, R. A. & Zakay, D. (1996). Models of Psychological Time Revisited. In H. Helfrich (Eds.), Time and Mind. Proceedings of the International Symposium on Time and Mind held in Dec. 1994 at the University of Regensburg. (pp. 171-195). Seattle; Toronto; Göttingen; Bern: Hogrefe & Huber Publishers.

Block, R. A. & Zakay, D. (1997). Prospective and retrospective duration judgments: A meta-analytic review. *Psychonomic Bulletin & Review,* 4 (2), 184-197.

Brinkenkamp, R. (2001). Test d2 Aufmerksamkeits-Belastungs-Test. 9., überarbeitetet und neu normierte Auflage. Hogrefe Verlage. Bern, Schweiz.

McClain, L. (1983). Interval estimation: Effect of processing demands on prospective and retrospective reports. Perception and Psychophysics, 34(2) 185-189.

Hicks, R. E., Miller, G.W., & Kinsbourne, M. (1976) Prospective and retrospective judgments of time as a function of amount of information processed. *American Journal of Psychology*, 89, 719-730.

James, W. (1890). The principles of psychology, vol. 1. New York: Henry Holt.

McKay, T.D. (1977) Time estimation: Effects of attentional focus and a comparison of interval conditions. *Perceptual and Motor Skills,* 45, 584,586.

Dzaack, J., Kiefer, J., & Urbas, L. (2005) An approach towards multitasking in ACT-R/PM. This volume. ACT-R Workshop 2005, Trieste.

Leuchter, S. (in preparation) Software Engineering Methods for User Modelling in dynamic Human-Machine Systems. Ph.D Thesis, TU Berlin.

Ornstein, R.E. (1969) *On the experience of time. Middlesex,* England: Penguin.

Poynter, W. D. (1983). Duration Judgement and the Segmentation of Experience. *Memory & Cognition*, *11* 77-82.

Schulze-Kissing, D., van der Meer, E. & Urbas, L. (2004). A Psychological Analysis of Temporal Errors in Human-Machine-Systems. Proceedings of the IFAC Symposium: Analysis, Design and Evaluation of Human-Machine-Systems. Atlanta, USA, 07.-09. September 2004.

Urbas, L., Schulze-Kissing, D., Leuchter, S., Dzaack, J., Kiefer, J., & Heinath, M. (2005a) *Programm-beschreibung D2-Drive-Aufmerksamkeitstest [Manual for D2-Drive Test of Attention]*. Berlin: ZMMS

Urbas, L., Schulze-Kissing, D., Leuchter, S., Kiefer, J. Dzaack, J., & Heinath, M. (2005b) Erfassung Residualer Ressourcen Profile zur Modellierung von FAS-Nutzung unter Fahrbedingungen [Acquisition of residual ressource profiles for modelling of information system use while driving]. In Proceedings of TEAP 2005. 4.-6. April 2005. Regensburg