

University of Basel
Department of Psychology
Division of Cognitive Psychology and Methodology

Superiority of Paper as Text Presentation Medium for
Effective and Efficient Learning –
Is it Just an Illusion?

Master's Thesis
submitted by

Katarzyna Iwona Hongler

in March 2015

supervised by
Silvia Heinz, M.Sc.
Dr. Alexandre N. Tuch
Prof. Dr. phil. Klaus Opwis

Table of contents

Abstract	3
Introduction	4
Preference for Printed Texts	4
Effectiveness of Learning from Texts on Different Media.....	5
Efficiency of Learning from Texts on Different Media	7
Goals of the Study	9
Method	9
Design	9
Participants	10
Materials and Apparatus	10
Procedure	12
Results	14
Self-reported Media Preferences and Learning Habits	14
Effectiveness and Efficiency of Learning	18
Affect during Learning	20
Discussion.....	22
References	29
Appendices	34
Appendix A: Online Pre-survey	34
Appendix B: Learning Texts	39
Appendix C: NASA TLX Questionnaire	45
Appendix D: SAM Questionnaire	47
Appendix E: Online Quiz and Post-survey	48
Appendix F: Written Instruction for the Lab Experiment	53
Appendix G: Written Instruction for the Quiz	54

Abstract

Despite the ubiquity of digital devices, printed texts are still students preferred medium when learning for academic exams, as they are expected to warrant superior effectiveness and efficiency. To test this prevalent expectation, undergraduate psychology students learned from two academic texts displayed on paper, desktop or tablet. Their objective learning effectiveness was examined with a quiz while their objective efficiency was expressed as learning duration in minutes. Additionally subjectively perceived effectiveness (estimated correct answers in the quiz) as well as efficiency (estimated learning duration and NASA TLX mean unweighted workload score) was recorded. The participants rated their affective state during learning on the SAM scales and the NASA TLX frustration subscale. The reading medium had no effect on the objective and subjective effectiveness and efficiency or affect during learning. Thus the reported expectations for superiority of paper regarding learning effectiveness and efficiency were not confirmed. The findings have implications for students' learning habits and promotion of paperless university initiatives.

Superiority of Paper as Text Presentation Medium for Efficient and Effective Learning – Is it Just an Illusion?

Reading texts constitutes one of the main learning activities for students in academic institutions. Textbooks, scholarly research articles, scripts, and course handouts, all printed on paper, have been the prevailing materials for accessing necessary information to study and learn. Advances in computer technology made digital devices not only widely affordable and ubiquitous but also mobile enough to shift the importance away from paper as a medium of written information towards digitally presented texts. In spring 2014 98% of Swiss population younger than 30 years used the internet daily (Bundesamt für Statistik, 2014). It is therefore not surprising that many universities have been following the call for ecological sustainability (Flamm, 2014; Sustainable Development Solutions Network, 2014). They have been launching initiatives to go paperless and hereby to ease the environmental burden caused by excessive paper usage in the academia (e.g. Maastricht University Green Office, 2014; Drake University, 2014).

Preference for Printed Texts

A recent survey (Preller, Zwahlen, & Grösser, 2014) of 730 students at the University of Applied Sciences in Bern, Switzerland, revealed, however, that only 40% of the students would accept strictly electronically distributed supplementary learning materials, and just 25% would be willing to read textbooks exclusively on digital devices. Those students reported that they would print the digitized texts and study for exams on print, though. This finding stresses the predominant preference for printed learning materials. One of the reasons given by the students for such attitudes was the subjective notion that studying paperbound was associated with more efficient and better learning outcomes (Preller et al., 2014). Another survey conducted on 162 academics (students and researchers) in Australia and Germany (Franze, Marriott, & Wybrow, 2014) revealed that 56% of the participants preferred reading academic texts on paper rather than on a digital medium. The main reason (44%) indicated

was the ease of direct text annotation, while only 9% of the respondents considered printed texts to be easier to comprehend. Chou (2012) found in a qualitative study of 5 graduate students, who studied in English as their second language, that they preferred studying from printed texts when they were preparing for a course. They would, however, settle for a digital text, if they did not have to take notes, needed time-accurate information or decided just to skim a text. Overall, whenever the students wanted to read passages carefully, they preferred them printed on paper.

The widespread preference for studying from printed texts rather than their digital versions could be formed by previous experiences, as shown for other preferences in the past (e.g. Hoeffler & Ariely, 1999; Hoeffler, Ariely, & West, 2006). Therefore it might be expected that the objective learning outcomes and learning durations achieved using different media of text presentation would reflect the superiority of learning on paper.

Effectiveness of Learning from Texts on Different Media

Noyes and Garland (2008) reviewed empirical literature on reading from paper vs. screen. In the early studies, which were conducted before 1992, the evaluation of reading speed, reading accuracy, and reading comprehension was found to generally favour the print conditions when compared with desktop screen conditions. However, later studies yielded inconsistent results. The authors hypothesized that the differences might have resulted from advances in digital technology, the different demands of learning tasks given in the experiments, and the diverse measures of learning outcome (Noyes & Garland, 2008). Recent studies continue to report inconsistent results and their findings are difficult to compare and to generalize as they use diverse methodological approaches. Mangen, Walgermo, and Brønnick (2013) investigated differences in reading comprehension of 72 Norwegian high-school students. The participants read texts either on paper or on a computer screen. Each text was followed by a comprehension test, the time was limited and the students were allowed to consult the text while answering the test questions. Students, who read on

paper, scored significantly better than students, who read the texts on a screen. However, the ease of finding the appropriate text passages to answer the questions might have been a confounding variable, possibly favouring the printed texts.

Another research group (Rockinson-Szapkiw, Courduff, Carter, & Bennett, 2013) gave a convenience sample of 538 American undergraduate and graduate students a choice of learning either from a conventional textbook or an e-book for the course of one semester in different subjects of study. Four times more students opted for the conventional printed textbooks than for their digital versions. The e-books were accessed on laptops (49.1%), tablets (23.6%), e-readers (16%), and desktop computers (7.5%) and were pooled in one group. There was no significant difference in the course grades or in the subjectively perceived learning outcomes between the paperbound and the e-book groups.

Margolin, Driscoll, Toland, and Kegler (2013) investigated three distinct reading conditions (paper, desktop, e-reader) on a group of 90 undergraduate psychology students. The participants read five narrative and five expository texts on not study related topics, each comprising of about 540 words. The students were allowed to read at their own pace, to annotate, and to underline the texts but they could not consult the texts while answering the test questions. Each of the ten text passages was immediately followed by questions probing for the comprehension of the antecedent text. No significant differences in the comprehension were found between the text types and the presentation media.

The above mentioned research varies considerably in many ways, esp. in the choice of the sample, the experimental or quasi-experimental approach, the diversity of digital devices used, the selection of topics and texts' lengths, the time point and way of the evaluation of learning, the permission of individual reading habits, and the implementation of self-reports versus objective measurements. Hence, the evidence so far is difficult to compare and cannot be unequivocally generalized to students' learning for academic exams.

Efficiency of Learning from Texts on Different Media

Noyes, Garland, and Robbins (2004) measured the subjective efficiency using the NASA Task Load Index during a reading experiment with 30 undergraduate students. They found no significant differences in the workload scores between the paper and computer reading conditions and no differences in the learning outcomes.

None of the cited studies attended to the issue of learning duration. They implemented maximum time limits - the same for all investigated reading media per study - but they did not monitor the actual learning durations. One of the best predictors for retained information, however, is the amount of time spent learning it. The fact has been well documented since Ebbinghaus's (1885) famous discoveries related to the human memory and confirmed for performance and learning outcomes in diverse areas of expertise (Ericsson, Krampe, & Tesch-Römer, 1993; Fredrick & Walberg, 1980). Therefore, time might turn out to be a confounding variable in the above studies, esp. considering the earlier findings that reading on paper was faster than on screen (Noyes & Garland, 2008). The preference for paper as reading and learning medium was explained with its perceived and expected higher learning efficiency (Preller et al., 2014), which highlights the need for research on possible differences in learning durations on different media.

As shown further above, the perceived effectiveness of learning did not correspond to the observed effectiveness; therefore a similar disassociation of subjective and objective efficiency might also be possible and needs to be explored. Wilson, Lisle, Kraft, and Wetzel (1989) demonstrated that "...people sometimes assimilate the value of a stimulus to their expectations, even when the objective value of the stimulus is discrepant from these expectations". Correspondingly the preference for printed media persists despite an absence of its unambiguous superiority. Then it might be possible that in this case the subjective judgments, attitudes or beliefs of learners inform and guide the choices of learning media, as was shown for learning strategies (Dunlosky, Mueller, & Tauber, 2014; Metcalfe, 2009;

Metcalfe & Finn, 2008). Affect in general plays an important role in the construction of preferences and informs our judgements (Clore & Huntsinger, 2009; Glasman & Arbabzadeh, 2006; Peters, 2006) and could therefore shape the manifested preferences for printed texts.

Sackett, Meyvis, Nelson, Converse, and Sackett (2010) demonstrated how affect was associated with subjective time judgments. Time durations are perceived as shorter for positive, satisfactory experiences compared to dull, boring experiences, which are perceived as longer. Interestingly there are two ways to estimate durations while performing a distinct primary task. An individual can be asked to estimate task's duration only upon completion of the task, which is called *retrospective* time estimation. On the other hand, informing a person before starting a primary task that time estimation will be required at the end of it, is called *prospective* time estimation. Liikkanen and Gómez (2013) hypothesized an interrelation of affect and the method of time estimation. They indicated a need for research in this matter, as the direction of the effect was unknown and it could impact studies, which simultaneously explore subjective experiences of time duration and affective state.

Table 1

Aspects of learning from texts and their measures as used in the present study. Adapted from Hornbæk (2006).

Aspects of learning from texts	Measures	
	Objective	Subjective
Effectiveness (Outcomes)	quiz score	estimated quiz score
Efficiency	duration in min	estimated duration in min, workload (NASA TLX mean unweighted workload score)
Attitudes		affective state (SAM & NASA TLX frustration subscale)

Goals of the Study

The current study focuses on learning from academic texts as presented on paper, desktop and tablet. It explores, if the preferences for learning media and learning habits, reported by the sample of university students correspond to the findings reported in the literature. The main goal then is to investigate possible differences in their objective and subjective learning effectiveness and efficiency on the three reading media, as they may inform the reported preferences. The affective state during learning on the three different media may manifest the attitudes towards their use and is therefore concurrently monitored. The learning texts are chosen to be relevant to the students' subject of study to ensure interest in the concepts to be learned and to boost the ecological validity of the findings. This study is limited to learning by reading solely, and supportive editing of texts is explicitly excluded to prevent confounding effects. As a second line of investigation the impact of the method of time estimation on the affective state, which was suggested but not reported so far, is examined. Because of the equivocal and in part even lacking evidence in literature, no directed hypotheses are formulated. Table 1 lists the aspects of learning, which are screened in the present study, and the applied measures.

Method

Design

The present study used a 3 x 2 mixed design, with media of text presentation (paper, desktop, tablet) as between-subject factor and time judgment paradigm (retrospective after the first text = text 1; prospective after the second text = text 2) as within-subject factor. The dependent variables were measured and estimated learning durations in minutes; the measured and estimated quiz scores; the ratings on the SAM and the NASA TLX scales. In addition media preferences and learning habits were determined via self-report before the experiment.

Participants

Seventy-two undergraduate psychology students (56 female, 16 male) from University Basel volunteered to participate in the study. The participants ranged in age from 19 to 43 years, with a mean age of 24 years ($SD = 4.6$). None of the participants reported handicapping vision problems. To ensure a sufficient knowledge of terminology only students, who were enrolled at least in the third undergraduate semester, were eligible for participation. All students got credits for participating in the study. Additionally, there was an incentive of CHF-50-vouchers held out for the seven students, who would score best in the quiz at the end of the study. The participants were randomly assigned to one of the three experimental conditions, however care was taken for similar gender distribution in all conditions (paper: 5 male; tablet: 5 male; desktop: 6 male).

Materials and Apparatus

A pre-survey (see Appendix A) investigated participants' preferred media for learning from texts, the purposes of use of desktops, laptops, and tablets, their subjective abilities and frequencies of annotating, underlining, and marking of texts digitally and on paper. The participants also estimated their effectiveness and efficiency of learning from digitally presented texts versus texts printed on paper.

The learning texts contained of a slightly modified reprint of a chapter on cognitive effort in decision making from a German textbook “Die Psychologie der Entscheidung. Eine Einführung.” by Jungermann, Pfister, and Fischer (2010; see Appendix B). As the topic is not a part of the compulsory curriculum, it could be assumed that the participating students were not familiar with the subject. The book chapter was divided into two parts, which could stand for themselves. The texts were presented digitally as two separate PDF files or as two printed texts on paper. The text A contained of 1638 words, text B of 983 words. They were presented to the participants in fully counterbalanced order to control for length and order effects.

A German version of NASA Task Load Index (NASA TLX) with 6 subscales (mental, physical and temporal demands, performance, effort, frustration; Hart & Staveland, 1988) was implemented to rate the participants' perceived workload during both texts (see Appendix C). The subscales allowed ratings from zero to 20 with the higher numbers indicating higher demands. The ratings on the six NASA TLX subscales were analysed as raw data without the originally described weighing procedure (Hart & Staveland, 1988), because such approach has been shown to be shorter and more convenient for the participants while at the same time being as reliable as the original method (Nygren, 1991; Pfendler, 1990). The NASA TLX mean unweighted workload score was calculated for every participant by averaging the ratings on the six subscales. The frustration subscale, which represents a characteristic of an individual (Geddie et al., 2001), was additionally considered separately to operationalize the affect of frustration.

A German version of the 9-point Self-Assessment-Manikin Scale (SAM; Bradley & Lang, 1994) was used to assess participants' affective state (valence, arousal, and dominance) during both texts (see Appendix D). Low ratings indicated dissatisfaction on the valence subscale, boredom/ calmness on the arousal subscale and loss of control on the dominance subscale. High ratings indicated satisfaction on the valence subscale, interest/ excitement on the arousal subscale and perception of control on the dominance subscale.

The quiz, which probed recall as well as comprehension and transfer of the content learned in the lab comprised of 20 questions, 10 on each of the texts (see Appendix E). There was one open question on each text, requiring answers in just a few words. All other questions offered multiple choices with one or more correct answers. The quiz questions were constructed by the author with the exception of two questions, which were translations of questions published by Gregory and Kaprielian (2014). The questions were presented in fully randomised order.

A post-survey (see Appendix E) queried the participants about the perceived difficulty of texts' contents and quiz questions, and their motivation as compared to a regular exam. Additionally the participants were given the opportunity to comment, if the learning experience in the lab had corresponded with their usual learning habits or if they had missed something, and if so, what.

In the desktop condition the participants read the learning texts on a 21.5-inch Apple iMac with OS X 10.8.5 operating system, using Preview 6.0.1. The tablet condition was run on an iPad 2 with iOS 7.1.2, using Adobe Reader 11.3.1. In both digital conditions participants were allowed to individually adjust the height of their sitting position, the screen inclination and the zoom factor of the displayed texts for their convenience. For the paper condition the PDF texts were printed double-sided black on white DIN A4 sheets of paper in 12-point Times New Roman font.

The time was measured using a LabTimer App by Jeremy Murray run on an iPad mini with iOS 8.0, which enabled several simultaneous time measurements, while staying unnoticeable to the participants.

Procedure

At least 24 hours prior to the lab experiment all participants answered a pre-survey mounted on *SurveyMonkey* (a commercial survey instrument) as a first part of the study (see Figure 1). The second part took place in a lab at the Department of Psychology at the University of Basel. Participants were tested either individually or in groups of two; each participant had a separate cubicle. The participants were randomly assigned to one of the three experimental conditions. All participants were required to switch off their cell phones and to dispose any watch, they might have been wearing, in their bags. The clock was disabled in the settings in the desktop condition or hidden behind an opaque white tape in the tablet condition, to prevent time readings by the participants. Participants were provided with a written instruction (see Appendix F) and all signed a written consent. They had a maximum of

90 minutes for the whole laboratory procedure, which had been confirmed as sufficient duration in a pilot study with a convenience sample of four. The participants were neither allowed to annotate, to mark, nor to underline the texts. They were asked to read and re-read the texts in total or in parts as long and as often as they considered it necessary for them to understand and to learn the contents. The participants were informed that they would take a quiz the following day and that the seven participants, who would score best, would get a CHF-50-voucher each, at the end of the study. The participants were left unaware that their learning would be timed.

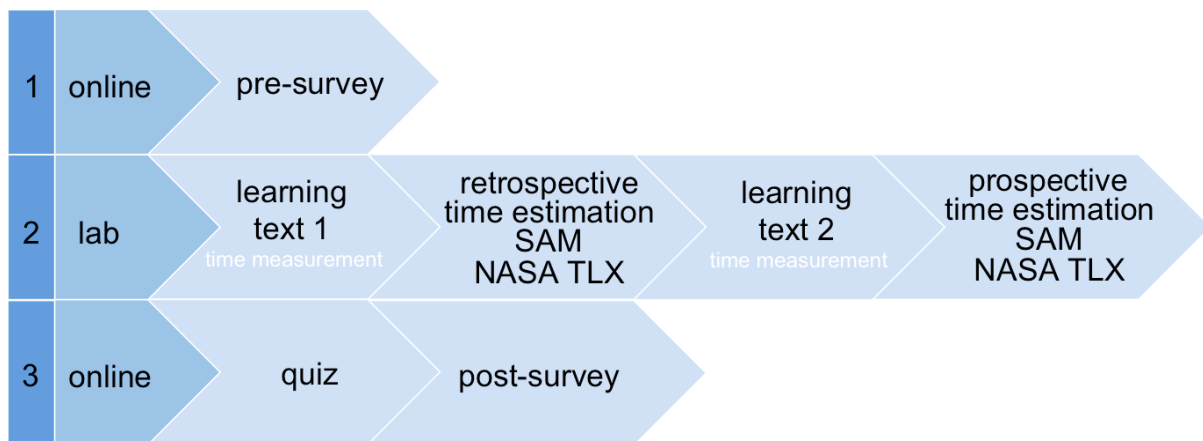


Figure 1. Sequence of the experimental procedure; time measurement by test administrator.

The first text was handed out to the participants or opened as a digital file, and the time measurement started. The timing was stopped once the participant indicated being finished with the first text. The text was then withdrawn or the PDF file closed, and a paper questionnaire handed out. The participants had to estimate how long they learned from the first text (retrospective estimation), fill out the SAM and the NASA TLX scales, and to indicate, if they were familiar with the content of the text. Once the participants finished answering the questionnaire, they were provided with an additional short written instruction informing them that when they would finish learning from the second text, they would have to

estimate the learning duration. Then the second text was presented according to the participants' assigned condition and the time measurement was started again. When the participants indicated having finished learning from the second text, the timing was stopped and the participants had to fill out the same questionnaire as after the first text, including time estimation for learning from the second text (prospective estimation).

The texts A and B were handed out to the participants in a fully counterbalanced order to control for possible order and text length effects on the learning outcome. Thus both texts A and B were equally used as first (text 1) and second text (text 2). Before leaving the lab the participants were handed out detailed written information about the quiz on the following day (see Appendix G).

The last part of the study contained of an online quiz, followed by a short post-survey mounted on *SurveyMonkey*. Twenty-four hours after the lab session participants received a link via email to access the quiz and the adjoining survey. The participants had to answer the quiz within 12 hours after the email, hereby all participants answered the quiz within 24 and 36 hours after learning. Once they opened the link, they had 45 minutes to answer all questions. Following the quiz the participants were required to estimate the number of questions, they expected to have correctly answered, and to fill out the post-survey.

Results

The statistical analysis of the collected data was conducted using IBM SPSS Statistics v. 20.0.0 software. An alpha level of .05 was applied to all statistical tests in the study.

Self-reported Media Preferences and Learning Habits

In the pre-survey vast majority ($n = 55$) of the 72 participants reported that they preferred texts printed on paper over digitally presented texts when learning for exams and only three participants indicated a preference for learning on a desktop screen. Table 2 depicts the distribution of preferred learning media between the three experimental conditions.

Table 2

Number of participants, who reported a preference for the specified medium of text presentation in each experimental condition.

Medium	Condition		
	Paper	Tablet	Desktop
Paper	21	18	16
Tablet	2	3	3
Laptop	1	0	5
Desktop	0	3	0

All 72 participants were familiar with digital devices and used at least one type thereof regularly. Laptops were most often and desktops the least frequently used digital media for learning purposes, while tablets were overall the least used digital device (Table 3).

Table 3

Frequency of reported purposes of use of digital devices.

Device	Purpose of use		
	Learning	Generally	Never
Tablet	19	10	43
Laptop	47	22	3
Desktop	11	31	30

The self-reports further revealed that paper was perceived as by far the most effective and the most efficient medium for learning from texts, followed by texts presented on laptop screens (Figure 2). At the same time the participants indicated the highest dissatisfaction with their learning efficacy and efficiency when learning on laptops (Figure 2).

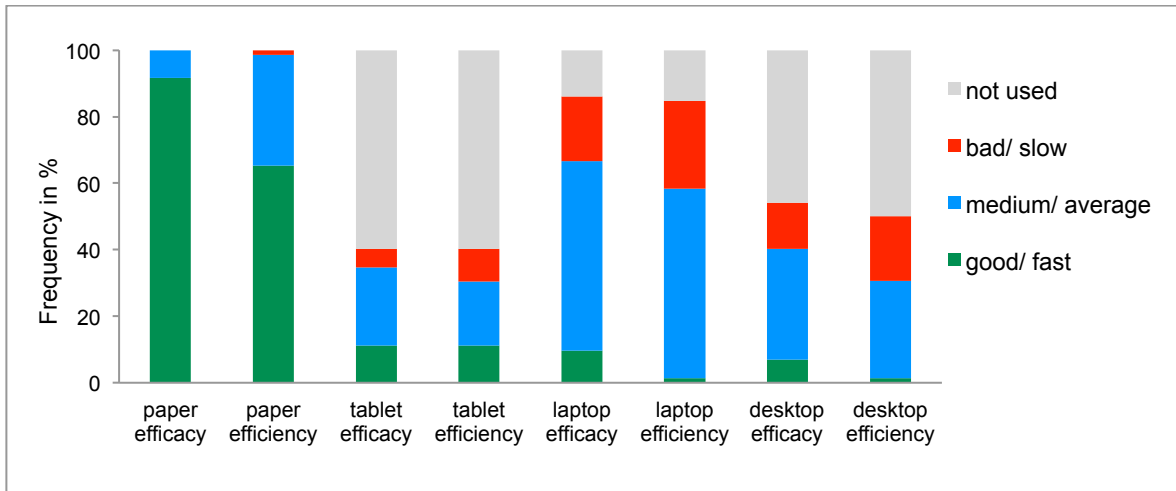


Figure 2. Self-reported subjective learning efficacy and efficiency from texts presented on paper, tablets, laptops and desktops ($N = 72$).

The participants reported different adoption of learning habits depending on the medium of text presentation. While all participants underlined or marked text passages on print, 38% never did so while reading texts digitally (Figure 3). A similar finding became apparent for annotating texts; only 1% never annotated printed texts but 42% never placed annotations on digital texts (Figure 3).

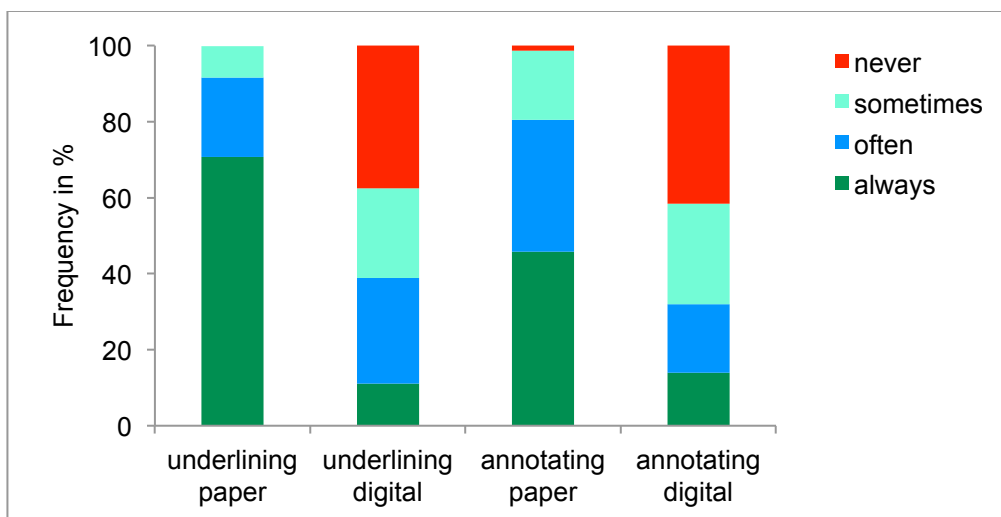


Figure 3. Self-reported habits of underlining and annotating printed and digitally presented texts ($N = 72$).

A minority of 11 % of participants assessed their abilities to mark, underline, or annotate PDF files digitally as rather advanced, 40 % reported having average, and 18% rather poor skills in this area, while 31% never used that feature at all.

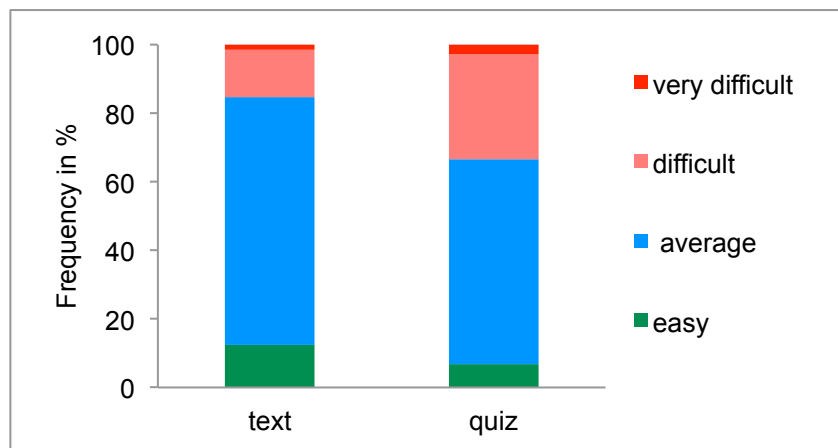


Figure 7. Perceived difficulty of texts and quiz questions; both reported after the quiz. $N = 72$.

In a post-survey right after the quiz the participants were asked to rate the perceived difficulty of the contents presented in the learning texts as well as the difficulty of the quiz questions. The texts were rated as easy to average by 85% of the participants while the quiz questions were perceived as easy to average by 67% of the participants (Figure 7). Most of the participants (74%) indicated lower motivation to learn during the experiment than for a regular exam, while 6% reported even higher motivation.

The learning in the lab was experienced as usual by 25% of the participants in the paper condition and by 21% in the tablet and desktop condition respectively. The remaining participants reported that they missed something compared with their usual learning experience. All of those 56 participants took the chance and commented in writing on what they lacked during reading and learning in the lab. For 32 individuals underlining and marking important passages in the texts would have been helpful and 31 participants would have liked to take notes or annotate within the texts. Fifteen participants missed repetitions of

learned concepts, distributed over longer periods of time and 13 explicitly mentioned their need for paper, any kind of pens or highlighters. Eleven individuals would usually write their own summaries, while eight would draw mind maps and charts. Other reasons for dissatisfaction, e.g. no coffee, too little breaks, no reading aloud, were brought up by just one to two participants each.

Effectiveness and Efficiency of Learning

The participants needed between 16 and 79 minutes total in order to learn from both texts and they answered between 4 and 15 questions correctly in the quiz, which excludes ceiling and floor effects. Seven participants (paper: $n = 3$; tablet: $n = 1$; desktop: $n = 3$) reported that they had heard about the concepts described in the learning texts before. They did not consider themselves knowledgeable in the specific topics, though.

To determine any differences in the objective learning effectiveness and efficiency between the three learning media the quiz scores and the total measured learning durations were analysed. In order to additionally explore effects of learning media on the subjective effectiveness and efficiency the estimated quiz scores and the total estimated learning durations were analysed. All of these statistical tests were conducted using one-way independent ANOVA. The total measured learning duration was calculated by summarizing the durations of learning from the first and second text. Correspondingly the total estimated duration is the sum of estimates of learning durations given after the first and second text. The means and standard deviations of the above measures as well as the test statistics are listed in Table 4.

The data analyses revealed no significant differences between the three reading conditions, neither for the quiz scores, nor for the total learning durations. Also the estimations of quiz scores and of total learning durations did not significantly differ between the three experimental groups.

Table 4

Means and standard deviations of total learning durations measured and estimated, of the quiz scores and their estimates in the three reading conditions (paper, tablet, desktop) along with the test statistics using one-way independent ANOVA.

		Paper		Tablet		Desktop		Statistics		
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>F</i> (2,69)	<i>p</i>	η^2
Quiz score	measured	10.58	2.21	9.17	3.35	10.13	2.42	1.72	.19	.047
	estimated	11.54	2.95	11.50	3.10	11.17	3.64	0.10	.91	.003
Duration in min	measured	38.38	13.59	39.08	15.53	39.25	11.14	0.03	.97	.001
	estimated	43.58	16.84	45.79	15.87	45.42	16.00	0.13	.88	.004

Note. For each condition $n = 24$

In order to examine the precision of the estimations of learning outcomes and durations, Pearson correlation coefficients were computed for all participants. The estimated quiz scores were significantly and moderately related to the measured quiz scores, $r(70) = .31$, $p = .004$ (one-tailed). There was a significant and strong relationship between the total estimated learning durations and the total measured learning durations, $r(70) = .70$, $p < .001$ (one-tailed). The obtained quiz scores were moderately and significantly correlated with the total measured learning durations, $r(70) = .40$, $p < .001$ (one-tailed), however the estimated quiz scores did not significantly correlate with the total estimated learning durations, $r(70) = .08$, $p = .24$ (one-tailed).

The NASA TLX mean unweighted workload score was used as an additional subjective measure of learning efficiency (see Figure 5 for means and confidence intervals). It was analysed with two-way 3 x 2 mixed ANOVA. There was no significant effect of reading condition or time estimation method on the perceived workload; neither the interaction of both factors had a significant effect (Table 5).

The medium of text presentation showed no significant effects on participants' learning effectiveness and learning efficiency, neither for the objective nor for the subjective measures thereof. The participants' estimations of quiz scores and learning durations were moderately resp. very accurate. The findings additionally indicate that the longer learning durations were associated with higher quiz scores, which confirms the importance of time spent on task for the learning outcomes (Ebbinghaus, 1885; Fredrick & Walberg, 1980).

Affect during Learning

The means and confidence intervals of ratings on the SAM subscales and the NASA TLX frustration subscale are depicted in Figures 6 and 5 respectively. In order to explore possible effects of the reading media (between subject variable: reading condition paper, desktop, tablet) and the time estimation method (within-subject variable: text 1 = retrospective estimation for the first text, text 2 = prospective estimation for the second text) on the readers' affective state the data were analysed using two-way 3 x 2 mixed ANOVA.

Table 5

Test statistics using two-way mixed 3 x 2 ANOVA for the ratings on the SAM and the NASA TLX frustration subscales and for NASA TLX mean unweighted workload scores in the three reading conditions (paper, tablet, desktop) after each learning text.

	Subscale	Condition		Text		Condition x Text	
		<i>F</i> (2,69)	<i>p</i>	<i>F</i> (1,69)	<i>p</i>	<i>F</i> (2,69)	<i>p</i>
SAM	valence	2.97	.06	3.18	.08	2.08	.13
	arousal	3.09	.05	6.09	.02	0.47	.63
	dominance	1.22	.30	4.23	.04	1.86	.16
NASA TLX	frustration	0.18	.84	4.62	.04	1.92	.15
	mean workload	0.28	.76	2.61	.11	2.20	.12

Note. *p* < .05 are in bold face.

There was no significant effect of reading condition on any of the SAM or NASA TLX frustration subscales (Table 5).

There was, however, a significant effect of time estimation method on ratings made on the SAM arousal and dominance subscales (Table 5), which were lower for text 2 than for text 1 (Figure 6). A significant effect of time estimation method was also observed on the NASA TLX frustration subscale, where participants gave significantly (Table 5) higher ratings for text 2 than text 1 (Figure 5).

The interaction effect of reading condition with the method of time estimation was not significant on any of the measures of affect.

Overall the medium of text presentation showed no significant effect on the participants' affective state. However ratings of affect differed significantly between the time estimation methods. The participants felt more interested, more in control and less frustrated in the retrospective time estimation condition than in the prospective time estimation condition.

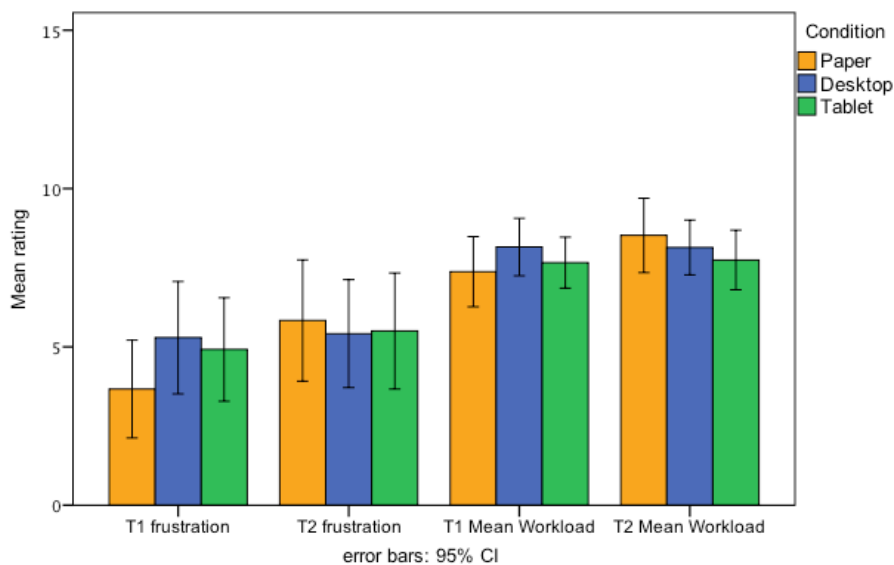


Figure 5. Mean ratings on the NASA TLX frustration subscale and NASA TLX mean unweighted workload score for the three reading conditions and both time estimation conditions text 1 (T1) and text 2 (T2). For each reading condition $n = 24$.

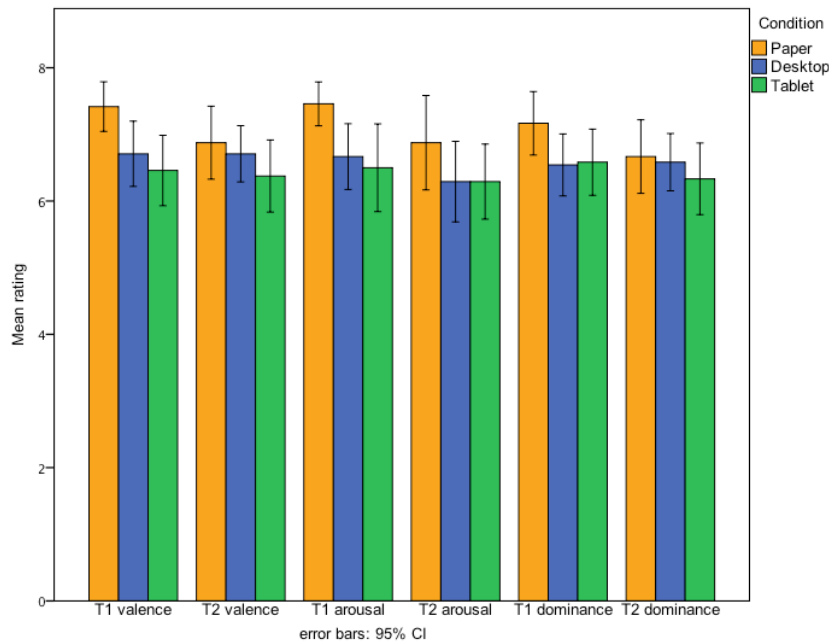


Figure 6. Mean ratings on the SAM subscales (valence, arousal, and dominance) for the three reading conditions and both time estimation conditions text 1 (T1) and text 2 (T2). For each reading condition $n = 24$.

Discussion

The present study confirmed the recent reports (Chou, 2012; Franze et al., 2014; Preller et al., 2014; Rockinson-Szapkiw et al., 2013) that given a choice, students in the academia universally prefer learning from texts printed on paper rather than their digitized versions. The observed results, however, did not reflect the participants' expectations that reading on paper would lead to better learning outcomes and be more efficient than reading texts on digital media.

Participants in all three experimental reading conditions - paper, desktop, tablet - scored comparably in the quiz and also their estimated numbers of correctly answered questions did not differ. Ceiling as well as floor effects of the quiz questions could be excluded as a reason for this finding. Besides 85% of participants considered the learning texts as easy to average and 67% reported the same about the test questions. However almost three quarters of participants felt less motivated than for usual exams and this despite contents

relevant for the subject of study and despite monetary incentives. This may be due to the still limited ecological validity of the current investigation. The habitual learning environments of the participating students rarely if ever resemble laboratory settings. Furthermore learning is usually self-regulated, distributed over time, probably more goal-directed and it contains repetitions, as indicated by about a fifth of the participants.

Hence, it is possible that the prevalent low motivation levelled out any otherwise existing differences in learning on different media. Implementation of texts, which would have been more relevant for the actual exams might have had enhanced the learning motivation but would possibly at the same time have allowed use of additional learning materials outside the lab. Such confounding variable would be difficult to control but could be included in future studies to extend the current findings. Nevertheless the present results align with those detailed for academic settings by other researchers (Margolin et al., 2013; Noyes et al., 2004; Rockinson-Szapkiw et al., 2013) and indicate that the medium of text presentation does not influence the learning outcomes of university students.

The same pattern pertains to learning efficiency. Contrary to their self-reported predictions the experimental groups in the present study did not differ in their learning durations or their subjective estimates thereof. They also perceived similar workloads during the learning tasks, which confirms prior evidence (Noyes et al., 2004). Even though the learning time was not limited in this study, the participants were informed that they had 90 minutes for the lab session and many participants had to attend to their otherwise scheduled duties after the experiment. Therefore the learning time cannot be considered as completely self-regulated despite not being tightly restrained. This might have affected the learning durations, which varied widely between 16 and 79 minutes, though. Future studies could include fully self-regulated learning durations next to generously but externally limited durations, like in the present study, to investigate possible differences in temporal learning behaviour.

Taken together the findings suggest that neither the objective nor the subjective learning effectiveness and efficiency depend on the reading medium and cannot explain the prevalent preference for reading on paper in the academic context. It appears therefore counterintuitive that the three reading media did not even impact the participants' affective state during learning, which consequently also cannot account for its preference.

The question still to be answered is why most participants prefer reading printed texts to digitized texts and why they expect them to provide superior effectiveness and efficiency. Contradictory evidence seems aplenty. The answer may lay in the learning habits displayed on different reading media. While all participants reported editing printed texts in some way, about 40% did not engage in such activities at all when reading digitally. Furthermore about 80% of the participants stated that they missed something during this learning experiment as compared to their usual learning scenarios.

Almost all activities, which were reported missing, had something to do with active elaborations of the read texts, such as marking, underlining, annotating, summarizing, drawing mind maps, etc. During all those activities the learning text has to be actively attended to, which leads to deeper information processing and as a consequence to longer lasting memory traces (Craik & Lockhart, 1972; Tyler, Hertel, McCallum, & Ellis, 1979). It seems that when learning for exams, students engage in activities enhancing deeper information processing. They preferably perform them on paper even if they read digitally.

This behaviour might lead to confounding of reading texts as such with active elaboration of those texts' contents on paper and finally to the heuristics or belief that learning on paper is generally more effective and efficient. Thereby 'paper' may become a cue triggering the heuristic of superior effectiveness and efficiency of paper as learning medium. Such *metacognitive illusions* (defined as systematic errors in metacognitive monitoring of cognitive processes including learning; Serra & Melcalfe, 2009) are suggested to explain the dissociations between predictions made before learning and judgements given after a test

(Dunlosky et al., 2014), which was also found in the present study. Before the experiment the participants used the ‘paper-learning’ heuristic to predict more effective and efficient learning on paper. During the learning sessions in the lab, however, nobody could edit the texts in any way, which equally deprived participants in all reading conditions of the paper and pencil experience. It is therefore possible that it was the absence of this ‘paper’ cue in the context of that particular learning setting, which made all participants estimate similar quiz scores after the quiz, independent of the reading condition and contrary to prior predictions. It appears probable that the participants fell prey to their metacognitive illusions and, triggered by ‘paper’ as a cue, made inaccurate predictions before the learning session.

Choices of learning strategies depend in large part on metacognitive judgments (Dunlosky et al., 2014; Metcalfe, 2009; Metcalfe & Finn, 2008) and may probably also determine preferences for particular learning media, but, as described above, they are prone to be biased. Future research should therefore extend the present study to include accompanying editing activities, as they may explain the unfavourable adoption of digital devices for learning purposes. Involving individuals, who read on paper or digitally in editing on paper as well as digitally would be of particular interest. A recent study suggests that note taking longhand during a lecture is more effective than note taking on laptops (Mueller & Oppenheimer, 2014) and is worth a replication in different, esp. self-regulated, learning settings.

Interestingly only 11 % of the participants in the current study evaluated their abilities to mark, underline, and annotate PDF files digitally as rather advanced whereas almost half of the participants either never used this feature at all or rated their digital editing abilities as poor. Including abilities of editing texts presented on digital devices in future investigations could help to establish the reasons for the differences in learning habits between printed and digitized texts.

Students participating in the present study turned out to use desktops very rarely for learning purposes, which makes the above findings less conclusive. Replacing the desktop with a laptop condition would have much better corresponded with the participants habits and would have allowed more univocal interpretation of the results. Furthermore, light could have been shed on the surprising finding that despite widespread use of laptops, the participants were least satisfied with their effectiveness and efficiency as a learning medium.

Moreover, the usability of applications for editing digitally presented texts needs to be explored and possibly improved, as it was demonstrated not to sufficiently meet users' needs and expectations (e.g. Arif, Stuerzlinger, Mazalek, Kim, & Lee, 2015; King & Toland, 2014; Matulic & Norrie, 2012; Plimmer & Apperley, 2007) and may be another reason for the observed reluctance to use those applications. On the back of this, testing different digital input devices might shed further light on the reasons for the limited use of editing features of the current applications. Some participants in the lab stated that at least holding a pen would have facilitated their learning. Maybe pen-like devices allowing handwriting could enhance editing activities on digital devices and hereby enforce their wider adoption.

As a second line of investigation the study found differences in the affective state of the learners depending on the concurrent method of time estimation. The participants felt better, i.e. more interested, more in control and less frustrated in the retrospective than prospective method. The retrospective time judgment was made without being informed about its need in advance, while the prospective judgment was made with prior information about its need. The two time judgment methods are assumed to rely on and engage different resources (Block, Hancock, & Zakay, 2010). The prospective method requires division of attention between reading and time monitoring and hence exploits a part of attention. The retrospective estimate, on the other hand, is made solely from memory. The difference in the utilisation of different resources, memory versus attention, was hypothesized to explain the impact of the method of time estimation on affect (Liikkanen & Gómez, 2013).

However, such interpretation of the obtained results needs to be considered very cautiously as the prospective time estimation method in the present study coincided with the second text, which was the last in the sequence. Therefore fatigue progressively emerging during the experimental procedure might be a confounding factor, influencing the participants' affect. Interestingly, the method of time judgment did not affect the perceived workload, though, which hereby undermines the possible impact of fatigue. This issue should be addressed in future experiments, which could exclusively implement the prospective time estimation method for both learning tasks and compare the induced affect between them. Such procedure is not possible for retrospective judgements due to the fundamental feature of the method. A participant can be surprised just once to give a retrospective time estimate for a task, the following estimates in the same experimental context would certainly be confounded by prior experience.

This interesting possible effect of time estimation method on the affective state definitely requires further investigation. A confirmation of such an effect and its direction would inform and specify future studies, esp. in the area of usability and user experience, which often concurrently explore affective states and time estimations (Hornbæk, 2006; Liikkanen & Gómez, 2013).

From the perspective of the present findings the initiatives to go paperless in the academia should become more successful in future, because replacing printed texts with their digitized versions does not appear to compromise students' learning effectiveness, efficiency and affect during learning. Nonetheless, it has to be remembered that university students are probably primarily concerned with obtaining knowledge and expertise and not with demonstration of sustainable behaviour. Therefore in order to promote replacement of paper by digital media for learning purposes, it seems necessary to explicitly address students' main concerns and preferences first, followed by highlighting the ecological benefits of such replacement as a second step. For example raising awareness for the demonstrated

equivalence of printed and digitized texts with respect to effectiveness and efficiency of learning, while at the same time pointing to the opportunity to reduce the excessive paper usage by adoption of digital media, could foster the realization of sustainability initiatives. It seems possible that additional promotion of research and further development of digital input devices and text editing applications to better meet learners' needs (Arif et al.; 2015; King & Toland, 2014; Matulic & Norrie, 2012; Plimmer & Apperley, 2007) could trigger a transformation of students' learning preferences and hereby catalyse the establishment of ecologically sustainable paperless universities.

References

- Arif, A. S., Stuerzlinger, W., Mazalek, A., Kim, S., & Lee, G. (2015, April). A smart-restorable backspace to facilitate text entry error correction. In *CHI 2015 Workshop on Text Entry on the Edge*, Seoul, South Korea. To appear. Retrieved on March 15, 2015 from http://www.asarif.com/pub/workshop/SmartRestorableBksp_CHI15Workshop.pdf
- Bundesamt für Statistik. (2014). Informationsgesellschaft-Gesamtindikatoren: Haushalte und Bevölkerung – Internetnutzung. Retrieved on January 10, 2015 from http://www.bfs.admin.ch/bfs/portal/de/index/themen/16/04/key/approche_globale.indicator.30106.301.html?open=1,5#5
- Block, R. A., Hancock, P. A., & Zakay, D. (2010). How cognitive load affects duration judgments: A meta-analytic review. *Acta psychologica*, 134(3), 330-343.
- Bradley, M. M. & Lang, P. J. (1994). Measuring emotion: SAM and the semantic differential. *Journal of Experimental Psychiatry & Behavior Therapy*, 25, 49-59.
- Chou, I. C. (2012). Understanding on-screen reading behaviors in academic contexts: a case study of five graduate English-as-a-second-language students. *Computer Assisted Language Learning*, 25(5), 411-433.
- Clore, G. L., & Huntsinger, J. R. (2009). How the object of affect guides its impact. *Emotion Review*, 1(1), 39-54.
- Craik, F. I. M. & Lockhart, R. S. (1972). Levels of processing: a framework for memory research. *J. Verb. Learn. Verb. Behav.*, 11, 671-684.
- Dunlosky, J., Mueller, M. L., & Tauber, S. K. (2014). The contribution of processing fluency (and beliefs) to people's judgements of learning. In D. S. Lindsay, C. M. Kelley, A. P. Yonelinas, H.L. Roediger, III (Eds.), *Remembering: Attributions, Processes, and Control in Human Memory* (pp. 46-64). New York, NY: Psychology Press.
- Drake University. (2014). Paperless Initiative. Retrieved on December 18, 2014 from <http://www.drake.edu/sustainability/paperlessinitiative/>

- Ebbinghaus, H. (1885). *Über das Gedächtnis. Untersuchungen zur experimentellen Psychologie*. Leipzig: Verlag von Duncker & Humber. Retrieved on February 3, 2013 from <http://www.uni-leipzig.de/~psycho/wundt/opera/ebbing/memory/GdaechtI.htm>
- Ericsson, K. A., Krampe, R. Th., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review*, *100*, 363-406.
- Flamm, M. W. (2014). "Going Paperless": The Case for Electronic Submission of Student Work. *College Teaching*, *62*(1), 1-2.
- Franze, J., Marriott, K., & Wybrow, M. (2014, September). What academics want when reading digitally. In *Proceedings of the 2014 ACM symposium on Document engineering* (pp. 199-202). ACM.
- Fredrick, W. C., & Walberg, H. J. (1980). Learning as a function of time. *The Journal of Educational Research*, *73*(4), 183-194.
- Geddie, J. C., Boer, L. C., Edwards, R. J., Enderwick, T. P., Graff, N., Pfendler, C., Ruisseau, J.-Y., Van Loon, P. A. (2001). *NATO Guidelines on Human Engineering Testing and Evaluation* (pp. 26-32). Neuilly-sur-Seine Cedex, France: Research and Technology Organization North Atlantic Treaty Organization.
- Glasman, L. R., & Albarracín, D. (2006). Forming attitudes that predict future behavior: a meta-analysis of the attitude-behavior relation. *Psychological bulletin*, *132*(5), 778.
- Gregory, B., & Kaprelian, V.S. (2014). Types of errors. Retrieved on Mai 24, 2014 from http://patientsafetyed.duhs.duke.edu/module_e/types_errors.html
- Hart, S. G., & Staveland., L. E. (1988). Development of a multi-dimensional workload rating scale: Results of empirical and theoretical research. In P. A. Hancock & N. Meshkati (Eds.), *Human Mental Workload* (pp. 139-183). Amsterdam. The Netherlands: Elsevier.
- Hoeffler, S., & Ariely, D. (1999). Constructing stable preferences: A look into dimensions of experience and their impact on preference stability. *Journal of Consumer Psychology*,

8(2), 113-139.

Hoeffler, S., Ariely, D., & West, P. (2006). Path Dependent Preferences: The Role of Initial Experience and Biased Search in Preference Discovery. *Organizational Behavior and Human Decision Processes*, 101(2), 215-229.

Hornbæk, K. (2006). Current practice in measuring usability: Challenges to usability studies and research. *International journal of human-computer studies*, 64(2), 79-102.

Jungermann, H., Pfister, H.-R., & Fischer, K. (2010). *Die Psychologie der Entscheidung. Eine Einführung* (pp. 26, 31-40). Heidelberg: Spektrum Akademischer Verlag.

King, K., & Toland, J. (2014). iPads and the paperless office: The impact of tablet devices on paper consumption in higher education. *Journal of Applied Computing and Information Technology*, 18(1).

Liikkanen, L. A., & Gómez, P. G. (2013, September). Designing interactive systems for the experience of time. In *Proceedings of the 6th International Conference on Designing Pleasurable Products and Interfaces* (pp. 146-155). ACM.

Maastricht University Green Office. (2014). *Annual Plan 2014. Inspired by Sustainability – Implementing the Roadmap 2030*. (pp. 8-9). Retrieved on December 4, 2014 from <http://greenofficemaastricht.nl/publications/>

Mangen, A. Walgermo, B.R., & Brønnick, K. (2013). Reading linear texts on paper versus computer screen: Effects on reading comprehension. *International Journal of Educational Research*, 58, 61-68.

Margolin, S. J., Driscoll, C., Toland, M. J., & Kegler, J. L. (2013). E-readers, Computer Screens, or Paper: Does Reading Comprehension Change Across Media Platforms?. *Applied Cognitive Psychology*, 27(4), 512-519.

Matulic, F., & Norrie, M. C. (2012, May). Supporting active reading on pen and touch-operated tabletops. In *Proceedings of the International Working Conference on Advanced Visual Interfaces* (pp. 612-619). ACM.

- Metcalfe, J. (2009). Metacognitive judgments and control of study. *Current Directions in Psychological Science*, 18, 159–163.
- Metcalfe, J., & Finn, B. (2008). Evidence that judgments of learning are causally related to study choice. *Psychonomic Bulletin & Review*, 15, 174–179.
- Mueller, P. A., & Oppenheimer, D. M. (2014). The Pen Is Mightier Than the Keyboard Advantages of Longhand Over Laptop Note Taking. *Psychological Science*, 25, 1159-1168.
- Noyes, J. M., & Garland, K. J. (2008). Computer-vs. paper-based tasks: Are they equivalent?. *Ergonomics*, 51(9), 1352-1375.
- Noyes, J., Garland, K., & Robbins, L. (2004). Paper-based versus computer-based assessment: Is workload another test mode effect? *British Journal of Educational Technology*, 35(1), 111-113.
- Nygren, TH. E. (1991). Psychometric properties of subjective workload measurement techniques: Implications for their use in the assessment of perceived mental workload. *Human Factors*, 33(1), 17-33.
- Peters, E. (2006). The functions of affect in the construction of preferences. In S. Lichtenstein & P. Slovic (Eds.), *The construction of preference*, (pp. 454-463). New York: Cambridge University Press.
- Plimmer, B., & Apperley, M. (2007, July). Making paperless work. In *Proceedings of the 8th ACM SIGCHI New Zealand chapter's international conference on Computer-human interaction: design centered HCI* (pp. 1-8). ACM.
- Pfendler, C. (1990). Zur Messung der mentalen Beanspruchung mit dem NASA-Task Load Index. *Zeitschrift für Arbeitswissenschaft*, 44, 158-163.
- Preller, M., Zwahlen, S., Grösser, S. (2014). *E-Reader im Studium- Evaluationsstudie zur Einführung elektronischer Literatur an der Berner Fachhochschule*. Retrieved on

January 23, 2014 from <http://www.wirtschaft.bfh.ch/uploads/media/2014-01-Studie-Papierverbrauch.pdf>

- Rockinson-Szapkiw, A. J., Courduff, J., Carter, K., & Bennett, D. (2013). Electronic versus traditional print textbooks: A comparison study on the influence of university students' learning. *Computers & Education, 63*, 259-266.
- Sackett, A. M., Nelson, L. D., Meyvis, T., Converse, B. A., & Sackett, A. L. (2010). You're Having Fun When Time Flies: The Hedonic Consequences of Subjective Time Progression. *Psychological Science, 21*(1), 111-117.
- Serra, M. J., & Metcalfe, J. (2009). Effective implementation of metacognition. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Handbook of metacognition in education* (pp. 278–298). New York: Routledge.
- Sustainable Development Solutions Network. (2014). *An Action Agenda for Sustainable Development. Report for the UN Secretary-General*. Retrieved on June 22, 2014 from <http://unsdsn.org/wp-content/uploads/2013/06/140505-An-Action-Agenda-for-Sustainable-Development.pdf>
- Tyler, S. W., Hertel, P. T., McCallum, M. C., & Ellis, H. (1979). Cognitive effort and memory. *Journal of Experimental Psychology: Human Learning and Memory, 5*, 607–617.
- Wilson, T. D., Lisle, D. J., Kraft, D., & Wetzel, C. G. (1989). Preferences as expectation-driven inferences: effects of affective expectations on affective experience. *Journal of personality and social psychology, 56*(4), 519.

Appendix A

Online Pre-survey

1. Bitte, zeige an, wie Du die folgenden digitalen Geräte nutzt. Pro Gerät sind mehrere Antworten möglich.

	privat	fürs Studium (z.B.: Texte Lesen, Präsentation erstellen, Literaturrecherche, etc.)	zur Prüfungs- vorbereitung	beruflich (z.B.: Praktikum, Nebenjob)	benutze ich überhaupt nicht
Desktop-Computer	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Laptop-Computer	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Tablet-Computer	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Wenn Du möchtest, dann kannst Du hier Deine Antwort erläutern.

2. Im Durchschnitt, an wie vielen Tagen in der Woche liest du digitale PDFDateien präsentiert an an einem Monitor (z.B. mit Adobe Reader, Foxit Reader, GoodReader, o.a.) ?

- an 6-7 Tagen in der Woche
- an 4-5 Tagen in der Woche
- an 2-3 Tagen in der Woche
- an einem Tag in der Woche
- weniger als an einem Tag in der Woche.
- nie

3. Meine Fertigkeiten im Unterstreichen und Hinzufügen von digitalen Notizen an PDF-Dateien (z.B. mit Adobe Reader, Foxit Reader, GoodReader, o.a.) sind:

- eher gross
- durchschnittlich
- eher gering
- Diese Funktion benutze ich an digitalen PDF-Dateien gar nicht.

4. Wenn Du zur Prüfungsvorbereitung Texte liest, welches ist dann Dein BEVORZUGTES Medium?

- Monitor eines Desktop-Computers
- Monitor eines Laptop-Computers
- Tablet-Computer
- Papier (Ausdrucke, Kopien, Skripte, Bücher, etc.)

Wenn Du möchtest, dann kannst Du hier Deine Antwort erläutern.

5. Wenn Du aus auf PAPIER gedruckten Texten lernst, unterstreichst/ markierst Du dann Textpassagen?

- immer
- oft
- manchmal
- nie

Wenn Du möchtest dann kannst Du hier Deine Antwort erläutern.

6. Wenn Du aus DIGITAL präsentierten Texten lernst, unterstreichst/ markierst Du dann Textpassagen?

- immer
- oft
- manchmal
- nie

Wenn Du möchtest dann kannst Du hier Deine Antwort erläutern.

7. Wenn Du aus auf PAPIER gedruckten Texten lernst, fügst Du dann Notizen (im Text oder am Rand) ein?

- immer
- oft
- manchmal
- nie

wenn Du möchtest, dann kannst Du hier Deine Antwort erläutern.

8. Wenn Du aus DIGITAL präsentierten Texten lernst, fügst Du dann Notizen (im Text oder am Rand) ein?

- immer
- oft
- manchmal
- nie

wenn Du möchtest, dann kannst Du hier Deine Antwort erläutern.

9. Bitte, schätze Deine Lerneffektivität aus Texten ein und bezeichne die folgenden Medien in der Reihenfolge wie folgt:

- 3=so lerne ich gut/nachhaltig,**
- 2=so lerne ich mittelmässig,**
- 1=so lerne ich schlecht**
- 0=dieses Medium nutze ich beim Lernen nicht.**

Falls Du mehrere Medien als gleichwertig einschätzt, dann teile ihnen denselben Rang zu.

	3= gut/ nachhaltig	2= mittelmässig,	1= schlecht	0= nutze ich beim Lernen nicht
Auf Papier gedruckte Texte	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Texte auf dem Monitor eines Desktop-Computers	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Texte auf dem Monitor eines Laptop-Computers	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Texte auf einem Tablet-Computer	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

10. Bitte, schätze Deine Lerneffizienz aus Texten ein und bezeichne die folgenden Medien in der Reihenfolge wie folgt:

**3=so lerne ich schnell,
2=so lerne ich durchschnittlich schnell,
1=so lerne ich langsam,
0=dieses Medium nutze ich beim Lernen nicht.**

Falls Du mehrere Medien als gleichwertig einschätzt, dann teile ihnen denselben Rang zu.

	3= schnell	2= durchschnittlich	1= langsam	0= nutze ich beim Lernen nicht
Auf Papier gedruckte Texte	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Texte auf dem Monitor eines Desktop-Computers	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Texte auf dem Monitor eines Laptop-Computers	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Texte auf einem Tablet-Computer	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

11. Wie ist Deine bisherige Erfahrung mit Deiner SUBJEKTIVEN Einschätzung Deines Lernerfolges/ Deiner Prüfungsvorbereitung? Vor Prüfungen fühle ich mich MEISTENS

- kaum vorbereitet
 mittelmässig vorbereitet
 gut vorbereitet
 sehr gut vorbereitet

12. Meine objektiven Prüfungsergebnisse (=Noten) sind MEISTENS

- besser als ich erwartet habe
 meinen Erwartungen entsprechend
 schlechter als ich erwartet habe

13. Welches Fach studierst Du?

- Psychologie im Bachelorstudium
 Psychologie im Masterstudium
 Sonstiges (bitte angeben)

14. In welchem Semester studierst Du?

Studiensemester

15. Du bist

- weiblich
- männlich

16. Wie alt bist Du?

Dein Alter in Jahren

17. Hast Du eine Sehbehinderung, die mit einer Sehhilfe (Brille, Kontaktlinsen) NICHT ausgeglichen werden kann?

- Nein, ich habe keine Sehbehinderung, bzw. diese kann mit einer Sehhilfe ausgeglichen werden.
- Ja, ich habe eine Sehbehinderung, die mit einer Sehhilfe nicht ausgeglichen werden kann.

Appendix B

Learning Texts

A. Art und Umfang kognitiven Aufwandes

Der mit der Bearbeitung eines Entscheidungsproblems verbundene kognitive Aufwand hängt weitgehend davon ab, wie die entscheidungsrelevante Information repräsentiert ist bzw. ob überhaupt mentale Repräsentationen bereits vorliegen oder ob das notwendige Wissen erst angeeignet und strukturiert werden muß. Zwischen weitgehend automatisierten und mühelos ablaufenden Entscheidungen einerseits und ausführliche Informationssuche und -verarbeitung erfordernden Entscheidungen andererseits gibt es ein Kontinuum der kognitiven Anstrengung in der Art und dem Umfang der Nutzung kognitiver Ressourcen. Es können sowohl unterschiedlich viele Informationen herangezogen als auch unterschiedlich aufwendige Verarbeitungsprozesse eingesetzt werden (vgl. Svenson, 1990; Beach, Jungermann & de Bruyn, 1996). Mit dieser Dimension des kognitiven Aufwandes ist das Ausmaß an Reflexion und Bewußtheit korreliert, mit welcher Entscheidungen gefällt werden. In Anlehnung an Svenson (1990, 1996) unterscheiden wir vier Ebenen.

1. Routinisierte Entscheidungen

Die erste Ebene von Entscheidungen ist dadurch charakterisiert, daß die möglichen Optionen stets gleich sind und zwischen ihnen routinemäßig oder automatisch gewählt wird. In solchen Situationen sprechen wir nur dann von *Entscheidungen*, wenn diese das Resultat früherer, auf "höherer" Ebene angesiedelter Entscheidungen sind, die auf Grund häufiger Wiederholung routinisiert wurden.

Solche Entscheidungen verlangen den geringsten kognitiven Aufwand. Der Aufwand besteht im wesentlichen im Abgleich der gegebenen Situation mit vorgeschichteten Situationen und den in ihnen fixierten Entscheidungen; man spricht hier von einem *Matching*-Prozeß. Die Leichtigkeit dieses Abgleichs, d.h., ob die Situation als "passend" erkannt und damit der Automatismus aktiviert wird, hängt vor allem von der Ähnlichkeit der aktuellen Situation mit den im Gedächtnis gespeicherten Situationsprototypen ab. Bei hinreichend hoher Ähnlichkeit wird das gespeicherte Entscheidungsschema aktiviert und die gewohnte Wahl getroffen. Durch derartige *habituelle Präferenzen* werden aus der Menge von theoretisch möglichen Optionen diejenigen ausgefiltert, die bekannt sind und in ähnlichen Situationen bereits häufig gewählt wurden.

Modifikationen in der Wahl zwischen möglichen, im Gedächtnis gespeicherten Optionen werden durch differentielle Hinweisreize der Situation ausgelöst. Stößt die Durchführung des schematischen Ablaufs auf Hindernisse, d.h.: Ist der Normalfall nicht gegeben, wird der Entscheidungsprozeß abgebrochen, ein anderes Schema aktiviert oder auf eine höhere Ebene der Aufmerksamkeit und Reflexion gewechselt.

Eine routinisierte Entscheidungssituation ist die morgendliche Fahrt mit dem Wagen zum Arbeitsplatz. Die Entscheidung, welche Straßen genommen werden, ist fast vollständig habituiert ("erst Sophienstraße- dann rechts in die Oranienburger Straße - dann links in die Tucholskystraße ...") und wird lediglich durch feine Unterschiede in der Situationswahrnehmung (*situation awareness*) modifiziert ("Wenn der Wagen der Müllabfuhr auf der Tucholskystraße zu sehen ist, dann weiter auf der Oranienburger Straße bis zur Friedrichstraße"). Daß das "Fahrtstreckenschema" ablaufen kann, setzt die hinreichende Ähnlichkeit mit der jeweiligen morgendlichen Situation voraus. Ist die Ähnlichkeit zu gering, beispielsweise beim Aufenthalt in einer weniger vertrauten Stadt, ist die Ausführung des Schemas gestört und die Entscheidungen laufen nicht mehr oder nur noch teil-

weise automatisch ab; es muß, je nach Kenntnis der Stadt, eine Strecke ganz bewußt gewählt werden.

Die Situation kann das Verhalten allerdings auch so sehr bestimmen, daß inzwischen eingetretene Veränderungen der äußeren Situation nicht oder nur sehr verzögert wahrgenommen werden und das routinisierte Verhalten über längere Zeit hinweg suboptimale Konsequenzen hat. Erst die bewußte Wahrnehmung der Veränderung führt dann dazu, daß die Routine aufgegeben und eine neue Routine entwickelt wird. So fährt man oft noch längere Zeit einen größeren Umweg, bevor man merkt, daß die alte kürzere Strecke wieder befahrbar ist. Analoge Effekte von kognitiven Routinen sind in der Denkpsychologie vielfach untersucht worden (Schönpflug & Schönpflug, 1995, S. 287ff.). Der Vorteil solcher Routinen liegt darin, daß man Entscheidungen nicht stets wieder neu treffen muß und damit kognitive Ressourcen für andere Aktivitäten frei werden, beispielsweise während der morgendlichen Fahrt im Radio die Lesung eines Romans zu verfolgen. Der Nachteil liegt darin, daß die primäre Aktivität möglicherweise zu wenig Aufmerksamkeit erhält und dadurch Ereignisse nicht wahrgenommen werden, die eine Verhaltensänderung bzw. Umentscheidung verlangen. Mit diesem Problem beschäftigt sich u.a. die *Fehlerforschung*, auf die wir in Abschnitt 5 kurz eingehen.

2. Stereotype Entscheidungen

Stereotype Entscheidungen unterscheiden sich von routinisierten Entscheidungen durch zwei Aspekte: Erstens werden sie nicht durch die Gesamtsituation, sondern durch die Art der möglichen Entscheidungsoptionen ausgelöst, und zweitens gibt es einen minimalen Bewertungsprozeß. Dadurch ist ein höherer kognitiver Aufwand erforderlich, weshalb solche Entscheidungen meistens als "bewußte" Entscheidungen erlebt werden.

Stereotype Entscheidungen betreffen bestimmte Entscheidungsbereiche, beispielsweise Konsumententscheidungen ("Was will ich im Restaurant essen?" oder "Welchen Wein soll ich kaufen?"). Obwohl die Situation (z.B. das konkrete Restaurant) ganz verschieden sein kann, stammen die Optionen aus klar definierten Bereichen (z.B. Menüs).

Die Bewertungen erfolgen jedoch in hohem Maße stereotyp, d.h. nach erlernten Bewertungsschemata, die nicht mehr neu geprüft werden. Die Bewertung reduziert sich so auf den unmittelbaren Gesamteindruck oder auf wenige hervorstechende Merkmale der Optionen (z.B. bei Menüs: Preis, Art der Nachspeise, Fisch oder Fleisch usw.); Prozesse der Abwägung (*trade-offs*) sind eingeschliffen oder werden, falls die Merkmalskombinationen unvertraut sind, durch einfache Regeln gelöst.

Stereotype Entscheidungen sind durch Erfahrungen oder durch Gefühle (oder beides) bestimmt. Die Präferenz wird nicht durch eine bewußte Analyse der einzelnen Merkmale der Optionen gebildet, sondern durch ein *holistisches, intuitiv* erscheinendes Urteil; man spricht hier von einem holistischen Affekturteil, das so gut wie keine kognitive Verarbeitung voraussetzt. Die Wahrnehmung vieler Objekte, vor allem von Personen, führt zu einem unmittelbaren Eindruck, ob das Objekt oder die Person "gefällt" oder "sympathisch" ist. Möglicherweise ist für solche globalen Bewertungen ein vom kognitiven System unabhängiges Affektsystem verantwortlich. Zajonc (1980) hat postuliert, daß kognitive und affektive Prozesse bei solchen unmittelbaren Präferenzurteilen weitgehend unabhängig funktionieren. Derartige Phänomene werden meistens anhand des *Vertrautheitseffektes* untersucht. Werden einer Versuchsperson (Vp) beliebige neuartige Stimuli (z.B. abstrakte geometrische Figuren) in unterschiedlicher Häufigkeit dargeboten, und wird die Person anschließend gebeten, die Figuren hinsichtlich ihres "Gefallens" zu bewerten, so zeigt sich, daß Figuren, die häufiger dargeboten wurden, den Personen besser gefallen. Stimuli werden allein dadurch, daß man sie oft wahrgenommen hat, stärker präferiert. Dies ist, wie Experimente von Zajonc (1980; Kunst-Wilson & Zajonc, 1980) gezeigt haben, von der bewußten Wiedererkennung von Stimuli weitgehend unabhängig. Mit einer Entscheidung, die auf solchen holistischen Urteilen des unmittelbaren Gefallens beruht, werden Optionen selegiert, die

vertraut und bekannt sind.

3. Reflektierte Entscheidungen

Reflektierte Entscheidungen sind dadurch charakterisiert, daß keine habituellen oder stereotyp abrufbaren Präferenzen für die Optionen vorhanden sind. Der Entscheider denkt explizit über seine Präferenzen nach, er sucht nach Informationen in seinem Gedächtnis und gegebenenfalls auch in seiner Umgebung, und er bildet aus diesen Informationen seine Präferenzen. Dieser Prozeß umfaßt mindestens die Bewertung der Merkmalsausprägungen in Hinblick auf ihre Wünschbarkeit, kann aber auch die Analyse der Optionen auf relevante Merkmale sowie die Integration der Bewertungen einbeziehen. Dies erfordert einen wesentlich höheren kognitiven Aufwand als die bisher beschriebenen routinisierten und stereotypen Entscheidungen.

Reflektierte Entscheidungssituationen stehen im Zentrum der entscheidungs- psychologischen Forschung. Meistens wird von vorgegebenen Optionen ausgegangen und untersucht, in welcher Weise die Bewertungsprozesse ablaufen und nach welchen Regeln der Entscheider seine Auswahl trifft.

Die Reflexion einer Entscheidung muß nicht unbedingt zu einer Wahl führen. Entscheidungen können auch abgebrochen oder aufgeschoben werden. Das "Nicht-Fällen" von Entscheidungen kann nach Corbin (1980) mehrere Formen annehmen: Entscheidungen können beispielsweise *verweigert* werden; die Entscheidungssituation wird dann als solche nicht mehr akzeptiert, und man "tut nichts". Entscheidungen können auch *aufgeschoben* werden, oft mit der Hoffnung, daß sich eine Lösung wie von selbst ergibt ("aussitzen"). Da reflektierte Entscheidungen sich meistens auf Probleme beziehen, deren Wichtigkeit für den Entscheider relativ hoch ist (z.B. Kaufentscheidungen teurer Objekte), kann der Prozeß zusätzlich stark durch affektive und motivationale Faktoren beeinflusst werden.

Die Charakterisierung einer Entscheidung als *reflektiert* bedeutet nicht, daß emotionale Faktoren keine Rolle spielen, und auch nicht, daß alle Aspekte gründlich bedacht werden oder daß der Prozeß kein Hin und Her der Gefühle und Gedanken zeigt. Es bedeutet lediglich, daß über die Entscheidung bewußt nachgedacht wird, wie unvollständig und schwankend auch immer.

4. Konstruktive Entscheidungen

Charakteristisch für konstruktive Entscheidungen sind zwei weitere Aspekte: Erstens sind die Optionen entweder nicht vorgegeben oder nicht hinreichend genau definiert. Zweitens sind die für die Entscheidung relevanten persönlichen Werte entweder unklar oder müssen erst generiert werden. Dadurch verlangen Entscheidungen auf dieser Ebene den höchsten kognitiven Aufwand; es kommen zusätzlich kognitive Prozesse ins Spiel, insbesondere Prozesse der Suche nach Information. Dieser Entscheidungstyp ist in der Entscheidungsforschung bislang kaum behandelt worden (Fischhoff, 1996).

Die Konstruktion von *Optionen* ist charakteristisch für viele Alltagsentscheidungen. Bei klar definiertem Ziel (z.B. "ich will Urlaub machen") müssen die möglichen Optionen erst in einem konstruktiven Prozeß erzeugt werden. Die Informationssuche betrifft dabei sowohl eine Suche nach Optionen als auch eine Suche nach und Generierung von damit verbundenen Konsequenzen. Solche Entscheidungssituationen sind den *synthetischen* Problemen der Denkpsychologie analog (vgl. Dörner, 1976). Diese Suchprozesse können manchmal auch zu ungewöhnlichen und nicht vorher-gesehenen Entscheidungen führen.

Ist in einer Situation nicht nur die Wahl der Option unklar ("Was soll ich tun?"), sondern auch die Richtung, in die eine Wahl gehen soll ("Was will ich eigentlich?"), ist dem Prozeß der Wahl zwischen Optionen ein Prozeß der Entscheidung über die Ziele vorgeschaltet. Solche Situationen sind den *dialektischen* Problemen analog (vgl. Dörner, 1976).

Sowohl bei konstruktiven als auch bei reflektierten Entscheidungen kommt es häufig vor, daß bereits während des Generierungs- und Repräsentationsprozesses tentative, vorläufige Wahlen stattfinden, ohne also den Entscheidungsprozeß unbedingt zu beenden oder zu unterbrechen. Die Qualität dieser vorläufigen Wahl kann den Umfang und die Art weiterer kognitiver Such- und Konstruktionsprozesse bestimmen (Montgomery, 1983). Bis es also zur *faktischen* oder endgültigen Entscheidung mit einer Festlegung auf eine Option (*commitment*) kommt, können eine ganze Reihe weiterer tentativer Wahlen gefällt und wieder verworfen worden sein.

Es gibt weitere Unterschiede in Art und Umfang des kognitiven Aufwandes bei Entscheidungsprozessen. Die folgende Übersicht zeigt einige dieser Merkmale für die vier Entscheidungsebenen.

Ebene:	routinisiert	stereotyp	reflektiert	konstruktiv
Bewußtheit	nein	niedrig	hoch	hoch
Anforderung an Aufmerksamkeit	sehr gering	gering	hoch	sehr hoch
Generierung neuer Informationen	nein	nein	ja	ja
Zeitdauer	schnell	schnell	schnell-lange	lange
Flexibilität	kaum	gering	hoch	sehr hoch
Vorstrukturiertheit	sehr hoch	hoch	hoch-mittel	gering
Gedächtnis-repräsentation	S-R-Assoz./ Gewohnheits- hierarchien	Schemata, Skripte	Ziele, Konsequenzen	allgemeines Weltwissen
kognitive Prozesse	Matching	Schema- aktivierung	Bewertung	Konstruktions-/ Inferenzprozesse

B.

Beispiel 1: Eine falsche Entscheidung

Am 13. Januar 1982 startet eine Boeing-737 der Air Florida mit 79 Insassen vom Washington National Airport in Richtung Florida. Wegen heftigen Schneefalls mußte der Flughafen in den letzten zwei Stunden vorübergehend gesperrt werden. Der Betrieb rollt wieder an, nachdem die *runways* (Start- und Landebahnen) und *taxiways* (Rollbahnen) vom Schnee befreit worden sind und der Schneefall ausgesetzt hat. Die Besatzung ist nur wenig erfahren mit winterlichem Luftverkehr (Air Florida!). Während der zweistündigen Wartezeit in der Parkposition am Gate ist das Flugzeug bereits dreimal vom zuständigen Bodenpersonal enteist worden. Es wird nun als erstes Flugzeug in die Abflugreihenfolge eingegliedert und wartet auf einem Rollweg als führendes Flugzeug in einer längeren Warteschlange auf die Startfreigabe der Luftverkehrskontrolle, die wegen anfliegenden Verkehrs noch einige Zeit auf sich warten läßt. Die Crew hat die Enteisungssysteme an Bord des Flugzeuges am Boden nicht aktiviert. Daraufhin bildet sich an den Vorderkanten der Flügel bereits wieder ein Eisansatz, den die Crew vom Cockpit aus nicht erkennen kann, den sie aber ahnt. Zusätzlich vereist das linke engine inlet pressure probe (ein technisches Bauteil, das über die Schubkraft des jeweiligen Triebwerkes informiert, das aber bei Vereisung fehlerhafte Informationen, in der Regel einen zu großen Schub, anzeigt).

Nach der Startfreigabe rollt die B-737 unverzüglich zur Startbahn und beginnt den Startvorgang. Bei der Beschleunigung fallen dem Co-Piloten, der den Start durchführt, größere Unstimmigkeiten zwischen den Triebwerksanzeigen eines Triebwerkes auf. Trotz dieser Unstimmigkeiten wird der Start der B-737 nicht abgebrochen. Der weitere Ablauf ist wie folgt: Nach dem Start baut die Maschine nur ungenügend Geschwindigkeit auf, rotiert mit der Nase leicht nach oben und erleidet unmittelbar danach einen Strömungsabriß an beiden Tragflächen. Das Flugzeug hält sich nicht mehr in der Luft und kollidiert mit einer vielbefahrenen Straßenbrücke, reißt zwei Autos von der Fahrbahn und fällt in den eisigen Fluß unter der Brücke. 75 Menschen sterben (Kayten, 1994).

5. Parallele Differenzierungen in der Fehlerforschung

Die im vorherigen Abschnitt beschriebenen Ebenen weisen Parallelen auf zu den Ebenen der Aufgabenbearbeitung, die in den Modellen von Rasmussen (1986) und Reason (1994) im Rahmen der kognitionswissenschaftlichen Fehlerforschung unterschieden werden. So differenziert Rasmussen (Rasmussen & Jensen, 1974; Rasmussen, 1986) zwischen einer *fertigkeitsbasierten (skill-based) Ebene*, einer *regelbasierten (rule-based) Ebene* und einer *wissensbasierten (knowledge-based) Ebene*. Diese Ausführungsebenen werden, ähnlich wie bei Svenson (1990), hinsichtlich der Art der kognitiven Prozesse, des Grades der Bewußtheit bzw. Automatisiertheit und des Repräsentationsformates im Langzeitgedächtnis unterschieden.

Die *fertigkeitsbasierte Ebene* ist dadurch gekennzeichnet, daß mehr oder weniger automatisiert Handlungsrountinen und -schemata abgerufen werden, die im Langzeitgedächtnis gespeichert sind. Dieses Konzept entspricht dem Konzept der routinisierten Entscheidungen. Beispiele sind das Tippen eines Textes am PC oder Autofahren. Das versehentliche Löschen anstatt des Speicherns einer Datei oder die Bedienung eines falschen Hebels sind Fehler bei "der Ausführung und/oder Speicherung einer Handlungsfolge ..., ungeachtet der Frage, ob der Plan, dem sie folgen, angemessen war, um das gesetzte Ziel zu erreichen" (Reason, 1994, S. 28). Reason spricht bei *fertigkeitsbasierten Fehlern* auch von *Schnitzern (lapses)* oder *Patzern (slips)*. Im Beispiel des Piloten im Beispiel 1 hätte dann ein fertigkeitsbasierter Fehler vorgelegen, wenn der Pilot versehentlich (z.B. durch eine falsche Bewegungsabfolge) den Schubkrafthebel eines Triebwerkes in die falsche Richtung bewegt und damit die Schubkraft des Triebwerkes - statt zu erhöhen - verringert hätte.

Auf der *regelbasierten Ebene* wird mit Hilfe von "Wenn-dann"-Regeln entschieden. Die Anwendung dieser Regeln verlangt Inferenzprozesse, weil geprüft werden muß, ob die Voraussetzungen für die Anwendung der Regel erfüllt sind. Diese Ebene ähnelt der Ebene der stereotypen Entscheidungen, denn auch dort ist ein

Minimum an Bewertungsprozessen nötig, um Entscheidungen zu treffen. *Regelbasierte Fehler* beruhen auf der falschen Einschätzung von Situationen mit der Konsequenz der Auswahl falscher Regeln oder dem Abruf zur Problemlösung ungeeigneter Prozeduren. Unterstellen wir einmal, daß der Pilot in unserem Beispiel wußte, daß bei Frost und längerer Wartezeit vor dem Abflug die Enteistungssysteme an Bord anzuschalten sind. Wenn also kein Mangel an Wissen oder kein fehlerhaftes Wissen vorlag, dann war das Nicht-Einschalten dieser Systeme eine Fehlhandlung auf der regelba-sierten Ebene. Die WENN-DANN- Regel war bekannt: <WENN Frost und längere Wartezeit, DANN Einschalten der Enteistungssysteme>. Aber vielleicht hatte der Pilot die Situation, also die WENN-Bedingung, nicht richtig klassifiziert, die Stärke des Frostes oder die Länge der Wartezeit unterschätzt. Und dies führte dann dazu, daß die in dieser Situation korrekte Wenn- Dann- Verknüpfung nicht realisiert, also die Enteistungssysteme nicht eingeschaltet wurden.

Hätte dagegen der - sehr unwahrscheinliche - Fall vorgelegen, daß der Pilot gar nicht über das entsprechende Wissen verfügte, also nicht über die <WENN Frost, DANN Enteistungssysteme einschalten>-Regel verfügt hätte, dann hätte es sich um einen *wissensbasierten* Fehler gehandelt. Auf der *wissensbasierten Ebene* werden Probleme dann gelöst, wenn diese neuartig sind. Hier können nicht einfach gespeicherte Regeln angewendet werden, sondern es müssen neue Problemlösungen konstruiert werden. Dies erfordert die Nutzung sowohl von deklarativem als auch von prozeduralem Wissen. Wissensbasierte Fehler entstehen durch unvollständiges oder fehlerhaftes deklaratives und/oder prozedurales Wissen. Die wissensbasierte Ebene weist Ähnlichkeiten zu den Ebenen der reflektierten und konstruktiven Entscheidungen bei Svenson auf. Auch sie können nur unter Nutzung von deklarativem *und* prozeduralem Wissen getroffen werden und erfordern mehr und umfangreichere kognitive Prozesse als die Entscheidungen auf den anderen beiden Ebenen. So gibt es bei reflektierten Entscheidungen keine abrufbaren Präferenzen zwischen den Optionen. Die Präferenzen müssen erst gebildet werden, was die Bewertung der Optionen hinsichtlich ihrer Konsequenzen und deren Wahrscheinlichkeiten erfordert. Bei den konstruktiven Entscheidungen müssen darüber hinaus die Optionen selbst erst in einem konstruktiven Prozeß erzeugt werden.

Trotz dieser Parallelen gibt es bislang kaum Verknüpfungen der theoretischen Konzepte von Entscheidungspsychologie und Fehlerforschung, obwohl die Nutzung der Modelle der jeweils anderen Disziplin zur Erklärung von Phänomenen und Befunden aus dem eigenen Gebiet durchaus interessant sein könnte. Ansätze in dieser Richtung finden sich in dem Bereich, der sich unter der Bezeichnung *naturalistic decision making* mit sogenannten *real world problems* beschäftigt; in einem Band von Klein, Orasanu, Calderwood und Zsombok (1993) findet man mehrere Arbeiten, u.a. auch von Rasmussen, die Fehlerforschung und Entscheidungsforschung miteinander verbinden.

Appendix C

NASA-TLX Questionnaire

Markiere, bitte, auf den folgenden Skalen, in welchem Masse Du Dich in den 6 genannten Dimensionen von der ERSTEN Lese-/Lernaufgabe beansprucht bzw. gefordert gefühlt hast.

Beispiel:

**Geistige Anforderung**

Wie viel geistige Aktivität war gefordert? (z.B. Denken, Entscheiden, Rechnen, Erinnern, Beobachten, Suchen, usw.) War die Lese-/Lernaufgabe leicht oder anspruchsvoll, einfach oder komplex?

**Körperliche Anforderung**

Wie viel körperliche Aktivität war erforderlich? (z.B. Drücken, Ziehen, Drehen, Kontrollieren, Aktivieren, usw.) War die Lese-/Lernaufgabe locker oder anstrengend, erholsam oder mühselig?

**Zeitliche Anforderung**

Wie viel Zeitdruck verursachte die Lese-/Lernaufgabe? War das Tempo langsam und geruhsam oder schnell und hektisch?

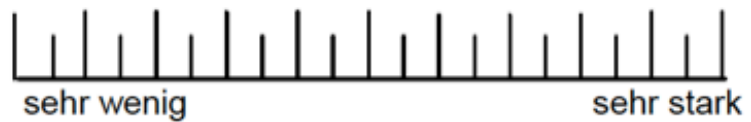


Eigene Leistung

Wie erfolgreich hast Du die geforderte Lese-/Lernaufgabe Deiner Meinung nach erfüllt? Wie zufrieden bist Du mit der Aufgabenbewältigung?

**Anstrengung**

Wie stark musstest Du Dich anstrengen (geistig und/oder körperlich), um Deinen Grad der Aufgabenerfüllung zu erreichen?

**Frustration**

Wie unsicher, entmutigt, irritiert, gestresst und verärgert hast Du Dich während der Lese-/Lernaufgabe gefühlt?



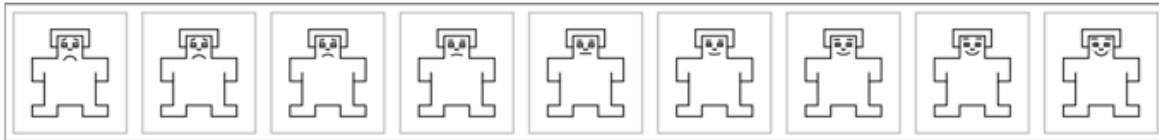
Appendix D

SAM Questionnaire

Während des Lernens aus dem ERSTEN Text fühlte ich mich:

Bitte, kreuze das entsprechende Symbol an.

Allgemeines Wohlbefinden



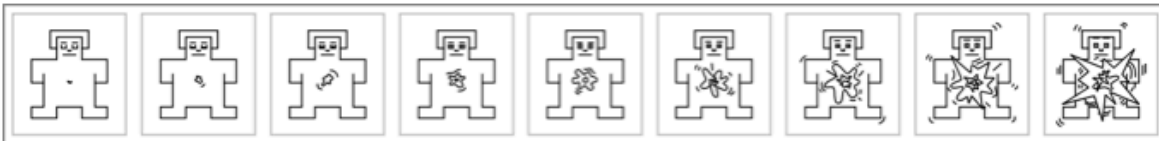
sehr
unzufrieden

sehr
zufrieden

Während des Lernens aus dem ERSTEN Text fühlte ich mich:

Bitte, kreuze das entsprechende Symbol an.

Aktiviertheit / Leistungsbereitschaft



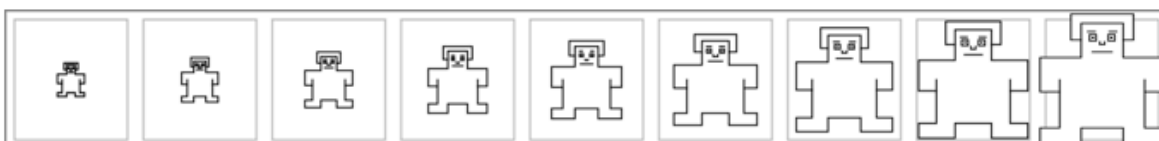
müde
gelangweilt

wach
interessiert

Während des Lernens aus dem ERSTEN Text fühlte ich mich:

Bitte, kreuze das entsprechende Symbol an.

Kontrolle / Dominanz



nichts
unter Kontrolle

alles
unter Kontrolle

Appendix E

Online Quiz and Post-survey

1. Welche Aussage bezüglich der Art und des Umfanges der Nutzung kognitiver Ressourcen bei unterschiedlichen Entscheidungsebenen trifft zu?

- Der kognitive Aufwand ist immer gleich.
- Der kognitive Aufwand ist für jede Art der Entscheidung klar abgrenzbar.
- Der kognitive Aufwand liegt auf einem Kontinuum.

2. Ordne folgende Entscheidungsebenen nach der Grösse des kognitiven Aufwandes. Ordne dem grössten Aufwand eine 4 und dem geringsten Aufwand eine 1 zu.

<input type="text" value="4"/>	stereotype Entscheidungen
<input type="text" value="3"/>	konstruktive Entscheidungen
<input type="text" value="2"/>	routinisierte Entscheidungen
<input type="text" value="1"/>	reflektierte Entscheidungen

3. Nenne den Hauptvorteil und den Hauptnachteil routinisierter Entscheidungen.

Hauptvorteil

Hauptnachteil

4. Bei welchen Entscheidungen werden holistische Affekurteile am ehesten gefunden?

- routinisierte Entscheidungen
- stereotype Entscheidungen
- reflektierte Entscheidungen
- konstruktive Entscheidungen

**5. Was ist typisch für stereotype Entscheidungen?
(mehrere Antworten möglich)**

- Sie werden durch die Gesamtsituation ausgelöst.
- Es findet ein minimaler Bewertungsprozess statt.
- Sie werden durch mögliche Entscheidungsoptionen ausgelöst.
- Es findet kein Bewertungsprozess statt.

6. Ordne jeder Kaufsituation jeweils eine der Entscheidungsebenen zu.

	stereotype Entscheidung	reflektierte Entscheidung	routinierte Entscheidung	konstruktive Entscheidung
Kauf der Lieblingsmarmelade im neuen Supermarkt.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Kauf eines Sofas für die neue Wohnung.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

7. Anja brauchte eine neue Sommerhose. Sie ging in die Stadt fest entschlossen die modische hellblaue skinny Jeans zu kaufen, die sie neulich im Schaufenster einer Edelboutique gesehen hatte. Kurz vor der Kasse machte Anja jedoch kehrt und hing die hellblaue Jeans zurück. Sie kaufte eine rote, dünne Baumwollhose mit der sie bis heute sehr zufrieden ist. Um welche Entscheidung handelt es sich hierbei am ehesten?

- routinisierte Entscheidung
- stereotype Entscheidung
- reflektierte Entscheidung
- konstruktive Entscheidung

8. Bei welchen Entscheidungsprozessen finden vorläufige Wahlen statt, ohne dass der Entscheidungsprozess unterbrochen oder beendet wird?

(mehrere Antworten möglich)

- stereotype Entscheidungen
- konstruktive Entscheidungen
- routinisierte Entscheidungen
- reflektierte Entscheidungen

9. Bei welchen Entscheidungen sind die relevanten persönlichen Werte nicht von Anfang an näher präzisiert?

(mehrere Antworten möglich)

- konstruktive Entscheidungen
- reflektierte Entscheidungen
- stereotype Entscheidungen
- routinisierte Entscheidungen

10. Welche Entscheidungsebene(n) ist (sind) in der Entscheidungsforschung bisher kaum behandelt worden?

(mehrere Antworten möglich)

- routinisierte Entscheidungen
- stereotype Entscheidungen
- reflektierte Entscheidungen
- konstruktive Entscheidungen

11. Welche Aussage(n) trifft (treffen) am ehesten zu?

(mehrere Antworten möglich)

- Es gibt viele Parallelen zwischen der Entscheidungspsychologie und der Fehlerforschung.
- Es gibt kaum Parallelen zwischen der Entscheidungspsychologie und der Fehlerforschung.
- Es gibt kaum Verknüpfungen der theoretischen Konzepte der Entscheidungspsychologie und der Fehlerforschung.
- Es gibt viele Verknüpfungen der theoretischen Konzepte der Entscheidungspsychologie und der Fehlerforschung.

**12. Die Ausführungsebenen in der kognitionswissenschaftlichen Fehlerforschung von Rasmussen werden unterschieden hinsichtlich:
(mehrere Antworten möglich)**

- der Art der kognitiven Prozesse
- dem Grad der emotionalen Intensität
- dem Grad der Bewusstheit
- des Repräsentationsformates im Langzeitgedächtnis

13. Welche der unten genannten Ausführungsebenen kommt nicht in dem Fehlermodell von Rasmussen vor?

- fertigkeitsbasierte Ebene
- reflexionsbasierte Ebene
- regelbasierte Ebene
- wissensbasierte Ebene

14. Ordne die zwei Ausführungsebenen des Fehlermodells von Rasmussen einer der Entscheidungsebenen von Svenson zu.

	konstruktive Entscheidung	reflektierte Entscheidung	routinisierte Entscheidung	stereotype Entscheidung
fertigkeitsbasierte Ausführungsebene	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
regelbasierte Ausführungsebene	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

15. Wie nennt man zwei Fehlertypen, die auf der fertigkeitsbasierten Ebene auftreten können?

Fehlertyp 1

Fehlertyp 2

16. Wissensbasierte Fehler passieren am ehesten

- bei routinisierten Handlungen.
- wenn Bewertung einer Situation notwendig wird.
- wenn neuartige Probleme gelöst werden müssen.
- wenn Inferenzprozesse notwendig werden.

17. Reto fährt seit 9 Jahren regelmässig und unfallfrei Auto. Vor ihm entlang der rechten Fahrbahnbegrenzung fährt ein Gelegenheitsfahrradfahrer, der nie einen Führerschein besessen hat. Als Reto an ihm vorbei fahren will, biegt der Radfahrer plötzlich links ab und wird von Retos Auto erfasst. Der Fehler, der RETO unterlaufen ist, ist am ehesten

- fertigkeitbasiert
- regelbasiert
- wissensbasiert

18. Reto fährt seit 9 Jahren regelmässig und unfallfrei Auto. Vor ihm entlang der rechten Fahrbahnbegrenzung fährt ein Gelegenheitsfahrradfahrer, der nie einen Führerschein besessen hat. Als Reto an ihm vorbei fahren will, biegt der Radfahrer plötzlich links ab und wird von Retos Auto erfasst. Welchen Fehler beging der RADFAHRER am ehesten?

- einen fertigkeitbasierten Fehler
- einen regelbasierten Fehler
- einen wissensbasierten Fehler

19. Ordne folgende Situationen einer Fehlerebene zu.

	wissensbasierter Fehler	fertigkeitbasierter Fehler	regelbasierter Fehler
Andy geht an einem Feiertag den gewohnten Weg zur Arbeit statt wie geplant zum Park.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ein Gärtner pflanzt eine Azalee (braucht sauren Boden) im Garten seines Häuschens am Meer (ungeeigneter "salziger" Boden).	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ein Tischler-Lehrling schneidet ein Regalbrett zu dünn, weil er die Dicke des Sägeblattes nicht berücksichtigt hat.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

20. Eine erfahrene Krankenschwester verabreicht ein ungeeignetes Medikament, weil sie die falsche Spritze gegriffen hat. Um welchen Fehler handelt es sich hierbei?

- einen fertigkeitbasierten Fehler
- einen emotionsbasierten Fehler
- einen wissensbasierten Fehler
- einen regelbasierten Fehler
- keinen der oben genannten Fehler

21. Du hast gerade 20 Fragen zu den Texten, die Du gestern gelesen hast beantwortet. Was denkst Du, wie viele Fragen hast Du richtig beantworten können?

Anzahl richtig beantworteter Fragen aus

20

22. Ich fand die Inhalte der Lerntexte

- sehr einfach
- einfach
- mittelschwierig
- schwierig
- sehr schwierig

Wenn Du möchtest, dann kannst Du hier Deine Antwort erläutern.

23. Ich fand die Testfragen insgesamt

- sehr einfach
- einfach
- mittelschwierig
- schwierig
- sehr schwierig

wenn Du möchtest, dann kannst Du hier Deine Antwort erläutern.

24. Meine Motivation im Vergleich zu einer studienrelevanten Prüfung war

- grösser
- gleich gross
- geringer

Wenn Du möchtest, dann kannst Du hier Deine Antwort erläutern.

25. Bitte, schätze Deine Lernerfahrung während dieser Studie generell ein.

- Ich habe wie gewohnt aus den Texten gelernt.
- Mir hat Folgendes zur gewohnten Lernerfahrung gefehlt:

Appendix F

Written Instruction for the Lab Experiment

Herzlich Willkommen bei der Lehrmittelstudie.

Bitte, lies alle Instruktionen genau durch. Falls Dir etwas unklar ist oder Du Fragen hast, wende Dich an die Testleiterin.

Im Folgenden werden Dir nach einander zwei Texte aus der Entscheidungspsychologie präsentiert. Deine Aufgabe ist es sie so zu lesen und zu studieren, wie wenn Du für Prüfungen lernst. Nach jedem Text wirst Du noch zwei kurze Fragebögen ausfüllen müssen, die sich auf Deine Erfahrung und Dein Befinden während des Lesens/Lernens beziehen.

- **Nimm Dir die Zeit**, die Du brauchst, um die in den Lerntexten dargestellten **Zusammenhänge zu verstehen**.
- Du **darfst** den jeweiligen Text ganz oder in Auszügen **mehrmals lesen**.
- **Morgen** wird in einem **online-Test** Dein Wissen, das Du Dir heute aneignest, überprüft. Der Test besteht aus Multiple-Choice- und aus offenen Fragen, es werden sowohl Wiedergabe- als auch Transferfragen zu beantworten sein.

Dein Engagement wird sich lohnen, denn die testbesten 12% werden zusätzlich zu den 5 Unterschriften mit je einem Pro Innerstadt-Geschenkbögen im Wert von CHF 50.- belohnt werden.

Appendix G

Written Instruction for the Quiz

Morgen (ca. 24 Std. nach dem heutigen Termin) bekommst Du eine E-Mail mit dem Link zum Wissenstest, dessen vollständige Beantwortung Dir die 5 Unterschriften sichert. Dabei ist Folgendes zu beachten:

1. Der Link zum **Test** ist an Deine E-Mail-Adresse gebunden und **kann deshalb nur von der entsprechenden Mailbox aus geöffnet werden**. An eine andere Mailbox weitergeleitete Links funktionieren nicht.
2. Nach Eingang der E-Mail in Deiner Mailbox bleibt der **Link zum Wissenstest nur 12 Stunden lang aktiv**. Danach wird er automatisch deaktiviert und die Unterschriften werden nur teilweise erteilt.
3. Sobald Du den Link zum Test geöffnet hast, bleiben Dir **45 min Zeit, um die Fragen zu beantworten**. Es ist nicht möglich den Test zu unterbrechen und später fortzusetzen. Der Test muss an einem Stück bearbeitet werden. Solltest Du den Test innerhalb der vorgegebenen Zeit nicht vollständig beantworten, wird der Test automatisch abgebrochen und die Unterschriften werden nur teilweise erteilt.
4. Am Ende des Wissenstests befindet sich die Information wann und wo Du die Unterschriften abholen kannst.
5. Nach Beendigung der Studie (u.U. erst November 2014) werden die testbesten 10% der Teilnehmer ermittelt. Diese gewinnen zusätzlich je einen Pro Innerstadt-Geschenkbbon im Wert von CHF 50. Die Gewinner werden per E-Mail benachrichtigt.