



# UNIVERSITY OF TWENTE.

Faculty of Behavioral, Management &  
Social Science

## The effectiveness of reviews and the role of practice when learning statistics from instructional videos

Paul Dunkel

S1216392

M.Sc. Thesis

July 2018

---

### Supervisors:

Dr. Hans van der Meij  
Dr. Henny Leemkuil

Master Psychology  
Learning Sciences  
University of Twente

---

## ABSTRACT

**Background:** Many students perceive statistics as the most anxiety inducing course in their degree program. Nowadays, statistical practice is closely related to new software-programs such as SPSS which is often taught through video instructions. The design of these instructional videos needs to be adequately optimized and tailored for students learning statistics.

**Aim:** The present article investigates the effectiveness of reviews and the role of practice when learning statistics from instructional videos by making use of the Demonstration-Based Training (DBT) approach. The main goal was to assess whether the inclusion of a review and/or a practice component increases motivation and learning outcome.

**Method:** By means of an online-experiment with 70 students, videos were tested in a university-level statistics course. Students were randomly assigned to one of four conditions: a) review-practice b) review c) practice d) control. User logs were recorded to measure video engagement. Learning outcome was measured with a multiple-choice knowledge test, an SPSS performance test and an SPSS transfer test. Motivation was assessed with a questionnaire measuring task-relevance and self-efficacy.

**Results:** The findings suggest that a review alone does neither affect learning outcome nor motivation positively. However, this study found a strong positive effect of the practice component on a subsequent SPSS performance test and motivation in terms of self-efficacy. In addition, an interaction effect between review and practice for increasing self-efficacy could be found.

**Conclusion:** This study sheds more light on how instructional videos should be designed in contemporary classrooms. The contribution of review and practice was critically examined and offers paths for future research in multimedia-based learning.

## Table of contents

|   |    |
|---|----|
| <b>ABSTRACT</b> .....                               | 1  |
| <b>1. INTRODUCTION</b> .....                        | 3  |
| <b>2. THEORETICAL FOUNDATION</b> .....              | 5  |
| <b>2.1 Reviews</b> .....                            | 6  |
| 2.1.1. <i>Text-based summaries</i> .....            | 6  |
| 2.1.2 <i>Reviews in software-training</i> .....     | 7  |
| <b>2.2 Practice</b> .....                           | 7  |
| 2.2.1 <i>Worked-examples</i> .....                  | 7  |
| 2.2.2 <i>Practice in software-training</i> .....    | 8  |
| <b>3. RESEARCH MODEL &amp; DESIGN</b> .....         | 10 |
| <b>4. METHOD</b> .....                              | 12 |
| <b>4.1 Participants</b> .....                       | 12 |
| <b>4.2 Instructional materials</b> .....            | 12 |
| 4.2.1 <i>Review in IVs</i> .....                    | 13 |
| 4.2.2 <i>Practice in IVs</i> .....                  | 14 |
| <b>4.3 Instruments</b> .....                        | 15 |
| 4.3.1 <i>SPSS practice</i> .....                    | 15 |
| 4.3.2 <i>SPSS performance test</i> .....            | 15 |
| 4.3.3 <i>SPSS Transfer test</i> .....               | 15 |
| 4.3.4 <i>Knowledge test</i> .....                   | 16 |
| 4.3.5 <i>Motivation Questionnaire</i> .....         | 16 |
| <b>4.4 Procedure</b> .....                          | 16 |
| <b>4.5 Data Analysis</b> .....                      | 17 |
| <b>5. RESULTS</b> .....                             | 18 |
| <b>5.1 Engagement</b> .....                         | 18 |
| <b>5.2 Learning outcome</b> .....                   | 19 |
| <b>5.3 Motivation</b> .....                         | 21 |
| <b>5.4 Relation among dependent variables</b> ..... | 21 |
| <b>6. DISCUSSION</b> .....                          | 23 |
| <b>6.1 Limitations and future research</b> .....    | 24 |
| <b>6.2 Conclusion</b> .....                         | 25 |
| <b>References</b> .....                             | 26 |
| <b>Appendix</b> .....                               | 31 |

# 1. INTRODUCTION

Statistics may be one of the most demanding and rigorous courses during scientific studies and evokes cognitive as well as emotional distress in many students (Onwuegbuzie & Wilson, 2003). Reasons for the perceived cognitive distress are that students not only have to memorize concepts, theories, principles and formulas but also have to conduct analyses and formulate hypotheses which makes learning statistics a difficult task (Matthew & Clark, 2003). These cognitive challenges in turn lead to emotional distress within the students when facing a statistical task. Empirical evidence suggests that students in nonmathematical studies perceive statistics courses as the most anxiety-inducing course in their degree program (Chew & Dillon, 2014).

Modern statistical practice is closely related to new software programs such as the Statistical Package for Social Science (SPSS) which needs to be learned by students in order to conduct statistical analyses properly (Goud, 2010; Baglin & Da Costa, 2014). It has been suggested that the usage of such statistical software programs contributes additionally to the perceived distress many students experience (DeVaney, 2010). Thus, next to the cognitive skills, students nowadays also need to develop some technological skills which even raises the challenge students' encounter when facing a statistical task (Baglin & Da Costa, 2014).

A widely used method to ease and facilitate the learning of statistical principles and software in higher education is the use of video tutorials (Kay & Kletskin, 2012), often referred to as "how-to" or "instructional" videos. Instructional Videos (IVs) transfer content of a specific theme via Demonstration-Based Training (DBT). Their primary purpose is to enable, support or guide task completion (Van der Meij & van der Meij, 2016). Probably IVs are best known to students through the website Youtube, launched in 2005, which is currently the second most popular website in the internet (Alexa Top 500 Global Sites, 2018). The popularity of this web page indicates the potential it brings into educational science.

Within the last decade, the usage of video tutorials in higher education has grown rapidly and videos are specifically designed for instruction (Brar & Van der Meij, 2017; Lloyd & Robertson, 2012). Research has shown that the implementation of such videos in the curriculum can result in significant gains in skills (Alpay & Gulati, 2010), test scores (Traphagen, Kusera & Kishi, 2010) and grades (Wieling & Hofman, 2010). Besides this, there are further advantages to IVs. For example, students can control their pace of learning, theory and practice can be combined, and the videos can be viewed anyplace, anywhere and anytime. Research has further shown that students have a preference for multimedia presentation when they find themselves in learning situations (Veronikas & Maushak, 2005; DeVaney, 2010). Additionally, video-based instructions improve students' motivation in terms of attention and result in more memorized content compared to content provided via traditional text-based instructions (Choi & Johnson, 2005).

To improve students understanding of principles of statistics and the pertaining software, methods of video instructions should be optimized to increase motivation and learning outcomes. Past research has already developed assumptions, principles and guidelines for a successful creation of video instruction. For example, Mayer (2008) provides 10 principles of multimedia instructional design and Koumi (2013) complements this framework by developing design

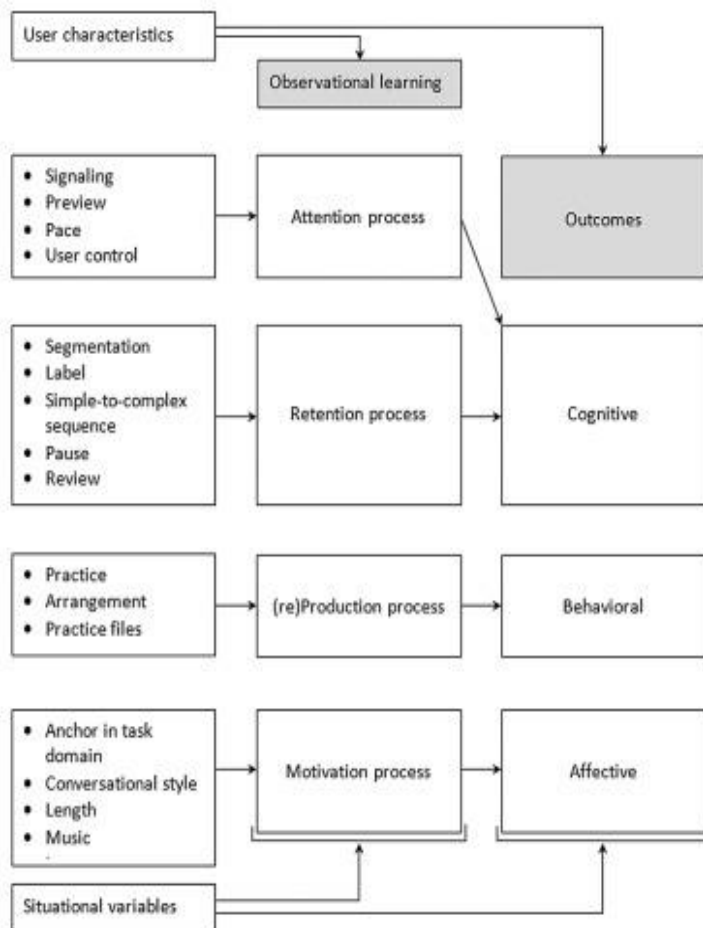
guidelines for educational multimedia materials. Regarding video tutorials in software training, Van der Meij and van der Meij (2013) introduced guidelines for their successful creation. For instance, the tutorials should preview the tasks and provide procedural rather than conceptual information.

The present research wants to extend the existing knowledge about creating successful IVs for software training. Two aspects which have received little to no attention in multimedia research are the effectiveness of reviews and the role of practice for enhancing learning outcomes and motivation among students working with SPSS. In particular, this study wants to investigate how the inclusion of a review and a practice component in an IV affects a student's learning outcome in terms of knowledge, performance and transfer as well as motivation in terms of task-relevance and self-efficacy.

## 2. THEORETICAL FOUNDATION

To create an IV for software training, this study utilized an adapted version of the Demonstration-Based Training (DBT) model from Brar and Van der Meij (2017; Fig. 1). Originating in Bandura's (1986) views on observational learning, DBT assumes that learning occurs through observation. In particular, DBT means acquiring knowledge, skills and attitudes by viewing examples, demonstrations or performances (Rosen, Salas, Pavlas, Jensen, Fu & Lampton, 2010). Based upon this, the model shows which instructional features can support the interrelated processes of attention, retention, production and motivation to facilitate positive outcomes in software training.

Generally video demonstrations are surely an easy and valuable technique to deliver information. However, solely observing a demonstration is no assurance for consecutive learning or maintaining the delivered information (Rosen et al, 2010). If the viewer passively observes the demonstrated content, its value is threatened. For that reason it is essential that the viewer actively watches the demonstration in order to process the content more deeply.



Active and deeper processing can be stimulated by making use of features from the DBT-model, such as including a review or a practice component after the demonstration (Van der Meij & Van der Meij, 2013). To investigate the value of reviews and practices for learning, this study used prerecorded IVs as a mean to demonstrate statistical principles and concepts in software training. Surprisingly, no research has been done yet where both a practice and a review component are present in an IV. Consequently, the present study investigates the single and interactive effects of a review and practice component on students' motivation and learning outcome.

**Fig. 1.** DBT-model of the connection between conditions, instructional features, learning processes and outcomes in software-training (adapted from Brar & van der Meij, 2017)

## 2.1 Reviews

Referring to the DBT-model, there are several reasons to assume that reviews in software-training could have a positive influence on the retention process which in turn might influence learning outcome positively. First, including a review in IVs could strengthen the retention process by summarizing the key points, giving the user an overview of the main issues in a procedure to be learned. Especially in cases where the learning content is diverse and complex, summarizing key points could help the learner in organizing the content. Second, when an IV is featured with a review, the user can compare and correlate it with the summary he or she self-constructed while watching the demonstration. If the user notices any disparity, he or she can replay the section of the IV in question. Third, a review is a short repetition of the demonstration giving the user a second chance to learn. This might be the case if the user was distracted at any point while watching the demonstration. In addition, a short repetition of the main steps can strengthen memorization (Brar & Van der Meij, 2017).

The effectiveness of reviews in IVs for software training has hardly been examined in the past. Past research focused more on the effectiveness of predefined text-based summaries than on multimedia-based summaries. To support the benefits of the inclusion of reviews in software-training, the effectiveness of text-based summaries needs to be investigated. Therefore, the research was extended to a related field, text-based summaries, with the intention to transmit insights from this field to multimedia-based reviews.

### 2.1.1. Text-based summaries

The effectiveness of text-based summaries was investigated decades ago. Hartley and Trueman (1982) reviewed four empirical studies on text-based summaries which also included an investigation of summaries' placements within texts. First, they mentioned a study published by Christensen and Stordahl (1955) where a reliable effect was absent and a study published by Vezin, Berge and Mavrelis (1973) where a positive effect for the inclusion of an end summary could be assessed.

Next, they referenced a study from McLaughlin Cook (1981) which compared the effectiveness of (a) a summary after a text with (b) a summary at the beginning of a text and (c) no summary at all. The (a) summary after a text condition yielded the best text recall. It was assumed that the absence of a positive effect for the (b) summary at the beginning condition might be due to the fact that readers overlooked the summary. Consequently, he conducted a study where he subdivided the condition into summary at the beginning on same page and summary at the beginning on a different page. Summary at the beginning on a different page and summary at the end showed significantly higher text recall than summary at the beginning on same page and no summary at all condition.

In view of these considerations, Hartley and Trueman (1982) conducted five successive empirical studies on how the placement of a summary affects retention and recall. The general finding was that summaries enhanced retention for summarized content. No significant difference between the placement, at the end or beginning, could be assessed. It can be concluded that similar

effects can be expected for the effectiveness of multimedia-based reviews as they are the digital equals to text-based summaries.

### *2.1.2 Reviews in software-training*

Reviews can be considered as summarizations of the main steps for task completion as well as summarizations of main ideas for concepts. One recent study tested the effectiveness of reviews in IVs for the procedure to conduct a t-test with SPSS (Brar & van der Meij, 2017). The study compared a review condition with a no review condition and found no significant effect in favor of the review condition on a conceptual knowledge and SPSS performance test.

On the other hand, there is empirical evidence favoring the inclusion of reviews for software training in Microsoft Word. Two experiments yielded direct support for the inclusion of reviews (van der Meij & van der Meij, 2016 a,b). In both studies, a significant effect in favor of the inclusion of a review was found. Furthermore it was shown that the usage of reviews improved the self-efficacy of the participants.

Up to now, it seems unclear if and under which circumstances reviews are an effective feature to enhance motivation and learning outcomes. The sparse and contradictory results of the review experiments call for further investigation.

## **2.2 Practice**

According to the DBT-model, a practice component could have an influence on the (re)production process after watching a demonstration which in turn might have a positive impact on the whole learning outcome. The advantage of practice after giving instructions is that the user is stimulated to (re)produce contents and processes in order to deepen his or her understanding. This stimulates the learner to construct meaning and could therefore strengthen learning. By engaging in practice, students are able to apply knowledge through interaction with the learning material and connect with the information on a deeper level.

Empirical evidence about the effectiveness of practice after video demonstration for software training is sparse and ambiguous. Accordingly, the research was extended to a similar field, worked-examples. Again, insights from research in this field can be transmitted to the possible effectiveness of practices in IVs.

### *2.2.1 Worked-examples*

A worked-example provides an expert solution model to a problem and gives a step-by-step explanation on how to solve the problem. This is done by drawing the learners' attention to key features in a problem and provide them with task-specific information (Atkinson, Derry, Renkl, & Wortham, 2000). Many worked-examples also contain a practice component on a similar problem with the same design: First, students receive procedural information on how to solve a problem (the worked-example) and then engage in a practice component afterwards on a similar problem.

Many studies investigated the placement of the practice component in worked-examples research. Empirical studies generally suggest that practice after demonstration increases learning



for novices. In a study published by Reisslein, Atkinson, Seeling and Reisslein (2006) participants with low prior knowledge performed better with practice after and participants with high prior knowledge did better with practice before worked-examples in electric suits.

Wouters, Paas and van Merriënboer (2010) did research on the influence of practice with animated models for problem solving in probability calculus. Participants were divided into three conditions: (a) practice after worked-example (b) practice before worked-example and (c) restudy of worked-example. Learning outcome was measured by trained and transfer tasks. No significant difference between the conditions could be found. The lack of the positive effect for condition (a) practice-after worked-example was explained by the fact that the participants had relatively high prior knowledge.

In a study published by van Gog, Kester and Paas (2011) participants received four electrical troubleshooting tasks. The researchers compared (a) example only (b) practice only (c) example with practice-after and (d) example with practice-before. The results reveal a significantly higher score for the (a) example only and (c) example with practice after condition on an immediate post-test. No difference between condition (a) and (c) was found.

In conclusion, most studies found a positive effect for engaging in practice after demonstrating the example. A few studies found that studying the example only is equally effective as engaging in practice afterwards. The following section discusses the role of practice in the field of video based software training.

### *2.2.2 Practice in software-training*

In an experiment conducted by Ertelt (2007), one group of participants engaged in practice after they watched five demonstration videos on how to use the software program RagTime, a desktop publishing program. The control group did not engage in practice after watching the videos. The results showed a small but significant effect in favor of the practice condition on an immediate and delayed post-test. Furthermore, practice had a positive effect on a transfer test. According to the researcher, the inclusion of practice encouraged users to engage in more active and deeper processing.

In another recent experiment on practice with video-based software training, participants watched videos on formatting tasks in Microsoft Word (van der Meij, Rensink & van der Meij, 2018). There were two experimental conditions where the timing of practice varied (practice-video; video-practice) and one control condition with no practice at all. It was expected that in the condition where the practice preceded the video, the practice would have a motivational effect, leading to an increase in motivation within the participants to study the video. In addition, it was expected that the condition where the video was followed by the practice the highest learning outcome would occur. Both assumptions were not confirmed. The control group (video only) had comparable learning outcomes to the experimental conditions on an immediate post-test, delayed post-test and transfer test.

Van der Meij (2018) conducted another experiment to study the effectiveness of practice. The author used the same conditions as in the above mentioned experiment but added a fourth condition, where practice was preceded by the demonstration as well as followed by the

demonstration (practice-video-practice). In the experiment the participants had the task to format a Microsoft Word page. The results showed that engaging in practice increased training time and led to more negative mood states during training. In addition, the control group (video only) had comparable learning outcomes to the practice groups on an immediate and delayed post-test. The only clear advantage for the practice groups concerned the transfer test. Surprising results were found for the practice-video-practice condition. It was expected that the highest learning gain would be found in this condition, but instead this condition had the lowest performance scores on a practice test and the immediate post-test.

However, the results of the above described experiments are not sufficient to claim that the inclusion of a practice component is needless or even counterproductive. Rather, it suggests that practice is a more complex design issue than initially thought. Therefore, this study extends the existing empirical research by focusing on the effects of a practice component with a different content and target group.

### 3. RESEARCH MODEL & DESIGN

Based on the DBT model for software-training the following research model is derived for this study (Fig. 2). The model shows the proposed impact of the two independent variables review and practice on the two dependent variables motivation and learning outcome. The current study applied a 2x2 between-subject design comparing four conditions. In the first condition, a review as well as a practice component followed the video demonstration (review-practice). The second condition included a review but no practice component (review). In the third condition a practice component was included but no review (practice) and the fourth condition functioned as a control condition where neither a review nor a practice component was included (control). The setting in which this study was conducted limited the research questions for which the data could be gathered to the following research questions:

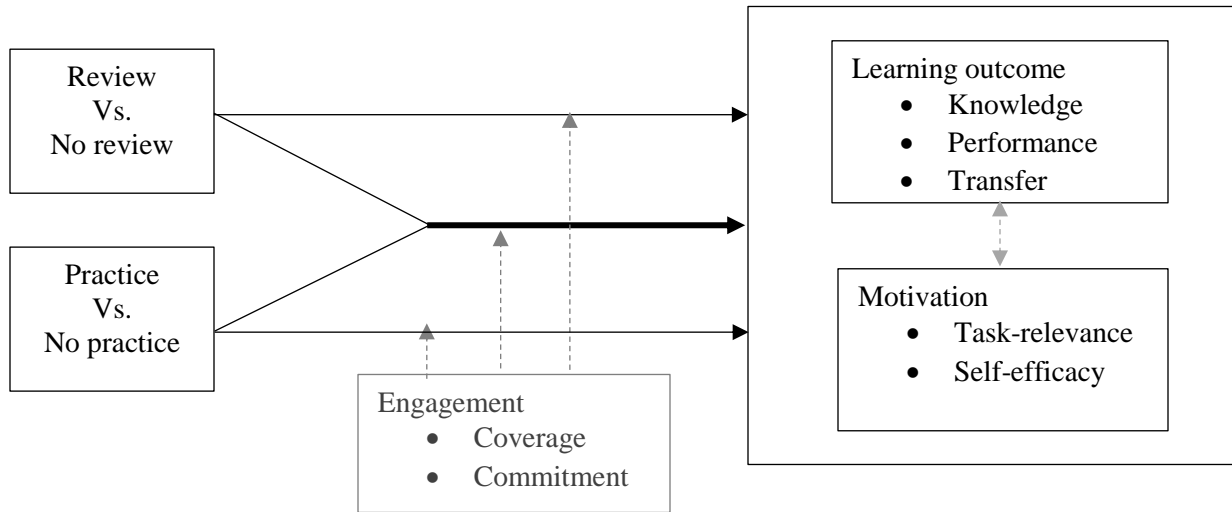


Fig. 2. Research Model

#### 1. To what extent are the videos engaging?

To investigate video viewing, the present study measured student engagement by coverage and commitment. Coverage refers to the number of seconds the video was set into play mode at least once (unique seconds). Commitment refers to the total number of seconds the video was set into play mode (total seconds; Brar & Van der Meij, 2017). It thus measures the total number of seconds when the play mode was activated more than once. Both measures are expressed as a percentage of the total length of the video. Engagement may be related to the extent to which the independent variables affect the dependent variables in this study.

*2. Is there a difference in students' learning outcome among the conditions?*

As one objective in this study is to investigate under which circumstances IVs are effective, the learning outcome was assessed. Learning outcome was measured on the basis of three tests: A knowledge test, an SPSS performance test and an SPSS transfer test. It is expected that the highest learning outcome will be in the condition where both a review and a practice component are present. Whereas a review could enhance the retention process, a practice might enhance the (re)production process.

*3. Is there a difference in student motivation among the conditions?*

As motivation is a stimulating factor behind learning processes, another objective in this study is to examine whether there is a difference in participants' motivation among the four conditions. Motivation was assessed by focusing on the perceived self-efficacy and task-relevance of the participants. It is expected that the condition where both a review and a practice component are present yields the highest self-efficacy because students may feel more confident in solving a subsequent task. Further, no difference among the conditions regarding task relevance is expected because a review or a practice component may not have an effect on the significance of the actual task.

*4. Is there a relationship among the dependent variables?*

In addition, the present study wants to investigate whether there is a relation among participants engagement, learning outcome and motivation. It is expected that a positive score in students' engagement and self-efficacy also results in a more positive learning outcome, resulting in a positive relationship between those measures.

## **4. METHOD**

### **4.1 Participants**

The participants consisted of pre-Master students who were enrolled in an introductory statistics course at a university in the Netherlands ( $n=70$ ). The sample was 37,1 percent male and 62,9 percent female. 38,6 percent ( $n=27$ ) of the participants did the pre-master for International Business Administration, 25,7 percent ( $n=18$ ) were enrolled in Psychology, 17,1 ( $n = 12$ ) percent in Educational Science and Technology, 12,9 percent ( $n=9$ ) in Communication Science and the remaining 5,7 percent ( $n=4$ ) in other degree programs. The participants were between 21 and 43 years old with a mean age of 24,2 years ( $SD= 3,9$  years).

The participants were randomly assigned and evenly distributed to the four conditions. Participation was voluntary and the students were told that they could stop with the experiment at any time if they felt uncomfortable. In addition, they were told that participation would probably prepare them well for an upcoming exam. All students who completed the experiment received a 10 € payment.

### **4.2 Instructional materials**

The videos in this study focused on descriptive statistics. The video content was especially tailored for one unit of an introductory statistics course at a university in the Netherlands. The content explained the meaning and calculation of different concepts and how to compute them with SPSS. The textbook “Discovering Statistics using IBM SPSS Statistics” by Field (2013) and “The Practice of Social Research” by Babbie (2015) served as a foundation. All videos were created with the help of the software program Camtasia.

Due to the diverse and complex content of descriptive statistics, the material was divided into four parts. The first video (3:58) focused on measures of central tendency, the mean and median. Video two (3:22) focused on measures of dispersion, quartiles and interquartile range. The third video (3:02) paid attention to boxplot and outliers and the fourth (3:00) to variance and standard deviation. For all conditions, stimulus material was held constant

All videos followed the same structure. First, the meaning of the concept in question was explained on the basis of a real-life example. The real-life example used for all videos was a data set of exam scores. The topic of exam scores was considered to be an interesting and relevant topic for students (Merrill, 2002). Second, the mathematical procedure of how to calculate the different concepts was demonstrated. Next, a demonstration of how to compute the concepts with the software program SPSS was performed. Finally, an explanation of how to interpret the SPSS output was given. The IVs taught declarative knowledge by explaining the definition of the concepts and by mentioning the categories to which the concepts belong. Procedural knowledge was taught by demonstrating the procedure of how to calculate the concepts with pencil and paper and how to compute them with SPSS.

SPSS provides multiple solution methods to compute a task. Users have the possibility to use the menu for creating the syntax of a statistical procedure or to write the syntax themselves.

According to Renkl (2014), tutorials should present only a single method and it is advised to choose the easiest one. Therefore, all IVs presented a single, menu-based method,

The adopted DBT model for software-training addresses several features and techniques to support processes of attention, retention, reproduction and motivation (Fig. 1). This study pays close attention to the effect of a review and a practice component. How the features and techniques other than review and practice were implemented in the IVs is explained in Appendix A.

#### 4.2.1 Review in IVs

A review video was presented as a stand-alone video at the end of the four IVs and functioned as a recapitulation of the key conceptual and procedural information. The total length of the review was 168 seconds. On the review's opening screen the word "review" was written so that it was clear for the audience what followed (Fig. 3). The opening screen was presented for 3 seconds.

The content of the four statistic videos was reviewed in consecutive order. For each concept, its definition appeared under the name of the concept. Shortly after that, the mathematical calculation was visualized step by step on the screen (Fig. 4). The visualization of the mathematical calculation was complemented with an audible narrative which repeated the procedure. The narrative instructions in the review video were formulated in a way to align with the viewer's presumed mental rehearsal, meaning they were personalized to take an "I" perspective (e.g.: "To calculate the mean, I add up all scores and divide the sum by the number of scores."). Signaling techniques were used to draw the user's attention to relevant parts of the screen. After that, the procedure on how to compute the given concepts with SPSS was demonstrated by presenting a recorded screen cast of the computing (Fig. 5). The demonstration was complemented with an audible narrative (e.g. "In SPSS I click on analyze ... descriptive statistics ... frequencies ... move the variable to the right box and click on statistics ... in the new window I tick mean and median and click on continue and then ok.")



Fig. 3. Screenshot from the review video's opening

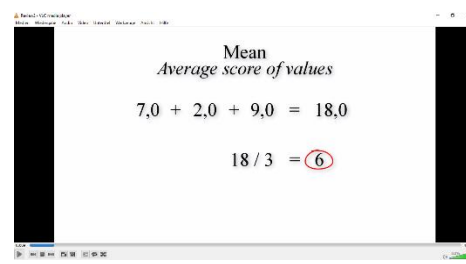
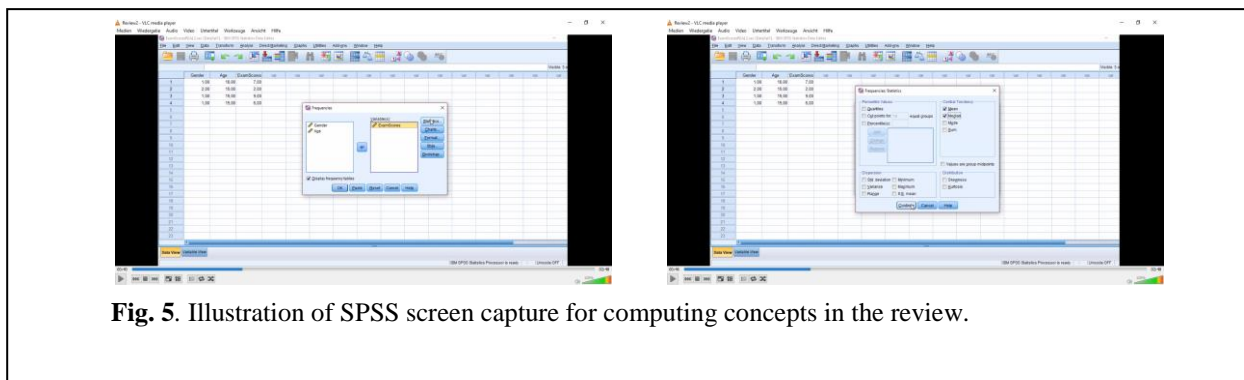


Fig. 4. Screenshot of how conceptual and procedural knowledge was reviewed in the review video



**Fig. 5.** Illustration of SPSS screen capture for computing concepts in the review.

#### 4.2.2 Practice in IVs

The practice component was presented randomly to half of the student's right after they were engaged in the IVs. These participants received the task description that the party commission of their university plans the next campus party and therefore the commission wants to find out about the beer consumption of the students. Consequently the participants had to solve six SPSS tasks regarding the beer consumption of a fictional data set of 50 students. To answer the questions, the participants had to provide the SPSS output as well as stating the value in question in a box below. To make the fictional data set more realistic, age, gender, education and nationality were also included.

The practice tasks were similar tasks to the SPSS performance test which followed the practice. The order of the tasks and the data set differed from the SPSS performance test. The practice was written on a separate word document which the students received via e-mail

(Appendix B). The fictional data set was also sent to the students via e-mail. The students had the option to receive feedback which was provided on the website where the other parts of the study were implemented (Fig. 6). The participants were told to consult the feedback or solve the tasks on their own. When the participants completed the practice they were told to press an arrow located below the feedback page. After clicking on the arrow, the participants were no longer able to consult the feedback for the practice.

**Question 1: What is the average number of beer glasses the students drank on their last party?**  
 --> Asked for the average, you have to determine the mean of your data.  
 1. Click on Analyze --> Descriptive statistics --> Frequencies...  
 2. Move the Glassesofbeer variable to the right box and click on the statistics button on the right.  
 3. In the new window you mark "Mean", click on continue and then ok.  
 SPSS output:

**Frequencies**

Statistics

| Glassesofbeer |         |        |
|---------------|---------|--------|
| N             | Valid   | 50     |
|               | Missing | 0      |
| Mean          |         | 5,7400 |

**Question 2: What is the number of glasses of beer for the student(s) on the 50th percentile?**  
 --> Asked for the 50th percentile, you have to determine the median of your data.  
 1. Click on Analyze --> Descriptive statistics --> Frequencies...  
 2. Move the Glassesofbeer variable to the right box and click on the statistics button on the right.  
 3. Mark "Median" and click on continue and then ok  
 SPSS output:

**Frequencies**

Statistics

| Glassesofbeer |         |        |
|---------------|---------|--------|
| N             | Valid   | 50     |
|               | Missing | 0      |
| Median        |         | 6,0000 |

**Fig. 6.** Illustration of how feedback was provided during the practice

## 4.3 Instruments

### 4.3.1 SPSS practice

In the SPSS Practice the participants could practice the demonstrated computations on their own (Appendix B). The SPSS practice contained six items serving similar tasks presented in the demonstrations (*1. What is the average number of glasses of beer the students drank at their last party?; 2. What is the number of glasses of beer for the student(s) on the 50<sup>th</sup> percentile? 3. What is the value of the 10<sup>th</sup> percentile? 4. What is the range of values for the interval between the 1<sup>st</sup> and 3<sup>rd</sup> quartile? 5. Make a boxplot from the number of glasses of beer. 6. Find the two values which indicate how spread out around the mean the scores are*). The practice tasks were given on a separate Word document which the participants received via e-mail.

To answer the questions, participants had to provide the SPSS output with the answer to the question as well as explicitly stating the value in question in a box below the SPSS output. 0 points were given if the stated value in the box was incorrect and the SPSS output did not contain the answer to the question. 1 point was given if only the provided SPSS output contained the answer or the stated value was correct. 2 points were given if the provided SPSS output as well as the stated value in the box were correct. The maximum score for the SPSS Practice was 12 points.

### 4.3.2 SPSS performance test

In the SPSS performance test students were asked to compute the trained SPSS analyses of the demonstrations without the option to consult feedback (Appendix C). The performance test consisted of six questions covering content from all four tutorials (*1. Make a boxplot from the Facebook-Friends; 2. Find the two values which indicate how spread out around the mean the scores are; 3. What is the average number of Facebook-Friends for these students?; 4. What is the value of the 10<sup>th</sup> percentile?; 5. What is the range of values for the interval between the 1<sup>st</sup> and 3<sup>rd</sup> quartile?; 6. What is the number of Facebook-Friends for the student(s) on the 50<sup>th</sup> percentile?*) The six items tested the participants' procedural knowledge.

The SPSS performance test was given on another separate Word document which the students received via e-mail. The scoring of the SPSS performance test was similar to the scoring of the SPSS practice with a score range from 0 – 2 for each item making a highest score of 12 points. The order of the items and the data set differed to the one provided in the SPSS practice. The data was a fictional set of 50 students with information about the number of Facebook-friends they have. The topic was chosen because the data set included values different from the values used in the demonstrations (e.g. 100-900 in the test compared to 2-9 in the demonstration). To make the data more realistic, the set contained also information about age, gender, nationality and education of the fictional group of students.

### 4.3.3 SPSS Transfer test

Three untrained items were also added to the SPSS performance test (Appendix C). The goal of the transfer test was to measure if the participants were able to transmit knowledge from one concept to a totally new concept which was not explained in the demonstration. In the SPSS transfer test participants were asked to compute three analyses on their own (*1. Mode is another*



*measurement of central tendency. The mode is the most frequently occurring value in a data set. Determine the mode of your data set using SPSS; 2. Range is another measurement of dispersion. The range is the distance between the highest and lowest score within a data set. Determine the range of your data set using SPSS; 3. Another graphical representation of data is a histogram. Create a histogram of the Facebook-friends using SPSS.)* The scoring of the transfer test was similar to the scoring of the SPSS practice and SPSS performance with a highest score of 6 points.

#### *4.3.4 Knowledge test*

Declarative knowledge was assessed using 11 multiple choice questions with four alternatives (Appendix D). Participants received one point for each correct answer making for a highest score of 11 points. The 11 items covered content from all four videos.

#### *4.3.5 Motivation Questionnaire*

Students also received a short motivation questionnaire (Appendix E). The motivation questionnaire consisted of 5 items measuring self-efficacy and 4 items measuring task relevance on a 7-point Likert scale. The questionnaire was an altered version of the Motivated Strategies for Learning Questionnaire (MSLQ) by Pintrich, Smith, Garcia and McKeachie (1991), which was especially tailored for this study (see Appendix B). An example of an item measuring self-efficacy was “I now know how to make a boxplot with SPSS”. An example of an item measuring task relevance was “I think that SPSS is relevant for my study.” The answer possibilities ranged from strongly disagree, to strongly agree. The reliability analysis for the four task relevance items yielded Cronbach’s  $\alpha = 0.67$  and for the five self-efficacy items a Cronbach’s  $\alpha = 0.69$ .

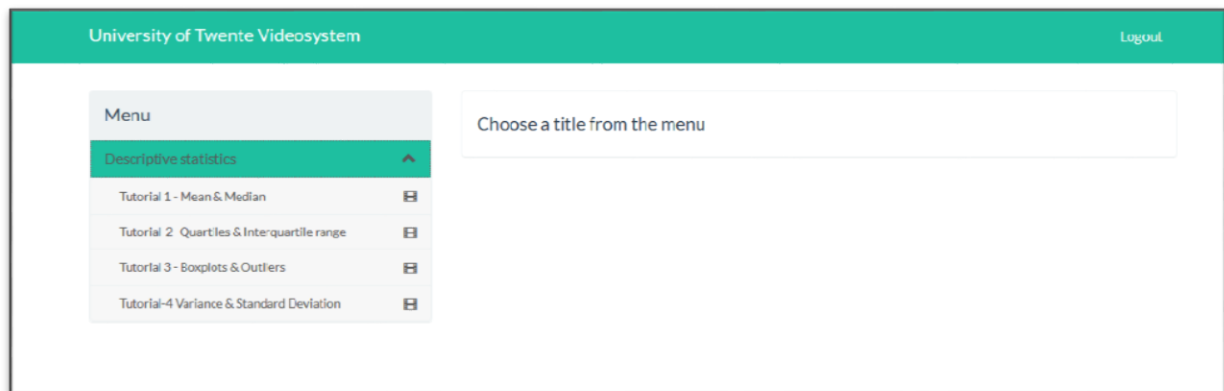
### **4.4 Procedure**

All students enrolled in the course were informed beforehand by their teacher that an SPSS pre-training for an upcoming course unit had been created. All students then received an e-mail to which they had to reply if they wanted to participate in the SPSS pre-training (see Appendix F).

All students who replied to the first e-mail then received another individual e-mail containing a link to the study and an individual log-in code for the website where the IVs were presented. After clicking on the link the students were first remitted to the website containing the IVs (Fig 12.) The website where the IVs were uploaded measured the viewing time for each participant in seconds automatically. The log-in codes randomly directed half of the students to the website including the additional review video and the other half to the website where the review video was not implemented.

Half of the students then engaged in practice. The practice was a Word file containing the practice items which were attached to the e-mail. After the practice, the SPSS performance test was presented. For the students who did not engage in the practice, the SPSS performance test was presented just after the demonstrations. After the SPSS performance test the multiple-choice knowledge test and motivation questionnaire were presented. At the end, they were informed to send the filled in SPSS performance test and SPSS practice back to the researcher. The results of

the knowledge test and motivation questionnaire were automatically stored online. One week after the pre-training, the students were informed via E-mail where they could pick up their payment.



**Fig. 7.** Screenshot of the website where the videos were presented

## 4.5 Data Analysis

To measure the video engagement of the participants, coverage (unique seconds) and commitment (total seconds) of the videos were analyzed separately for each condition. First, the mean for coverage and commitment was calculated for each condition and indicated as a percentage of the total length of the video. To get a first overview whether there was a difference among the conditions regarding the dependent variables, the descriptive statistics were calculated.

To assess the learning outcome for each condition, the sum for the SPSS practice, SPSS performance test and SPSS transfer test for each participant was calculated. In addition, the multiple-choice Knowledge test was re-coded in the way that participants received one point for each correct answer. Furthermore, the means score for the task-relevance and self-efficacy items were calculated.

After that, a bivariate correlation analysis between the engagement mean scores and the test scores was conducted. To measure the learning outcome, comparisons among the four conditions were tested with a two-sided Multivariate Analysis of Variance (MANOVA). To assess whether there was an effect of condition regarding motivation, another Multivariate Analysis of Variance (MANOVA) was conducted. For both MANOVA's, the underlying assumptions were tested and an effect size of  $\alpha = .05$  was chosen.

## 5. RESULTS

### 5.1 Engagement

To test how engaging the demonstration and review videos were, the mean percentages of coverage (unique seconds) and commitment (total seconds) in the four conditions were calculated. Table 1a presents the mean percentages for coverage in the different conditions, considering all four demonstration videos in combination. Additionally, the mean percentage for coverage of the review video in condition 1 and 2 is displayed. The table shows that coverage was very high with a mean score above 99 percent in each condition for the demonstration videos. Interestingly, the results also show that the mean score of coverage for the review video in condition 1 was merely 55.3 percent and in condition 2 51.8 percent. This indicates that on average participants watched almost each second of the four demonstration videos but only half of the additional review video.

**Table 1a.** Mean coverage\* (standard deviation) per condition for the demonstration and review videos

|                             | Demonstration videos Mean<br>(SD) | Review video Mean<br>(SD) |
|-----------------------------|-----------------------------------|---------------------------|
| Review*Practice<br>(N = 17) | 99,29%<br>(2,2)                   | 55,3%<br>(74,52)          |
| Review<br>(N = 18)          | 99,23%<br>(1,85)                  | 51,88%<br>(75,21)         |
| Practice<br>(N = 19)        | 99,28%<br>(1,71)                  | -                         |
| Control<br>(N = 16)         | 99,4%<br>(1,44)                   | -                         |

\*Note: a coverage score of 0% indicates that no single second of the video has been set in play mode, a coverage score of 100% indicates that each second in the video has been set in play mode.

Table 1b presents the mean percentage for coverage of the four demonstration videos separately. The table shows that each video in each condition was set almost completely into play mode with an excellent viewing time of at least 98 percent. The continuous results indicate that there is no difference in coverage among the conditions.

**Table 1b.** Mean coverage\* (standard deviation) per condition and instructional video

|                             | Video # 1<br>Mean<br>(SD) | Video # 2<br>Mean<br>(SD) | Video # 3<br>Mean<br>(SD) | Video # 4<br>Mean<br>(SD) |
|-----------------------------|---------------------------|---------------------------|---------------------------|---------------------------|
| Review*Practice<br>(N = 17) | 98,87%<br>(4,75)          | 99,62%<br>(1,35)          | 98,91%<br>(2,30)          | 99,87%<br>(0,38)          |
| Review<br>(N = 18)          | 99,67%<br>(1,23)          | 98,74%<br>(2,52)          | 98,61%<br>(2,82)          | 99,78%<br>(0,82)          |
| Practice<br>(N = 19)        | 99,21%<br>(2,38)          | 99,14%<br>(2,15)          | 98,99%<br>(1,77)          | 99,81%<br>(0,52)          |
| Control<br>(N = 16)         | 99,63%<br>(1,42)          | 99,13%<br>(1,84)          | 98,98%<br>(2,03)          | 99,83%<br>(0,48)          |

\*Note: a coverage score of 0% indicates that no single second of the video has been set in play mode, a coverage score of 100% indicates that each second in the video has been set in play mode.

Table 2a presents the mean percentage for commitment in the different conditions, considering all four demonstration videos in combination and the review videos. Due to the percentage of above 100 percent in each condition, the results indicate that some sections of the four demonstration videos were set into play mode more than once. The review videos' percentage indicates that some sections of the additional video were not set into play mode at all.

**Table 2a.** Total mean commitment score from the four videos combined and review per condition.

|                             | Total Mean<br>(SD)  | Review Mean<br>(SD) |
|-----------------------------|---------------------|---------------------|
| Review*Practice<br>(N = 17) | 130,1%<br>(82,95)   | 78,22%<br>(131,03)  |
| Review<br>(N = 18)          | 116,35%<br>(97,93)  | 64,21%<br>(109,51)  |
| Practice<br>(N = 19)        | 112,73%<br>(103,88) | -                   |
| Control<br>(N = 16)         | 121,73%<br>(87,13)  | -                   |

Note: A total score of above 100% indicates that some sections are viewed more than once.

Table 2b presents the mean percentage for commitment of the four demonstration videos separately. The table shows that on an average some sections of each video in each condition were set into play mode more than once. The continuous results indicate that there is no difference in commitment among the four conditions. Still, the mean percentage of commitment constantly increases with the demonstration videos, indicating that the more complex the content was, the longer the students watched the video.

**Table 2b.** Mean commitment (standard deviation) per condition and video

|                             | Video # 1<br>Mean<br>(SD) | Video # 2<br>Mean<br>(SD) | Video # 3<br>Mean<br>(SD) | Video # 4<br>Mean<br>(SD) |
|-----------------------------|---------------------------|---------------------------|---------------------------|---------------------------|
| Review*Practice<br>(N = 17) | 101,3%<br>(10,74)         | 130,39%<br>(85,59)        | 144,8%<br>(109,12)        | 152,97%<br>(126,33)       |
| Review<br>(N = 18)          | 102,03%<br>(29,14)        | 119,87%<br>(105,77)       | 121,91%<br>(117,23)       | 125,69%<br>(139,56)       |
| Practice<br>(N = 19)        | 105,36%<br>(22,73)        | 121,68%<br>(143,31)       | 112,25%<br>(121,87)       | 112,93%<br>(125,59)       |
| Control<br>(N = 16)         | 106,32%<br>(17,77)        | 128,51%<br>(110,42)       | 120,88%<br>(92,01)        | 135,32%<br>(128,32)       |

Note: A total score of above 100% indicates that some sections are viewed more than once.

## 5.2 Learning outcome

The descriptive statistics for the dependent variables per condition are displayed in Table 3. To test whether the two independent variables, review and practice have an effect on students learning outcome, a multivariate analysis of variance (MANOVA) was conducted. The results of the MANOVA are displayed in Table 4.

**Table 3.** Descriptive statistics of the dependent variables per condition

|                          | <u>Review*Practice</u> |      | <u>Review</u> |      | <u>Practice</u> |      | <u>Control</u> |      |
|--------------------------|------------------------|------|---------------|------|-----------------|------|----------------|------|
|                          | N = 17                 |      | N = 18        |      | N = 19          |      | N = 16         |      |
|                          | M                      | SD   | M             | SD   | M               | SD   | M              | SD   |
| <i>Practice*</i>         | 11.31                  | 1.62 | -             | -    | 11.21           | 1.08 | -              | -    |
| <i>Knowledge Test**</i>  | 8.06                   | 1.71 | 9.00          | 1.66 | 8.29            | 1.93 | 8.00           | 1.67 |
| <i>Performance Test*</i> | 11.59                  | 1.00 | 9.17          | 2.46 | 11.53           | 0.96 | 10.44          | 1.63 |
| <i>Transfer Test***</i>  | 5.00                   | 1.58 | 5.56          | 1.04 | 5.83            | 0.38 | 5.56           | 0.63 |
| <i>Motivation****</i>    |                        |      |               |      |                 |      |                |      |
| Task-relevance           | 5.97                   | 0.64 | 5.89          | 0.60 | 6.09            | 0.67 | 5.72           | 0.63 |
| Self-efficacy            | 5.62                   | 0.62 | 5.61          | 0.68 | 5.85            | 0.53 | 4.91           | 0.70 |

\* 6 items measured with 0 = false answer, 1 = half correctly and 2 = total correctly (highest score = 12 points)

\*\* 11 items measured with 0 = false answer and 1 = right answer (highest score = 11 points)

\*\*\* 3 Items measured with 0 = false answer, 1 = half correctly and 2 = total correctly (highest score = 6 points)

\*\*\*\* 9 items measured on a 7-point-Likert scale

There was a statistically significant difference in student learning outcome for exposition to a practice component,  $F(60,000) = 9.00$ ,  $p < .01$ , Wilk's  $\Lambda = .690$ ,  $\eta_p^2 = .310$ . In particular, a practice component outperformed no practice component on student scores in the performance test,  $F(1,66) = 20.41$ ,  $p = .000$ , 95% CI [1.020 ; 2.638],  $\eta_p^2 = .248$ . With a strong effect size this result indicates that including a practice component in an IV considerably increases student scores on an SPSS performance test.

In addition, there was an interaction effect between review and practice on student's learning outcome,  $F(60,000) = 3.68$ ,  $p < .05$ , Wilk's  $\Lambda = .845$ ,  $\eta_p^2 = .155$ . However, the test of the between-subject effects showed no significant difference on student scores on the three different tests when exposed to both a review and a practice component.

**Table 4.** Results of the two-way MANOVA for learning outcome

| Dependent variables     |             | F     | df   | Sig.              | $\eta_p^2$ |
|-------------------------|-------------|-------|------|-------------------|------------|
| <b>Learning outcome</b> |             |       |      |                   |            |
| Knowledge test          | Review      | 0.53  | 1,66 | .470              | .248       |
|                         | Practice    | 0.35  | 1,66 | .558              |            |
|                         | Interaction | 2.60  | 1,66 | .112              |            |
| Performance test        | Review      | 2.81  | 1,66 | .099 <sup>+</sup> |            |
|                         | Practice    | 20.41 | 1,66 | .000**            |            |
|                         | Interaction | 2.51  | 1,66 | .118              |            |
| Transfer test           | Review      | 2.38  | 1,66 | .128              |            |
|                         | Practice    | 0.44  | 1,66 | .510              |            |
|                         | Interaction | 2.70  | 1,66 | .105              |            |

Note: \*  $p < .05$ , \*\*  $p < .01$ , +  $p < .10$

### 5.3 Motivation

To find an answer to the question whether review and/or practice have an effect on the motivation of the participants, another multivariate analysis of variance (MANOVA) was conducted. The results are displayed in Table 6.

There was a statistically significant difference in motivation whether exposed to a practice component in an IV or not,  $F(61,000) = 5.45$ ,  $p < .05$ ; Wilk's  $\Lambda = .877$ ,  $\eta^2_p = .123$ . The test showed that including a practice component leads to a significantly higher scores in student self-efficacy compared to not including a practice component,  $F(1,66) = 8.05$ ,  $p < .01$ , 95% CI = [.127 ; .733],  $\eta^2_p = .115$ . With a strong effect size the result shows that letting students practice the tasks after watching a demonstration increases their perceived self-efficacy. Besides this, there was an interaction effect of review and practice on student self-efficacy,  $F(1,66) = 11.07$ ,  $p < .01$ , 95% CI = [.141 ; .872],  $\eta^2_p = .151$ . With a strong effect size the result indicates that including both a practice as well as a review component increases students' perceived self-efficacy. Regarding the manipulations of the IVs, no significant differences of students' perceived task-relevance between the four conditions was found.

**Table 6.** Results of the two-way MANOVA for motivation

| Dependent variables |             | F     | df   | Sig.   | $\eta^2_p$ |
|---------------------|-------------|-------|------|--------|------------|
| <b>Motivation</b>   |             |       |      |        |            |
| Task-relevance      | Reviewyn    | 0.06  | 1,66 | .811   |            |
|                     | Practiceyn  | 2.40  | 1,66 | .126   |            |
|                     | Interaction | 0.64  | 1,66 | .426   |            |
| Self-efficacy       | Reviewyn    | 1.65  | 1,66 | .204   |            |
|                     | Practiceyn  | 8.05  | 1,66 | .006** | .115       |
|                     | Interaction | 11.07 | 1,66 | .001** | .151       |

Note: \*  $p < .05$ , \*\*  $p < .01$ , +  $p < .10$

### 5.4 Relation among dependent variables

To assess the relation among the dependent variables a bivariate correlation analysis was conducted. The results can be found in Appendix G.

The results revealed that there was a strong statistical correlation between the scores on the practice and the performance test,  $r(33) = .59$ ,  $p < .01$ . The higher the participants scored on the practice, the higher they scored on the SPSS performance test. Further, the scores on the knowledge test correlated strongly with self-efficacy,  $r(66) = .37$ ,  $p < .01$ . Scores on the performance test correlated with the total viewing time of video 3,  $r(65) = .25$ ,  $p < .05$  and video 4,  $r(69) = .32$ ,  $p < .01$ . In addition, scores on the performance test correlated with self-efficacy,  $r(66) = .25$ ,  $p < .05$ . Besides this, self-efficacy had a strong correlation with task-relevance,  $r(66) = .33$ ,  $p < .01$ . For the transfer test, however, no statistically significant correlations between the transfer test, the engagement data and self-efficacy could be assessed.

In conclusion the correlations among the dependent variables show that participants who had a high perceived self-efficacy did better on all tests, except for the transfer test. Moreover, participants with high self-efficacy perceived the tasks as more relevant. In contrast, the longer the participants watched the videos did not lead to higher perceived self-efficacy. Further, the longer the participants watched the IVs did not generally lead to a higher scores on the tests. The only significant positive correlation could be found between the performance test and the commitment of video 3 and 4.

## 6. DISCUSSION

The present study's purpose was to contribute to the existing knowledge about learning via demonstration videos in software training. The focus was on the effectiveness of reviews and the role of practice in instructional videos given that this topic is of high relevance for students learning statistics. Using the DBT model as a foundation, this study looked at the potential optimization of IVs for increasing student motivation and learning outcomes. The measured user logs revealed excellent viewing times by the participants, meaning that the IVs were very successful in gaining and maintaining the participants' attention.

The implemented features were derived from a study which adapted the DBT model to fit the aims of software-training (Brar & van der Meij, 2017). The used features in turn were derived from existing guidelines and frameworks in multimedia learning (Mayer, 2014; Koumi, 2013, van der Meij & van der Meij, 2013). The high engagement rate with regard to the demonstration videos used in this study supports the positive effectiveness of such features and techniques when learning from IVs in software-training.

Contrary to the excellent viewing time of the four demonstration videos, the measured user logs of the additional review video showed much lower viewing times. A reason for that could be that the four demonstration videos were adequately designed so that the participants saw no need in watching an additional review video. Another reason could be that due to the fact that the review video was a stand-alone video which served content from all four tutorials, the participants didn't want to search for the exact position they wanted to re-watch. Instead, they may just have re-watched certain sections of the demonstration. The discrepancy between coverage and commitment could also be an indication for this. For future research it might be important to ask the participants about their opinion towards the review video directly after the research in order to find out whether they perceived it as valuable or not. Compared to the total engagement of the students in the four demonstration videos, both coverage and commitment of the review video was lower. That means that students may have watched the review video less intense than the demonstration videos. The statistical power of the reviews' effectiveness may therefore be limited. Consequently, the related statistical results should be interpreted with caution and not be overly generalized to a broader context.

Next to the assessment of student engagement, this study investigated whether there was a difference in their learning outcomes with respect to the four conditions. This research found no evidence that the inclusion of a review component has a positive effect on student's knowledge, performance and ability to transfer. One possible explanation might be the low viewing time of the review videos. This finding aligns with the outcome of Brar and Van der Meij (2017). It can be concluded that a review is not a necessity in order to optimize IVs in software training as long as the actual video demonstrations are adequately designed.

In contrast, this study found evidence that the inclusion of a practice component increases student's scores on the SPSS performance test. Still, a practice component does not increase their knowledge or ability to transfer. As the viewing time among the conditions was similar, it can be assumed that those effects can solely be traced back to students practicing the tasks. On the one



hand, this finding aligns with the result found by Ertelt (2007) where a positive effect for the inclusion of a practice on an immediate and delayed post-test was found. On the other hand, the finding contrasts with Ertelt's (2007) results as the present study could not find a positive effect of a practice on a transfer test. In addition, the findings from the present study contrast with the outcome of a study showing no positive contribution of a practice on subsequent tests with Microsoft Word (Van der Meij, 2018; Van der Meij, Rensink & van der Meij, 2018). However, the present study differs from these two studies, being of higher complexity for the task domain. The videos in the present study targeted statistics which can be considered a more complex issue than formatting a Microsoft Word page. The conclusion that can be drawn is that the more complex the demonstrated content is, the more beneficial is the inclusion of a practice component afterwards. In particular, a practice component could facilitate the probability of a successful (re)production process and in turn enhance student task performance when learning statistics from IVs

Another objective of this study was to assess whether there is a difference in student motivation among the four conditions. As expected, the inclusion of a practice and a review component has no influence on how students evaluate a task. Furthermore, this study found no evidence that a review component enhances motivation in terms of self-efficacy. However, this study found that a practice alone and a practice in combination with a review component enhances students' self-efficacy strongly. This outcome stands in contrast with the findings of another recent study in software training where the exposure to a practice component resulted in more negative mood states of the participants (Van der Meij, 2018). The result is in accordance with Van der Meij and Van der Meij (2016; a,b) where the inclusion of a review significantly improved the perceived self-efficacy of the participants. Apparently, if students are exposed to both, a review and a practice component their perceived self-efficacy improves even more strongly.

Beyond the single and interactive effects of review and practice, this study examined whether there is a relationship among the dependent variables. The findings showed that the students' perceived self-efficacy has a positive relationship with their learning outcome. More precisely, the higher their self-efficacy, the higher were their scores on the knowledge and performance test. It can be concluded that combining a review with a practice component in an IV could indirectly effect student learning outcome by raising their perceived self-efficacy. .

## **6.1 Limitations and future research**

Although this study delivered some meaningful findings, it is not without limitations and offers implications for future research. A first limitation is the experimental design the participants were exposed to which could have had an influence on how they processed the demonstrated information. In addition, the sample consisted of 70 (44 female; 26 male) students from one single statistics course of a university in the Netherlands distributed over four conditions. Future research might include a more diverse set of students, a larger sample size with an even gender distribution and participants from various universities in order to generalize the results for a broader context.

Another limitation considers the number of items that were used to measure performance and the ability to transfer. The number of items for the performance test (six items) and transfer

test (three items) can be considered as small. To answer each item the participants had to engage themselves with the software program SPSS, a software the students were not used to. As it may have taken the participants a certain amount of time to answer each item, the goal was to include not too many items to avoid a high drop-out rate. However, future research might include a larger number of items for each test to get a clearer picture of the learning outcome.

The IVs used in this study focused on descriptive statistics which can be considered as the basics in statistics. This means that the findings can only partly be transmitted to the whole statistics topic. Future research might test the effectiveness of reviews and the role of practice by focusing on a broader range of statistical computing's and calculations.

The user logs which recorded the video coverage (unique seconds) and video commitment (total seconds) measured the seconds the videos have been set into play mode. Of course, this is no guarantee that the viewers actively watched each second of the videos. To record actual viewing times, measurements like eye-movement records are needed. This is a complex and costly measurement which was not accessible during the present study.

## **6.2 Conclusion**

This research intended to identify the effectiveness of reviews and the role of practice in software-training. The study extended the previous work on multimedia-based learning by providing insights about the circumstances under which reviews and practice are effective when learning statistics from IVs. A review alone does not have a positive effect on learning outcome and motivation as long as the IVs are adequately designed. In contrast, when students engage in a practice after watching IVs it has a strong effect on a subsequent trained performance task. When students are exposed to both a review and a practice component this study revealed that the combination of these two has a strong effect on the students' perceived self-efficacy which in turn could have a positive influence on the learning outcome.

## References

- Alexa Top 500 Global Sites. (2018). Alexa.com. Retrieved 16 January 2018, from <http://www.alexa.com/topsites>
- Alpay, E., & Gulati, S. (2010). Student-led podcasting for engineering education. *European Journal of Engineering Education*, 35(4), 415–427. doi:10.1080/03043797.2010.487557
- Atkinson, R. K., Derry, S. J., Renkl, A., & Wortham, D. (2000). Learning from examples: Instructional principles from the worked examples research. *Review of Educational Research*, 70(2), 181-214. doi:10.3102/00346543070002181
- Babbie, E. R. (2015). *The practice of social research* (14<sup>th</sup> ed.). Boston, MA: Nelson Education.
- Bandura, A. (1986). *Social foundations of thought and actions: A social cognitive theory*. Englewood Cliffs, NJ: Prentice Hall.
- Baglin, J., & Da Costa, C. (2014). How do students learn statistical packages? A qualitative study. In *Topics from Australian Conferences on Teaching Statistics* (pp. 169-187). Springer, New York, NY.
- Brar, J., & van der Meij, H. (2017). Complex software training: Harnessing and optimizing video instruction. *Computers in human behavior*, 70, 475-485.
- Chew, P. K., & Dillon, D. B. (2014). Statistics anxiety update: Refining the construct and recommendations for a new research agenda. *Perspectives on Psychological Science*, 9(2), 196-208.
- Choi, H. J., & Johnson, S. D. (2005). The effect of context-based video instruction on learning and motivation in online courses. *The American Journal of Distance Education*, 19(4), 215-227.
- Christensen, C. M., & Stordahl, K. E. (1955). The effect of organizational aids in comprehension and retention. *Journal of Educational Psychology*, 46(1), 65-74.
- DeVaney, T. A. (2010). Anxiety and attitude of graduate students in on-campus vs. online statistics courses. *Journal of Statistics Education*, 18(1), 1-15.
- Ertelt, A. (2007). *On-screen videos as an effective learning tool. The effect of instructional design variants and practice on learning achievements, retention, transfer, and motivation*. (Doctoral dissertation), Albert-Ludwigs Universität Freiburg, Germany.
- Farkas, D. K. (1999). The logical and rhetorical construction of procedural discourse. *Technical Communication*, 46, 42-54.

- Field, A. (2013). *Discovering statistics using IBM SPSS statistics* (4th ed.). London: Sage Publications.
- Fulford, C. P. (1992). *Systematically designed text enhanced with compressed speech audio*. Paper presented at the Annual Meeting of the Association for Educational Communications and Technology, Washington, D. C.
- Ginns, P., Martin, A. J., & Marsh, H. W. (2013). Designing instructional text in a conversational style: A meta-analysis. *Educational Psychology Review*, 25(4), 445-472.
- Gould, R. (2010). Statistics and the modern student. *International statistical review*, 78(2), 297-315.
- Guo, P. J., Kim, J., & Rubin, R. (2014). *How video production affects student engagement: An empirical study of MOOC videos*. Paper presented at the L@S '14, Atlanta, GA
- Hartley, J., & Davis, I. K. (1976). Preinstructional strategies: The role of pretests, behavioral objectives, overviews and advance organizers. *Review of educational research*, 46(2), 239-265.
- Hartley, J., & Trueman, M. (1982). The effects of summaries on the recall of information from prose: Five experimental studies. *Human Learning*, 1, 63 –82.
- Höffler, T. N., & Schwartz, R. N. (2011). Effects of pacing and cognitive style across dynamic and non-dynamic representations. *Computers & Education*, 57(2), 1716-1726.
- Kay, R., & Kletskin, I. (2012). Evaluating the use of problem-based video podcasts to teach mathematics in higher education. *Computers & Education*, 59(2), 619-627.
- Kosslyn, S. M., Kievit, R. A., Russell, A. G., & Shephard, J. M. (2012). PowerPoint® presentation flaws and failures: a psychological analysis. *Frontiers in psychology*, 3, 230.
- Koumi, J. (2013). Pedagogic design guidelines for multimedia materials: a mismatch between intuitive practitioners and experimental researchers. *Journal of Visual Literacy*, 32(2), 85-114.
- Lang, A., Park, B., Sanders-Jackson, A. N., Wilson, B. D., & Wang, Z. (2007). Cognition and emotion in TV message processing: How valence, arousing content, structural complexity, and information density affect the availability of cognitive resources. *Media Psychology*, 10(3), 317-338.
- Leopold, C., Sumfleth, E., & Leutner, D. (2013). Learning with summaries: Effects of representation mode and type of learning activity on comprehension and transfer. *Learning and Instruction*, 27, 40-49. doi:10.1016/j.learninstruc.2013.02.003
- Lloyd, S. A., & Robertson, C. L. (2012). Screencast tutorials enhance student learning of statistics. *Teaching of Psychology*, 39(1), 67-71.

- Margulieux, L., Guzdial, M., & Catrambone, R. (2012). *Subgoal-labeled instructional material improves performance and transfer in learning to develop mobile applications*. Paper presented at the 9th International Conference on International Computing Education Research (ICER'12), Auckland, New Zealand.
- Mathews, D., & Clark, J. (2003). Successful students' conceptions of mean, standard deviation, and the Central Limit Theorem. *Manuscript submitted for publication*, 1-12.
- Mayer, R. (2014). *The Cambridge handbook of multimedia learning* (2<sup>nd</sup> ed.). New York, NY: Cambridge University Press.
- Mayer, R. (2008). Applying the science of learning: evidence-based principles for the design of multimedia instruction. *American Psychologist*, 63(8), 760-769.
- Mayer, R. E., & Pilegard, C. (2014). Principles for managing essential processing in multimedia learning: Segmenting, pre-training, and modality principles. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd ed., pp. 316-344). New York, NY: Cambridge University Press.
- McLaughlin Cook, N. (1981). Summaries: Further issues and data. *Educational Review*, 33(3), 215-222. doi:10.1080/ 0013191810330305.
- Merkt, M., & Schwan, S. (2014). How does interactivity in videos affect task performance?. *Computers in Human Behavior*, 31, 172-181.
- Merrill, M. D. (2002). First principles of instruction. *Educational Technology Research & Development*, 50(3), 43-59. doi:10.1007/BF02505024
- Onwuegbuzie, A. J., & Wilson, V. A. (2003). Statistics Anxiety: Nature, etiology, antecedents, effects, and treatments--a comprehensive review of the literature. *Teaching in Higher Education*, 8(2), 195-209.
- Pintrich, P. R., Smith, D., Garcia, T., & McKeachie, W. (1991). *A Manual for the Use of the Motivated Strategies for Learning Questionnaire (MSLQ)*, The University of Michigan, MI
- Ploetzner, R., & Lowe, R. (2012). A systematic characterisation of expository animations. *Computers in Human Behavior*, 28(3), 781-794.
- Reichelt, M., Kämmerer, F., Niegemann, H. M., & Zander, S. (2014). Talk to me personally: Personalization of language style in computer-based learning. *Computers in Human behavior*, 35, 199-210.
- Reisslein, J., Atkinson, R. K., Seeling, P., & Reisslein, M. (2006). Encountering the expertise reversal effect with a computer-based environment on electrical circuit analyses. *Learning and Instruction*, 16, 92-103. doi:10.1016/j.learninstruc.2006.02.008

- Renkl, A. (2014). Toward an instructionally oriented theory of example-based learning. *Cognitive Science*, 38, 1-37. doi:10.1111/cogs.12086
- Rosen, M. A., Salas, E., Pavlas, D., Jensen, R., Fu, D., & Lampton, D. (2010). Demonstration-based training: A review of instructional features. *Human factors*, 52(5), 596-609.
- Schreiber, B. E., Fukuta, J., & Gordon, F. (2010). Live lecture versus video podcast in undergraduate medical education: A randomised controlled trial. *BMC medical education*, 10(1), 68.
- Spanjers, I. A., van Gog, T., Wouters, P., & van Merriënboer, J. J. (2012). Explaining the segmentation effect in learning from animations: The role of pausing and temporal cueing. *Computers & Education*, 59(2), 274-280.
- Traphagan, T., Kusera, J. V., & Kishi, K. (2010). Impact of class lecture webcasting on attendance and learning. *Educational Technology Research and Development*, 58(1), 19–37. doi:10.1007/s11423-009-9128-7
- Van der Meij, H. (2014). Developing and testing a video tutorial for software training. *Technical Communication*, 61(2), 110-122.
- Van der Meij, H., & Carroll, J. M. (1998). Principles and heuristics for designing minimalist instruction. In J. M. Carroll (Ed.), *Minimalism beyond the Nurnberg Funnel* (pp. 19-53). Cambridge, MA: MIT Press.
- Van der Meij, H. & van der Meij, J. (2013). Eight guidelines for the design of instructional videos for software training. *Technical Communication*, 60(3), 205-228.
- Van der Meij, H., & van der Meij, J. (2014). A comparison of paper-based and video tutorials for software learning. *Computers & education*, 78, 150-159.
- Van der Meij, H., & van der Meij, J. (2015). A test of the design of a video tutorial for software training. *Journal of Computer Assisted Learning*, 31(2), 116-132.
- Van der Meij, H., & van der Meij, J. (2016a). The effects of reviews in video tutorials. *Journal of Computer Assisted Learning*, 32(4), 332-344.
- Van der Meij, H., & van der Meij, J. (2016b). Demonstration-Based Training (DBT) for the design of a video tutorial for software instructions. *Instructional Science*, 44, 527-542.
- Van der Meij, H., Rensink, I., & van der Meij, J. (2018). Effects of practice with videos for software training. *Computers in Human Behavior*.
- Van der Meij, H. (2018). Cognitive and motivational effects of practice with videos for software-training. *Technical Communication*.

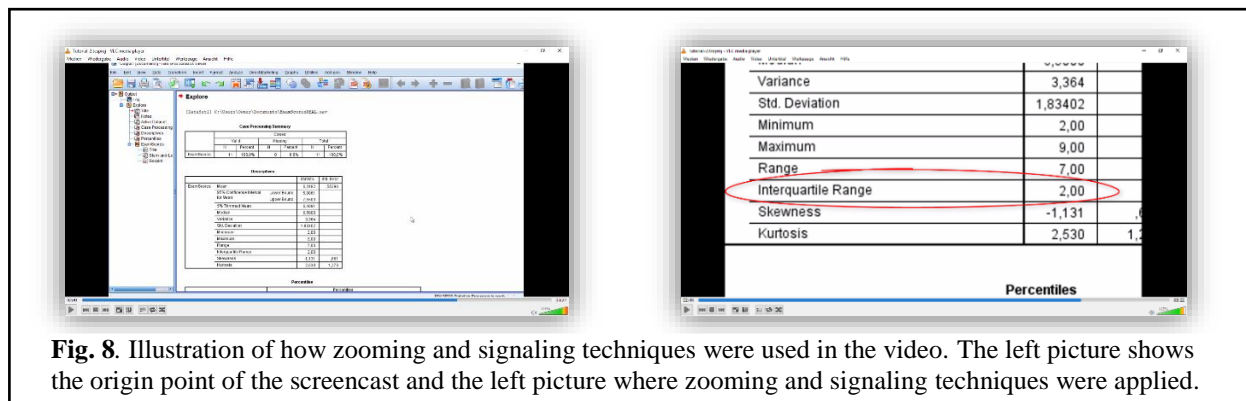
- Van Gog, T., Kester, L., & Paas, F. (2011). Effects of worked examples, example-problem, and problem-example pairs on novices' learning. *Contemporary Educational Psychology*, 36, 212-218. doi:10.1016/j.cedpsych.2010.10.004
- Van Merriënboer, J. J. G., & Sweller, J. (2005). Cognitive load theory and complex learning: Recent developments and future directions. *Educational Psychology Review*, 17(2), 147-177. doi:10.1007/s10648-005-3951-0
- Vezin, J. F., Berge, O., & Mavrelis, P. (1973). Role du résumé et de la répétition en fonction de leur place par rapport en texte. *Bulletin de Psychologie*, 27, 163-167
- Veronikas, S. W., & Maushak, N. (2005). Effectiveness of audio on screen captures in software application instruction. *Journal of Educational Multimedia and Hypermedia*, 14(2), 199.
- Wieling, M., & Hofman, W. (2010). The impact of online video lecture recordings and automated feedback on student performance. *Computers & Education*, 54(4), 992- 998. doi:10.1016/j.compedu.2009.10.002
- Wistia. (2012). Does length matter? Retrieved from [http:// wistia.com/blog/does-length-matter-it-does-for-video-2k12edition](http://wistia.com/blog/does-length-matter-it-does-for-video-2k12edition)
- Wouters, P., Paas, F., & van Merriënboer, J. J. G. (2010). Observational learning from animated models: Effects of studying-practicing alternation and illusion of control on transfer. *Instructional Science*, 38, 89-104. doi:10.1007/s1121-008-9079-0
- Yue, C. L., Bjork, E. L., & Bjork, R. A. (2013). Reducing verbal redundancy in multimedia learning: An undesired desirable difficulty?..*Journal of Educational Psychology*, 105(2), 266.

## Appendix

### *Appendix A: Implemented features and techniques used in the videos*

#### *Features enhancing attention*

To support attentional processes in a bottom up fashion, the IVs used **signaling techniques**. Zooming was used to draw the user's attention to the relevant part of the screen. This was for example the case if the narrator explained to the audience where to find the outcome value in the SPSS output. In addition to zooming, signals were drawn on the interface whereby dynamism occurred that attracted attention (Fig. 8). These signals were drawn in red color to enhance the attention drawing effect (Kosslyn, Kievit, Russell & Shephard, 2012).



**Fig. 8.** Illustration of how zooming and signaling techniques were used in the video. The left picture shows the origin point of the screencast and the left picture where zooming and signaling techniques were applied.

As signaling techniques support attentional bottom-up processes, the inclusion of a **preview** supports attentional processes in a top-down fashion. A preview is a short presentation which informs the viewer beforehand about the topic or goal of a subsequent demonstration. In this study, previews were used at the starting screen for each video. For example, if the video discussed the concepts quartiles and interquartile range, the names of the concepts were written on the starting screen as well as the category to which the concepts belonged (Fig. 9).

In addition, the narrator complemented the visual preview with an auditory preview (“Welcome back. This video will focus on measures of spread. We will discuss quartiles and the interquartile range”). The short previews at the opening of each video lasted between 5-9 seconds.

Besides this, at the end of the first three videos a short preview of the upcoming video was implemented. The names of the upcoming concepts were written on the screen (Fig. 10). This was complemented with a narrative (“Thanks for watching. The next video will focus on boxplot and outliers”). The closing for each video lasted between 5-8 seconds.





**Fig 9.** Illustration of a video's starting screen where the upcoming discussed concepts are written on the screen as well as the category to which they belong.



**Fig. 10.** Illustration of a video's closing screen where the upcoming concepts of the following video are mentioned

The **pace** in all videos depended on the speaking rate of the male native English speaker. To measure the precise speaking rate the words per minute (wpm) were measured. To compute the exact wpm metric, the total number of narrated words was divided by the total duration of seconds for each video. The result was then multiplied with 60. The wpm count for the four tutorial videos was 122, 130, 135, 129 wpm and 114 wpm for the review video. Given the fact that the average rate of human speech is 125-150 wpm (Fulford, 1992) the pace of the videos can be qualified as slow to moderate. Due to the cognitive processes the audience was engaged in while following the demonstration, the goal for the speaking rate was to be not too fast because this could have led to cognitive overload in the participants (Lang et al., 2007).

Another feature to enhance attentional processes is the inclusion of a toolbar which gives the viewer **user control**. This can be achieved by giving the user the options to stop, pause, replay or rewind the video. User control is favored because it gives the users the opportunity to selectively watch the video based on their learning needs (Schreiber, Fukuta & Gordon, 2010). Empirical research values user control for learning from dynamic visualizations (Merkt & Schwan, 2014; Höffler & Schwartz, 2011). In the present study, the videos were presented on a special website where participants had the options to stop, pause or rewind certain parts of the video giving them complete control over the videos.

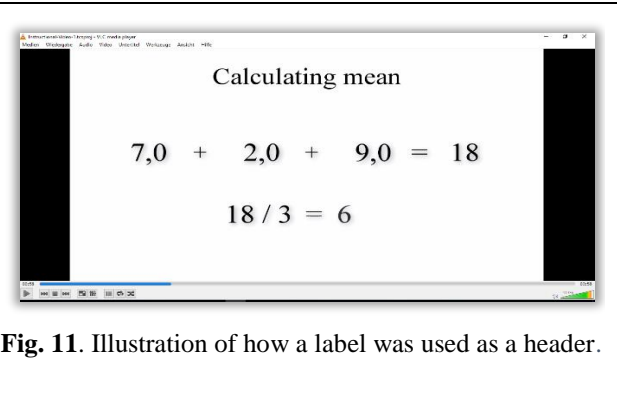
#### *Features enhancing retention*

The total descriptive statistics content was considered to be too diverse and complex to fit into one single video. Therefore, the content was **segmented** into four short video clips. The segmentation of the content led to relatively short video clips (around 3:30). Empirical research suggests that segmentation generally enhances learning from multimedia (Margulieux, Guzdial, & Catrambone, 2012; Mayer & Pilegard, 2014) and that shorter videos tend to prevent early drop out by viewers (Guo, Kim, & Rubin, 2014).

The four tutorials followed a **simple-to-complex sequence** (van Merriënboer & Sweller, 2005). For example, the calculation of the concepts of the first video takes fewer steps than the calculation of the concepts in the second video. Therefore, the content of the first video was considered to be simpler than the content of the second video.

Another feature to enhance retention is a **label**. A label summarizes the key point of a short video section, often consisting of one or two words. The inclusion of a label generates a “desirable difficulty” with the spoken narrative. This means that a small discrepancy between what the narrator says and what is written on the screen arises, stimulating the user to pay close attention

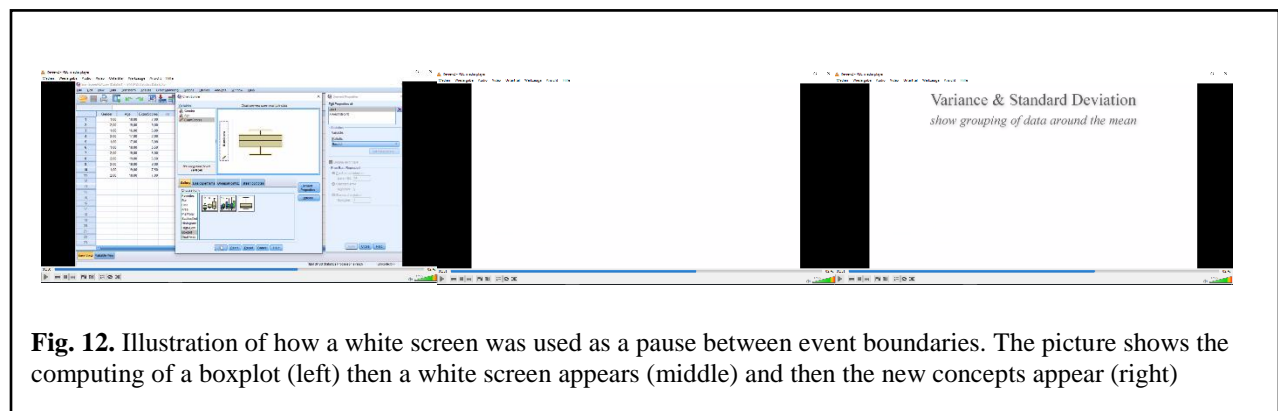
to both sources resulting in a deeper processing (Yue, Bjork & Bjork, 2013). In the present study labeling was used as a header for certain video sections. For example, if the narrator explained the calculation of a concept, a header was written on the top of the screen which served as a label (Fig. 11).



**Fig. 11.** Illustration of how a label was used as a header.

Retention processes can also be enhanced by the inclusion of **pauses**. Pauses are short breaks lasting about 2 seconds within a video. During a pause no visual or auditory

stimulus is presented giving the user the opportunity to process the before presented information. Furthermore a pause helps in organizing information. According to Spanjers, van Gog, Wouters and van Merriënboer (2012), pauses lead to a lower difficulty rating by students for dynamic visualizations and enhance learning. In the present research, pauses were used between event boundaries. If the narrator was finished with explaining a concept, the written text on the screen vanished until only a white screen was presented. Afterwards the label of the new concept appeared on the top of the screen (Fig. 12).



**Fig. 12.** Illustration of how a white screen was used as a pause between event boundaries. The picture shows the computing of a boxplot (left) then a white screen appears (middle) and then the new concepts appear (right)

### *Features enhancing motivation*

An instructional feature enhancing motivation in software training is **anchoring the tool in the task domain**. Domain anchoring was suggested by the minimalist approach to software training (Van der Meij & Carroll, 1998). In the current study, domain anchoring was attained by informing the participants that the video content will prepare them for an upcoming exam and that the demonstrated content will be useful for their further study. According to Merrill (2002), tasks need to be relevant, interesting and engaging for the user to stimulate learning.

Making use of a **conversational style** by the narrator has also a motivational effect. Conversational style means frequently using personal pronouns such as “I”, “you” and “we.” By using these words, a sense of social partnership between narrator and user is created. The related principle is known as the personalization principle. Recently, a meta-study suggested that a conversational style yields to more learning than a formal style, but that the positive effect disappeared if the instruction lasted longer than 35minutes (Ginns, Martin & Marsh, 2013). The meta-study further concluded that using a conversational style has only a small effect on motivation. However, an even more recent study showed a significant motivational effect with the conversational style (Reichelt, Kämmerer, Niegemann & Zander, 2014). In addition, action steps were given in the form of commands (Fig. 13). Research has shown that this is a recommended form of address (Farkas, 1999).

- ...

  1. Let's focus on how to determine the median.
  2. **Step 1: Put your scores in ascending order.**
  3. **Step 2: Find the middle value, this is your median.**
  4. The median of **our** three exam scores is seven.
  5. In this example, **we** determined the median for an even number of scores, which was simple, but what if **we** have an odd number of scores?
  6. Suppose there is a fourth student who took the exam. **We** have four scores now and therefore no value immediately visible.
  7. The procedure in this case is as follows:
  8. **Step 1: Put your scores in ascending order.**
  9. **Step 2: Put the two middle values and add them up.**
  10. **Step 3: Divide the result by two.**
  11. The median of **our** four exam scores is 6,5.

**Fig. 13.** Example of how conversational style and commands for action steps were used in the videos. (Bolt and numbering are only for reference purposes).

The **length** of a video is another factor which contributes to the motivation perceived by the viewer. Research demonstrated that long videos lead to an early dropout by viewers (Guo, Kim & Rubin, 2014; Wistia, 2012). To prevent such an effect, it is best to aim for a video length of 3-5min maximum. The videos in the present research fall into that range with a length of 3:58, 3:22, 3:02 and 3:00. The review video lies slightly below this range with a length of 2:48.

Research on the effect of **music** in multimedia design is sparse. A meta-study published by Kämpfe, Sedlmeier and Renkewitz (2010) which reviewed the few studies that compared conditions with and without music found that music had a small but negative effect on cognition but a positive effect on emotion.

In the current study music was played only at the start and at the end of each video. At the beginning, music was accompanied with a short introduction of the videos content (e.g., “Welcome back – this video will focus on measures of spread. We will discuss quartiles and the interquartile range”). At the closing, music accompanied a rounding off and for the first three

videos, a short preview for the next video (e.g., “And that’s it – thanks for watching. The next video will focus on boxplot and outliers”). The same uplifting instrumental music was used for the start and for the closing of the videos. The introduction and closing took 5-10s each for all videos.

## *Appendix B: SPSS practice*

The study associations Dimensie and Communiqué are planning the next party at the Vestingbar. To buy enough beer for the party, the associations want to find out about the beer consumption of the students. Therefore, they ask 50 students how many glasses of beer (0,33l) they drank on their last party.

The data of the 50 students are in the file “glassesofbeer.sav” Now open this file with SPSS (It is attached to your E-mail). Then come back to this document to continue.

**Before you start**, please state your username? (your login account from the E-mail, not your real name)

**Question 1:** What is the average number of glasses of beer the students drank at their last party?

*1.1 Move to SPSS, do your computing and copy and paste the SPSS output containing the answer below:*

*1.2 State the value in question in the box below:*

Value:

**Question 2:** What is the number of glasses of beer for the student(s) on the 50<sup>th</sup> percentile?  
*2.1 Move to SPSS, do your computing and copy and paste the SPSS output containing the answer below:*

*2.2 State the value in question in the box below:*  
Value:

**Question 3:** What is the value of the 10<sup>th</sup> percentile?  
*3.1 Move to SPSS, do your computing and copy and paste the SPSS output containing the answer below:*

*3.2 State the value in question in the box below:*  
Value:

**Question 4:** What is the range of values for the interval between the 1<sup>st</sup> and 3<sup>rd</sup> quartile?  
*4.1 Move to SPSS, do your computing and copy and paste the SPSS output containing the answer below:*

*4.2 State the value in question in the box below:*

Value:

**Question 5:** Make a boxplot from the glasses of beer.

*5.1 Move to SPSS, do your computing and copy and paste the SPSS output containing the answer below:*

**Question 6:** Find the two values which indicate how spread out around the mean the scores are.  
*6.1 Move to SPSS, do your computing and copy and paste the SPSS output containing the answer below:*

*6.2 State the values in question in the boxes below:*

Value:

Value:

- Please check if you have filled in your username.
- Now save this document and send it back to the researcher when you are done with the pre-training.
- Then move back to the website to continue.



### *Appendix C: SPSS performance and transfer test*

You want to investigate the Facebook activity of a group of 50 students. Therefore, you do research on the number of Facebook-Friends these students have.

The data of the 50 students are in the file “facebookfriends.sav” Now open this file with SPSS (It is attached to your E-mail). Then come back to this document to continue.

**Before you start**, please state your username (your login account from the E-mail, not your real name)

**Question 1:** Make a boxplot from the Facebook-Friends.

*Move to SPSS, do your computing and copy and paste the SPSS output containing the answer below:*

*SPSS Output:*

**Question 2:** Find the two values which indicate how spread out around the mean the scores are.  
*2.1 Move to SPSS, do your computing and copy and paste the SPSS output containing the answer below.*

|  |
|--|
|  |
|--|

*2.2 State the values in question in the boxes below:*

Value: 

|  |
|--|
|  |
|--|

Value: 

|  |
|--|
|  |
|--|

**Question 3:** What is the average number of Facebook-Friends for these students?

*3.1 Move to SPSS, do your computing and copy and paste the SPSS output containing the answer below:*

|  |
|--|
|  |
|--|

*3.2 State the value in question in the box below:*

Value: 

|  |
|--|
|  |
|--|

**Question 4:** What is the value of the 10<sup>th</sup> percentile?

*4.1 Move to SPSS, do your computing and copy and paste the SPSS output containing the answer below:*

*4.2 State the value in question in the box below:*

Value:

**Question 5:** What is the range of values for the interval between the 1<sup>st</sup> and 3<sup>rd</sup> quartile?

*5.1 Move to SPSS, do your computing and copy and paste the SPSS output containing the answer below:*

*5.2 State the value in question in the box below:*

Value:

**Question 6:** What is the number of Facebook-Friends for the student(s) on the 50<sup>th</sup> percentile?  
*6.1 Move to SPSS, do your computing and copy and paste the SPSS output containing the answer below:*

*6.2 State the value in question in the box below:*

Value:

**Question 7:** Mode is another measurement of central tendency. The mode is the most frequently occurring value in a data set. Determine the mode of your data set using SPSS.

*7.1 Move to SPSS, do your computing and copy and paste the SPSS output containing the answer below:*

*7.2 State the value in question in the box below:*

Value:

**Question 8:** Range is another measurement of dispersion. The range is the distance between the highest and lowest score within a data set. Determine the range of your data set using SPSS.

*8.1 Move to SPSS, do your computing and copy and paste the SPSS output containing the answer below:*

*8.2 State the value in question in the box below:*

Value:

**Question 9:** Another graphical representation of data is a histogram. Create a histogram of your data set using SPSS.

*9.1 Move to SPSS, do your computing and copy and paste the SPSS output containing the answer below:*

- Please check if you have filled in your username.
- Now save this document and send it back to the researcher when you are done with the pre-training.
- Then move back to the website to continue.

*Appendix D: Multiple-choice knowledge test*

1.

Quartiles

- ☐ are four equal segments of all scores
- ☒ **are three points that divide the scores into four equal parts**
- ☐ are needed to calculate the variance
- ☐ give you information about your highest and lowest scores

2.

Which of the following statements is true?

- ☐ True                      The mean is the same as the second quartile
- ☒ **True                      The third quartile is the same as the 75<sup>th</sup> percentile**
- ☐ True                      The first quartile is the same as the 10<sup>th</sup> percentile
- ☐ True                      The 50<sup>th</sup> percentile is the same as the mean

3.

In a boxplot

- ☐ The short line below the beige rectangle indicates Q1 (first quartile)
- ☒ **The beige rectangle itself indicates IQR (the interquartile range)**
- ☐ The short line above the beige rectangle indicates Q3 (third quartile)
- ☐ The black line in the middle indicates the mean

The body height of four family members are:

|          |        |
|----------|--------|
| father   | 180 cm |
| mother   | 160 cm |
| daughter | 170 cm |
| son      | 190 cm |

4.

The mean body height in this family is

- ☐ 160 cm
- ☐ 165 cm
- ☐ 170 cm
- ☒ **175 cm**

5.

In the same family the median body height is

- ☐ **175 cm**
- ☐ 170 cm
- ☒ 165 cm
- ☐ 180 cm

6.

The standard deviation

- ☒ **is expressed in the same unit as the original scores**
- ☐ gives less information than the variance
- ☐ takes fewer steps to calculate than the variance
- ☐ is preferably used in mathematical and statistical computing

7.

Which of the following statements is true?

- ☐ A score below the first quartile (Q1) is an outlier
- ☐ A score above the third quartile (Q3) is an outlier
- ☒ **The degree of dispersion (high or low) has an effect on the calculation of an outlier**
- ☐ An outlier is not shown in a boxplot

8.

When computing quartiles, knowing the median is important.

- ☒ **true**
- ☐ false
- ☐ it depends on the number of scores
- ☐ the mean is important

9.

Deviance is

- ☐ the same as the variance
- ☐ a score far from the median
- ☒ **the difference between each score and the mean**
- ☐ the difference between each score and the median

10.

When calculating the variance...

- ☐ ...you add all the deviances and divide by the number of scores minus 1
- ☐ ...you add all the deviances, square the result and divide by the number of scores minus 1
- ☐ ...you take the square root of the sum of all deviances
- ☒ **...you square all the deviances, add them and divide by the number of scores minus 1**

11.

The interquartile range

- ☐ is not shown in a boxplot
- ☐ is the only measure of dispersion
- ☒ **is the distance between the third and first quartile**
- ☐ gives you combined information about mean and median

### *Appendix E: Motivation questionnaire*

The following statements refer to your personal opinion about the tasks you performed. To what extent did you gain SPSS skills after completing the tasks?

Please rate how strongly you agree or disagree with each of the following statements by placing a check mark in the appropriate box. There are no right or wrong answers.

| <b>Statement</b>                                    | <b>Strongly disagree</b> | <b>Disagree</b> | <b>Slightly disagree</b> | <b>Neutral</b> | <b>Slightly agree</b> | <b>Agree</b> | <b>Strongly agree</b> |
|---|--------------------------|-----------------|--------------------------|----------------|-----------------------|--------------|-----------------------|
| I found the tasks important                         |                          |                 |                          |                |                       |              |                       |
| I found the tasks interesting                       |                          |                 |                          |                |                       |              |                       |
| I think that SPSS is handy for statistical analyses |                          |                 |                          |                |                       |              |                       |
| I think that SPSS is relevant for my study          |                          |                 |                          |                |                       |              |                       |
| The tasks were easy                                 |                          |                 |                          |                |                       |              |                       |
| SPSS is easy to use                                 |                          |                 |                          |                |                       |              |                       |
| I now know how to work with SPSS                    |                          |                 |                          |                |                       |              |                       |
| I can now compute a mean in SPSS                    |                          |                 |                          |                |                       |              |                       |
| I can now create a boxplot in SPSS                  |                          |                 |                          |                |                       |              |                       |



*Appendix F: E-mail for students*

Dear students,

In cooperation with the teacher of your course Research Methods & Descriptive Statistics, a brief pre-training (via Internet, so any time any place - no travel needed) was created for the course content (meeting 10) taught on March 9th. The pre-training consists of a few instructional (SPSS) videos as well as some statistical assignments.

You are invited to follow this pre-training. Participation is voluntary. It will prepare you for the lecture in March and you may also be better prepared for the SPSS exam in April. In addition, the UT offers you a cash payment of 10€ if you complete the pre-training.

Participation takes approximately 60 minutes. The pre-training opportunity is offered until March 8th.

If you are willing to participate, please reply to this E-mail and you will receive further instructions.

Kind regards,

*Paul Dunkel*

*Appendix G: Bivariate correlations among the dependent variables*

|                    | 1     | 2             | 3             | 4             | 5            | 6             | 7     | 8             | 9             | 10    | 11            | 12            | 13           | 14   | 15            | 16 |
|--------------------|-------|---------------|---------------|---------------|--------------|---------------|-------|---------------|---------------|-------|---------------|---------------|--------------|------|---------------|----|
| Variables          |       |               |               |               |              |               |       |               |               |       |               |               |              |      |               |    |
| 1. Video #1 unique | 1     |               |               |               |              |               |       |               |               |       |               |               |              |      |               |    |
| 2. Video #1 total  | .207  | 1             |               |               |              |               |       |               |               |       |               |               |              |      |               |    |
| 3. Video #2 unique | .016  | .202          | 1             |               |              |               |       |               |               |       |               |               |              |      |               |    |
| 4. Video #2 total  | -.169 | <b>.426**</b> | .078          | 1             |              |               |       |               |               |       |               |               |              |      |               |    |
| 5. Video #3 unique | .063  | .014          | -.005         | .093          | 1            |               |       |               |               |       |               |               |              |      |               |    |
| 6. Video #3 total  | .003  | <b>.451**</b> | .012          | <b>.580**</b> | <b>.276*</b> | 1             |       |               |               |       |               |               |              |      |               |    |
| 7. Video #4 unique | .028  | .268          | <b>.611**</b> | .155          | .090         | .104          | 1     |               |               |       |               |               |              |      |               |    |
| 8. Video #4 total  | -.035 | <b>.386**</b> | .054          | <b>.681**</b> | .220         | <b>.674**</b> | .236  | 1             |               |       |               |               |              |      |               |    |
| 9. Review unique   | .056  | .133          | .313          | <b>.418*</b>  | <b>.373*</b> | .218          | .357  | <b>.348*</b>  | 1             |       |               |               |              |      |               |    |
| 10. Review total   | .032  | .250          | .184          | <b>.439*</b>  | .311         | .271          | .307  | <b>.340*</b>  | <b>.897**</b> | 1     |               |               |              |      |               |    |
| 11. Practice       | -.116 | .149          | -.007         | .226          | .143         | .283          | -.229 | .306          | .076          | -.077 | 1             |               |              |      |               |    |
| 12. Knowledge      | .208  | -.256         | -.056         | -.078         | .157         | .173          | -.016 | .134          | -.030         | -.147 | .021          | 1             |              |      |               |    |
| 13. Performance    | -.066 | .083          | .067          | .126          | .102         | <b>.255*</b>  | -.018 | <b>.315**</b> | .261          | .231  | <b>.592**</b> | .180          | 1            |      |               |    |
| 14. Transfer       | -.074 | .076          | .114          | .081          | -.144        | -.019         | -.066 | .111          | -.185         | -.302 | -.235         | .149          | .172         | 1    |               |    |
| 15. Task-relevance | -.120 | .199          | .086          | <b>.303*</b>  | -.149        | .158          | .151  | .108          | .206          | .235  | -.079         | .066          | .051         | .078 | 1             |    |
| 16. Self-efficacy  | -.031 | -.147         | .108          | -.007         | .148         | .162          | -.130 | .037          | -.038         | -.181 | -.124         | <b>.366**</b> | <b>.251*</b> | .195 | <b>.326**</b> | 1  |

Note: \*p<.05 \*\*p<.01 (2-tailed)

