

The Effects of Practice Schedules on Self-Efficacy, Flow Experience, and Task Performance during a Video-Based Training: An Experimental Study

Master Thesis

Author: Jasmijn Maseland

Email: j.maseland@student.utwente.nl

1st Supervisor: Dr. Hans van der Meij

2nd Supervisor: Dr. Tessa Eysink

University of Twente

July, 2018

Abstract

Instructional videos are often used to model task performance. To learn from these videos, they need to offer more than an example of task completion. This study departed from a 'Demonstration-Based Training' (DBT) approach to optimize the video design for learning. Of these features, only practice schedules were manipulated. This study focuses on the effectiveness of two schedules: (1) blocked practice, (2) mixed practice. Blocked practice is based on a sequence in which task instruction is directly followed by practice on that task. In mixed practice several task instructions precede the practice. The research questions are: *(1) What is the effect of the blocked and mixed practice schedule on the self-efficacy of students during an initial and final self-efficacy questionnaire? (2) What is the effect of the blocked and mixed practice condition on flow experience of students during the training, and after the training (immediate, delayed, and transfer)? And, (3) What is the effect of the blocked and mixed practice condition on task performance during the training, and after the training (immediate, delayed, and transfer)?* 56 third and fourth grade primary school children (mean age 9.73 years) followed a video-based training on 'Word'. Participants were randomly assigned to one of the two conditions. The data showed a significant positive change in self-efficacy in both conditions. Also, blocked practice raised these scores more than mixed practice. The data for flow experience showed significantly higher scores on several measurement points for the blocked practice condition. Lastly, the study showed that the blocked practice condition outperformed the mixed practice conditions on a transfer test. For practice tasks and immediate and delayed retention no difference was found. The findings call for a replication study, but the provisional recommendation is to enable blocked rather than mixed practice in video-based software training.

Keywords: practice, contextual interference effect, blocked practice, mixed practice, instructional video, software training

Table of content

1. Project description	4
1.1 Problem statement.....	4
1.2 Theoretical framework.....	5
1.2.1 Practice schedules	5
1.3 Research design and question.....	6
2. Method	7
2.1 Respondents.....	7
2.2 Materials.....	7
2.2.1 Instructional materials.....	7
2.2.2 Measurement instruments.....	10
2.3 Procedure	13
2.4 Data analysis.....	14
3. Results	14
3.1 The effect of practice schedules on self-efficacy	14
3.2 The effect of practice schedules on flow experience per test	15
3.3 The effect of practice schedules on task performance	16
4. Discussion and conclusion	16
5. References	19
6. Appendices	23
Appendix A	23
Appendix B	24
Appendix C.....	25
Appendix D	27

1. Project description

1.1 Problem statement

During the years instructional video has become a popular medium to develop procedural knowledge (Mayer, 2008; Giannakos, 2013; Lloyd & Robertson, 2012). Procedural knowledge can be defined as the knowledge that is needed to execute different steps in a sequence to complete a task successfully (Rittle-Johnson, Siegler, & Alibali, 2001). For example, an instructional video from a software company often aims to develop procedural knowledge (Van der Meij & Van der Meij, 2016).

Also in schools, the importance of technology increases. Skillen (2008) states that learning can be enriched if pedagogical processes are combined with the use of technology. Additionally, Ertelt, Renkl, and Spada (2006) claim that instructional videos are of great importance in different learning contexts, especially when visualizations are indispensable. Besides, Ertelt et al. (2006) proved that instructional video can be used as an effective tool to teach procedural knowledge. In short, instructional video is already used to teach different content and subjects to children (Fokides, 2017).

Most instructional videos for procedural knowledge development present a model of task performance. This model shows how to perform each step in a procedure. Viewers can learn how to achieve the demonstrated task by mimicking the model. That is, they can learn from observation. Rosen, Salas, Pavlas, Jensen, Fu, and Lampton (2010) claim that observational learning benefits from an embellished design; showing merely a demonstration of action steps often leads to poor learning outcomes (see also Kim, Kim, Khera, & Getman, 2014; Ertelt et al., 2006; Hobbs, 1998; Marx & Frost, 1998). Rosen et al. (2010) suggests that to optimize learning, the demonstration design needs to address the fundamental processes involved. According to Bandura (1986), these processes are: (1) attention, (2) retention, (3) production, and (4) motivation. The first process, attention, requires that the learner selectively pays attention to important aspects of the learning material. The second process, retention, requires that the learner understands and remember the actions so that the task can later be achieved without support. The third process, production, requires that the learner literally reproduces the demonstrated and observed task. Lastly, the fourth process, motivation, underlies the other three processes. Motivation yields the stimulus for initial task engagement and persistence. The Demonstration-Based Training (DBT) approach has been proposed as the framework for coupling design features to each of the four processes (e.g., Brar & Van der Meij, 2017; Grossmann, Salas, Pavlas, & Rosen, 2013; Rosen et al., 2010). For instance, it is suggested that attention can be drawn to important screen objects by including highlights or cues. And retention can be supported by segmentation, meaning that instructional content is split into (more) manageable but still meaningful segments. The method section describes the main design features that were applied in the creation of the instructional video for the present study. The main feature investigated in this study is practice (as support for production).

Several studies showed that the incorporation of a moment of practice in combination with observational learning increases learning (e.g. Wouters, Paas, & Van Merriënboer, 2010; Ertelt, 2007; Van Gog, Kester, & Paas, 2011; Leppink, Paas, Van Gog, Van der Vleuten, & Van Merriënboer, 2014). In the literature two practice schedules are commonly distinguished: (1) blocked practice, and (2) mixed practice (Helsdingen, Van Gog, & Van Merriënboer, 2011). The first schedule, blocked practice, has a sequence in which an instruction of a task is directly followed by a moment of practice of that task. With mixed practice a different video-practice arrangement is adopted. First several task videos are

shown one after the other and only thereafter there is task practice (Helsdingen et al., 2011). Grossmann et al. (2013) mention that it remains unclear which schedule of practice is more effective for learning procedural knowledge. Therefore, this study investigates this issue.

The research goal of the study is to investigate the effectiveness of the two practice schedules for video-based software training involving a new domain. In this study that domain concerns Microsoft Office 'Word'. The study investigates more features than only the task performance. This research also investigates the flow experience and self-efficacy of the learners. In short, this research contributes to the research that already has been done on the effects of instructional video in combination with practice schedules with several new characteristics.

1.2 Theoretical framework

As mentioned earlier, instructional videos are rapidly gaining in popularity (Giannakos, 2013; Mayer, 2008; Lloyd & Robertson, 2012). Van der Meij and Van der Meij (2013) state that procedural knowledge can be acquired by observing a demonstration of performance in combination with instructional support. Grossmann et al. (2013) and Rosen et al. (2010) state that the framework Demonstration-Based Training (DBT) tries to accomplish this goal. Besides, practice during instructional video provides engagement of the learners in the learning process (Leppink et al., 2014). Practice is also one of the requirements of DBT to satisfy the fundamental process of 'production' (Bandura, 1986). Therefore, different schedules of practice are described.

1.2.1 Practice schedules

Helsdingen et al. (2011) mentioned two schedules of practice: (1) blocked practice, and (2) mixed practice. Blocked practice is based on a sequence in which task instruction is directly followed by a moment of practice on that task. This should lead to a better performance during training. In mixed practice several task instructions precede the task practices, which makes the assignment more challenging. The net effect is a stronger mental model that yields better retention and transfer (Helsdingen et al., 2011).

As mentioned above, blocked practice has been found to lead to better performance during training and a relative degradation of post-training performances and transfer. The opposite is shown for mixed practice whereby performance during training is degraded, but post-training performance and transfer of the learner is better. These results are also called the 'Contextual Interference Effect' (CI-Effect) (Helsdingen et al., 2011; Lee & Simon, 2004; Shea & Morgan, 1979; Wulf & Shea, 2002). The CI-effect has been observed in several domains, for example learning problem solving skills (Paas & Van Merriënboer, 1994), motor tasks (Cross, Schmitt, & Grafton, 2007; Guadagnoli & Lee, 2004; Simon, 2007; Welsher & Grierson, 2017; Neville & Trempe, 2017), cognitive operational tasks (Jamieson & Rogers, 2000), perceptual cognitive tasks (Broadbent, Causer, Williams, & Ford, 2017), foreign vocabulary learning (Schneider, Healy, & Bourne, 2002), sports practice (Farrow & Buszard, 2017), and troubleshooting tasks (De Croock, Van Merriënboer, & Paas, 1998).

According to Helsdingen et al. (2011) research there are two prominent explanations for the CI-effect. Firstly, Shea and Morgan (1979) stated the 'elaborative-processing hypothesis' as an explanation. This hypothesis states that mixed practice challenges the learner to make comparisons between the tasks (Helsdingen et al., 2011; Lin, Fisher, Winstein, Wu, & Gordon, 2008). Because the learner needs to identify each task variation

that is presented, the performances during the training are disadvantaged and the post-training performances and transfer are advanced. In blocked practice one task is trained. Besides, the learner only has to reproduce the demonstration of one task at a time. The consequence is that in blocked practice, the performance is better during the training, because the learner needs to have remembered only the steps to one particular task. For mixed practice counts that the learner is stimulated to create more elaborate and distinctive memorial representations due to the identification of each task variation, with the consequence that retention tests have superior performances. Also transfer tests are superior in mixed practice, because the learner already is aware of choosing a solution to the task (Helsdingen et al., 2011).

Secondly, Lee and Magill (1983, 1985) proposed the ‘forgetting-and-reconstruction hypothesis’ as an explanation. This hypothesis states that mixed practice challenges learners to switch strategy between practice tasks, creating a new strategy to complete a task successfully. That is, the learner constantly needs to adapt strategies to the to-be-performed tasks. This makes the performance during the training of lower quality but increases the quality in post-training and transfer tests. The opposite is seen in blocked practice, because the learner simply needs to recall the just-instructed task strategy. The learner is not challenged with competing strategies which results in more successful task performances during training, but which reduce learning of reconstruction strategies for different tasks.

1.3 Research design and question

Based on the problem statement and theoretical framework above this study aims to find an answer on the effectiveness of different schedules of practice during DBT through an instructional video about the software ‘Word’. This study used an experimental design with practice schedule as independent variable and performance task, flow experience, and self-efficacy as dependent variables. Two schedules of practice will be investigated: (1) blocked practice, and (2) mixed practice.

The experimental study consisted of a condition that represented the blocked practice schedule, and a condition that represented the mixed practice schedule. In order to investigate the effectiveness of the schedules of practice, the following research questions will be investigated:

Research question 1: What is the effect of the blocked and mixed practice schedule on the self-efficacy of students during an initial and final self-efficacy questionnaire? For motivation this study investigates self-efficacy, which can be defined as a person’s belief in the capacity to organize and execute the actions necessary to manage particular task outcomes (Bandura, 1997). One reason for choosing this motivational construct is that, to our knowledge, no earlier research on the CI-effect has investigated this variable. Another argument is that self-efficacy is a predictor of future actions such as persistence and greater effort expenditure in comparable settings (Bandura, 2012; Bandura & Locke, 2003). According to research on the CI-effect (Helsdingen et al, 2011; Lee & Simon, 2004; Shea & Morgan, 1979; Wulf & Shea, 2002) the tested hypothesis is that the blocked practice condition raises self-efficacy more than the mixed practice condition because final self-efficacy is measured immediately after training and therefore should not suffer from a relative degradation of post-training performances and transfer that blocked practice condition may experience.

Research question 2: What is the effect of the blocked and mixed practice condition on flow experience of students during the training, and after the training (immediate,

delayed, and transfer)? For cognitive load this study investigates flow experience, which can be defined as a state wherein an individual functions at his or her fullest capacity in combination with deep engagement with the learning task. Flow experience indicates whether or not students experience optimal concentration (Shernoff, Csikszentmihalyi, Schneider, & Shernoff, 2014; Yang & Tao, 2015). Although research on the CI-effect mentions differences in cognitive load during and after training for the two practice schedules, these differences appear not to have been measured. The tested hypothesis is that flow is highest during training in the blocked practice condition, but is lowest after training.

Research question 3: What is the effect of the blocked and mixed practice condition on task performance during the training, and after the training (immediate, delayed, and transfer)? To gauge learning the successful task performances are measured. In line with research on the CI-effect (e.g., Helsdingen et al., 2011; Shea & Morgan, 1979; Lee & Simon, 2004; Wulf & Shea, 2002; Paas & Van Merriënboer, 1994; Cross et al., 2007; Guadagnoli & Lee, 2004; Simon, 2007; Welsher & Grierson, 2017; Neville & Trempe, 2017; Jamieson & Rogers, 2000; Broadbent et al., 2017; Schneider et al., 2002; Farrow & Buszard, 2017; De Croock et al., 1998) there are measurements of trained tasks during (practice) and after training (immediate and delayed test). In addition, a transfer test is administered. The tested hypothesis is that the blocked practice condition leads to better performance during training, and lesser post training performances (immediate, and delayed) and transfer than the mixed condition (Helsdingen et al., 2011).

2. Method

2.1 Respondents

There were 56 students participating in the study. The students came from the third and fourth grade of a Dutch primary school. The respondents had an age between eight and twelve years old, with an average age of 10 years ($M = 9.73$; $SD = .73$). The study included 27 male students (48.2%) and 29 female students (51.8%). Most of the students were fluent Dutch speakers, other students mastered the basic skills of the Dutch language, and therefore were able to participate in the study. All students were novices or beginners in Microsoft Office Word 2010.

The students in each grade were randomly assigned to conditions. The study took place during the normal school time, between 08.30 and 14.00. All parents gave permission by e-mail before the study started. All used materials for this study were in Dutch, because the study was conducted in the Netherlands with Dutch students. Approval for the study was obtained from the Ethical Committee of the University of Twente.

2.2 Materials

2.2.1 Instructional materials

Instructional videos. To teach the respondents procedural knowledge about the software program 'Microsoft Office Word 2010', eight short instructional videos were constructed. These videos were organized in 'chapters' with paragraphs that gave access to the videos. The three chapters are: (1) starting and saving a Word file, (2) changing a text, and (3) the use of pictures in a Word file. Chapter 1 included instructional videos on: 1.1 starting an empty Word document (01:12), 1.2 opening an existing Word file (01:45), and 1.3 saving a Word file (02:32). Chapter 2 also presented three instructional videos: 2.1 deleting a

part of a text (03:25), 2.2 replacing a part of a text (03:17), and 2.3 copy and paste a part of a text (03:36). Chapter 3 contained two instructional videos on: 3.1 adding a picture in a Word file (03:31), and 3.2 changing the size of a picture (02:59). The exact distribution of the chapters and paragraphs of the instructional videos can be found in appendix A.

The instructional videos are designed according to the guidelines of the DBT-model (Brar & Van der Meij, 2017; Van der Meij & Van der Meij, 2013). Appendix B summarizes the framework. According to Brar and Van der Meij (2017) at least one design guideline per process needs to be implemented. Below we describe the guidelines that were adopted in creating the videos in this study.

To comply with the attention process, three guidelines are met: cueing, pace, and user control. Firstly, the instructional videos include **cueing**. According to Lemarié, Lorch, Eyrolle, and Virbel (2008) and Mayer (2008) cues point to the most important information in a video without adding any content. Examples of cues are: color coding, arrows, and circled or squared overlays (Brar & Van der Meij, 2017). There is statistical significant evidence that cues lead the attention of the user to the right location (Boucheix & Lowe, 2010) and raises learning (Richter, Scheiter, & Eitel, 2015). In the videos the following cueing techniques were used: arrows, circled overlays, and squared overlays. Secondly, the videos have been given a **moderate pace** (Brar & Van der Meij, 2017). According to Van der Meij and Van der Meij (2013) the use of a moderate space is recommended. Two reasons are given: (1) a pace that is too fast yields a cognitive overload of the working memory, and (2) a pace that is too slow ensures a boring video which consequently decreases the attention of the learner (e.g., Mayer, 2008; Boucheix & Guignard, 2005; Lang, Park, Sanders-Jackson, Wilson, & Wang, 2007; Boucheix, Lowe, & Bugaiska, 2015). Lastly, a toolbar is used to enable **user control**. The toolbar included options such as stop, pause, wind, and rewind. According to Witteman and Segers (2010) such options positively effect learning. Brar and Van der Meij (2017) also state that the learner can make the video fitting to their own cognitive capacity and learning needs.

To satisfy the retention process two guidelines are met: segmentation, and simple-to-complex task sequence. Firstly, **Segmentation** means dividing a longer video into smaller parts and adding a clear begin and end to each video (Brar & Van der Meij, 2017; Mayer, 2008). This guideline is met by creating short videos that are organized in chapters divided into paragraphs. Several studies have shown that segmentation enhances learning from instructional video (Margulieux, Guzdial, & Catrambone, 2012; Mayer, 2008). Secondly, the instructional videos are ordered with a **simple-to-complex sequence**. This means that the presentation of chapters and paragraphs is from simple to complex, which also should prevent or reduce the risk of early dropouts (Van Merriënboer & Sweller, 2005). For example, starting a Word document is presented at the beginning of the training, and adding a picture to a Word file appeared later in the training.

To meet the production process two guidelines are met: practice, and practice files. Firstly, the main goal of the study is to investigate two forms of practice schedules. Thereby, the guideline **practice** is automatically met. Practice is part of the complete training arrangement, and thereby is not a feature of the instructional video itself. According to Leppink et al. (2014) and Van Gog et al. (2011) practice is a user action that positively influences learning. Secondly, **practice files** are used to let the learner practice the demonstrated tasks. The focus in the practice files is on the demonstrated problem in the instructional video, all other information that might distract is kept to a minimum (Brar & Van der Meij, 2017; Van der Meij & Carroll, 1998).

To comply with the motivation process two guidelines are met: conversational style, and length. Firstly, a **conversational style** is used. According to Brar and Van der Meij (2017) an example of a conversational style is the use of personal and informal pronoun (e.g., I, you). The instructional videos included a conversational style as describe above. For example, 'You want to work with Word. Firstly, you need to open a new word file'. This sentence shows that an personal pronoun is used. Secondly, research shows that the maximum ideal **length** of an instructional video should be between 3 and 5 minutes (Wistia, 2012). Research by Guo, Kim, and Rubin (2014) shows that shorter videos lower dropout rates. All instructional videos are around 3 minutes, and therefore the guideline of video length is met.

Website. The instructional videos were available via a website. Each student received a participant number which was used as an username and password to log in during the training. Figure 1 shows the format of the website.

On the left side of the website there was a menu with a table of contents. The students could click on the instructional video they needed to watch. After clicking on an instructional video, it opened up on the right side of the website. While watching the instructional video, students could stop, pause, rewind, and forward the instructional video with the use of a toolbar at the bottom of the video. The toolbar popped up when the cursor was placed on the instructional video. Furthermore, the students could change the volume and click on the full-screen button.

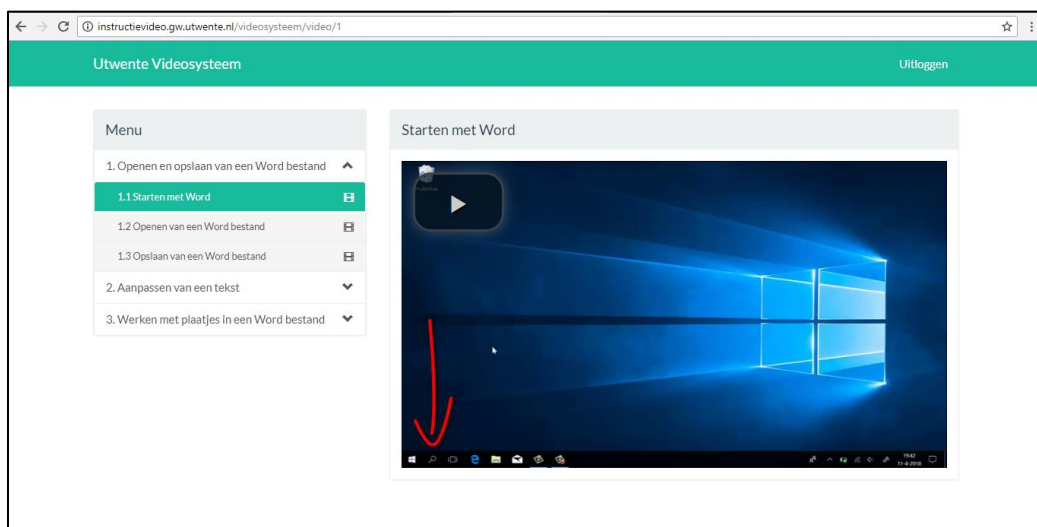


Figure 1. Screenshot of the website for the instructional video.

Student booklets. To guide the students' behavior during training they were given a booklet on paper telling them what to do (e.g., view or practice). The booklets for each condition differed from each other, of course. In the blocked practice booklet, an instruction to view a task video was directly followed by an instruction to engage in task practice (i.e., V1.1-P1.1-V1.2-P1.2-V1.3-P1.3). In the mixed practice student booklet, all instructional videos of a chapter directly followed each other, where after instructions for the practice tasks followed (i.e., V.1.1-V1.2-V1.3-P1.1-P1.2-P1.3). Practice tasks were explained in the student booklet, but executed on the tablet in the Word files. Both student booklets included the same flow-questions after each chapter.

Word files. To let the students practice during the training, Word files were created. The Word files represent the same problem as was demonstrated in the instructional video.

For every task of the training, immediate test, delayed test, and transfer test a different Word file is used. Word files differed in surface features, but had the same underlying structure. For example, during the training students were asked to change the title of the Word file from 'Tristan and Isolde' into 'Tristan and Isolde are in love'. For the immediate test the same task was asked, only a different name was used (i.e., from 'the frog' into 'The frog who wanted to be a bullock'). The Word files were available in the folder 'documents'.

Experimenter script. To guide the experiment an experimenter script was designed. The experimenter script showed exactly what should be said by the experimenter. In this way, the students of each condition receive exact the same instructions. Appendix C shows the experimenter script.

2.2.2 Measurement instruments

Tasks performance tests. Task performance success was measured during and after training: (1) practice tasks, (2) immediate test, (3) delayed test, and (4) transfer test. The first three tests are parallel, they assessed the students' task performance on the same tasks (same underlying problem, different surface features). These tests (practice tasks, immediate test, and delayed test) included each one item for each of the eight tasks whose completion was modeled in the instructional videos. The tests did not include an assessment of the first two trained tasks (i.e., start an empty Word document, and open an existing Word file) because the execution of these two tasks could not be registered during the experiment. Nevertheless, the students needed to perform the first two tasks to keep the training and test content the same. The transfer test includes items that are not directly trained but are comparable to the learning objectives of the instructional videos about the software Word. It included three items corresponding with the three main chapters of the experiment. The first item included a task wherein the participant should start the online template 'Neat and Pragmatic Resume' instead of starting up an empty Word document (see Figure 2). The second item included a task wherein the student should copy and paste a text from one to another Word file instead of copy and paste a text within a Word file. Lastly, the third item included a task wherein the student should make a picture smaller, instead of bigger.



Figure 2. Screenshot of the transfer test item 1.

A score of 0 points was awarded for a task that was not completed or not completed correctly. Participants received a score of 1 point when a (sub)task was performed correctly. Within each item, the number of possible points that could be earned depended on the number of subtasks involved. For example, the task belonging to the third paragraph of the first chapter (item 1.3) of the training does have a total points of three. Within task 1.3 the participant could receive 1 point for the following actions: the title is completed with the sentence 'are in love', the Word file is saved, and the word file is saved under the name 'Tristan and Isolde are in love' (see Figure 3). Points are awarded following the directions of a codebook. For the practice tasks, immediate test, and delayed test the maximum score was eight points each. For the transfer test, a maximum of seven points could be received. Scores were converted to percentages of possible points. Reliability analyses showed that the Cronbach's alpha was satisfactory to good for the four tests: practice tasks ($\alpha = .80$), immediate test ($\alpha = .79$), delayed test ($\alpha = .59$), and transfer test ($\alpha = .71$).

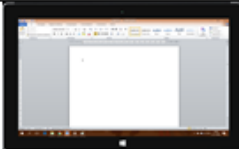
1.3 Opslaan van een Word bestand	
	Bekijk video 1.3 'Opslaan van een Word bestand'. Bekijk de video net zolang tot je weet hoe dit in Word gaat. Je mag pauzeren en terugspoelen.
	Opdracht 1.3 Opslaan van een Word bestand Je mag nu geen video bekijken. 1. Open het bestand 'Tristan en Isolde.docx' 2. Maak de titel af: Tristan en Isolde zijn verliefd. 3. Sla het bestand op met bestandsnaam: Tristan en Isolde zijn verliefd. 4. Sluit het bestand af.

Figure 3. Screenshot of the training item 1.3.

Self-efficacy questionnaire. Self-efficacy, a personal belief that someone is capable to complete a specified task successfully (Bandura, 1997), is measured with a paper-and-pencil questionnaire based on the Initial Experience and Motivation Questionnaire (IEMQ) (Van der Meij & Van der Meij, 2014). The self-efficacy questionnaire was administered twice, as initial questionnaire before training and as final questionnaire directly after training. The respondents were repeatedly asked the following question for the eight to-be-trained tasks of the instructional videos: 'How well do you think that you can perform this task?' (see Figure 4). A seven-point Likert-scale is used. Answers can range from (1) Very bad to (7) Very well. The mean score for the self-efficacy questionnaires will be reported. Reliability analyses showed that Cronbach's Alpha was good for the two questionnaires: initial self-efficacy test ($\alpha = .88$), and final self-efficacy test ($\alpha = .88$).

3. Opslaan van een Word bestand



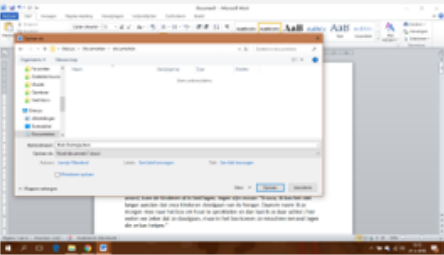
LEES TAAK

- **Kijk** naar afbeelding 1.
- Je **wilt** dit bestand opslaan in documenten.
- **Kijk** naar afbeelding 2.

1



2





VRAAG

	heel slecht	1	2	3	4	5	6	heel goed
1 Hoe goed kun je deze taak uitvoeren denk je?								

Figure 4. Screenshot the self-efficacy test item 3.

Flow experience questions. Lastly, a paper-and-pencil questionnaire based on the Experience Sampling Form (ESF) (Shernoff et al., 2014) is used to measure the flow experience of the respondents. Flow experience is defined as a state wherein an individual functions at his or her fullest capacity in combination with deep engagement with the learning task (Shernoff et al., 2014; Yang & Tao, 2015). The respondents need to fill in questions like: 'I knew exactly what to do by each step', 'I felt like I could make all the tasks easily', and 'Thinking was easy' (see Figure 5). A seven-point Likert-scale is used. Answers can range from (1) Completely not suits me to (7) completely suits me. Mean scores for flow experience per test were computed. The flow experience questionnaire was included in the student booklet during the training, immediate test, delayed test, and transfer test. Reliability analyses showed that Cronbach's alpha was excellent for each test: practice tasks ($\alpha = .85$), immediate test ($\alpha = .96$), delayed test ($\alpha = .95$), and transfer test ($\alpha = .92$). Besides, the Cronbach's alpha was also measured for each chapter of the training, immediate test, and delayed test. Reliability analyses for the training showed that Cronbach's alpha was satisfactory to good: chapter 1 ($\alpha = .77$), chapter 2 ($\alpha = .84$), and chapter 3 ($\alpha = .64$). Reliability analyses for the immediate test showed that Cronbach's alpha was good to excellent: chapter 1 ($\alpha = .90$), chapter 2 ($\alpha = .91$), and chapter 3 ($\alpha = .83$). Reliability analyses for the delayed test showed that Cronbach's alpha was excellent: chapter 1 ($\alpha = .92$), chapter 2 ($\alpha = .96$), and chapter 3 ($\alpha = .94$).

	Je hebt nu de opdrachten van hoofdstuk 1 gemaakt, je hebt: • een leeg Word bestand geopend. • een bestaand Word bestand geopend. • een Word bestand opgeslagen. Je krijgt nu 4 vragen over deze opdrachten. Omcirkel het antwoord dat het beste bij jou past.
---	---

	Klopt niet							Klopt wel
1. Bij iedere stap wist ik wat ik moest doen	1	2	3	4	5	6	7	
2. Ik had het gevoel dat ik de opdrachten goed kon maken	1	2	3	4	5	6	7	
3. Ik kon makkelijk denken tijdens de opdrachten	1	2	3	4	5	6	7	
4. De juiste gedachten kwamen vanzelf tijdens de opdrachten	1	2	3	4	5	6	7	

Figure 5. Screenshot of flow experience test during the training after chapter 1.

2.3 Procedure

The experiment was conducted in two sessions in a small classroom that seated a maximum of 18 students at a time (all from the same condition). The execution of the study took place during normal school days. The first session included the initial and final self-efficacy test, the training, and immediate test. The latter two also included the flow experience questionnaire. The second session consisted of the delayed test, and the transfer test, both also included the flow experience questionnaire. Each participant worked on a tablet with earplugs and a computer mouse. The Word files were already uploaded in the document folder by the experimenter before the start of the experiment. Also, the website was already opened and the student booklets, tests, pencil, and eraser were handed out before the start of the study.

After getting the respondents out of their classroom the experimenter gave a five-minute introduction telling them that the training was about Word consisting of eight instructional videos, corresponding practice tasks, and tasks that tested for what they remembered after the training. The introduction included also an explanation of the website and practice tasks following the experimenter script. Thereafter, the initial self-efficacy test was administered (max 5 minutes). Then, the participants started the training. With help of the student booklet, the students worked independently during training. In the blocked practice condition an instructional video (V) was immediate followed by corresponding task practice (P) (e.g., V1.1-P1.1-V1.2-P1.2-V1.3-P1.3). In mixed practice all instructional videos (V) of a chapter were viewed before the corresponding task practices (P) were performed (e.g., V1.1-V1.2-V1.3-P1.1-P1.2-P1.3). After each chapter, the students filled in the flow experience test. The instructor observed if the students were independently making the practice tasks as the instructional student booklet said. The training was ended after 45 minutes. Next, students took a short break of fifteen minutes. Thereafter, the students filled in the final self-efficacy test (max 5 minutes) and took the immediate test, plus flow experience questionnaire (max 20 minutes).

Exactly one week later, the second session took place. the students did a delayed test (max 20 minutes), and a transfer test (max 15 minutes), both including a flow experience questionnaire.

2.4 Data analysis

The data were analyzed with the program IBM SPSS Statistics version 23. First, a check on the random distribution of participant characteristics (i.e., age, gender) was done on the data. The chi-square test showed that there was no significant difference between the conditions with regard to gender, $\chi^2(1) = 1.79, p = .18$. However, unexpectedly, an ANOVA test on age showed that there was a significant difference between the conditions, $F(1, 55) = 4.08, p = .048$. The blocked practice condition had a mean age of 9.92 ($SD = .76$), and the mixed practice condition had a mean age of 9.54 ($SD = .67$). Therefore, the variable 'age' was treated as a covariate in all analyses.

Also, a check for outliers and distributions was done. The data is corrected for outliers which accounts for slight differences in the degrees of freedom. Tests on assumptions on normality of distribution, and homogeneity of variance (i.e., Levene test) revealed no violations for the three dependent variables (i.e., self-efficacy, task performance, and flow performance). Therefore, ANCOVAs could be used.

Flow measures were recorded after each chapter during and after training. Data analyses showed that this led to a considerable loss of data (incomplete datasets). In addition, the alpha level should be reduced due to repeated testing. Therefore it was decided to compute and report an overall score for flow. Mean scores for flow experience per test were computed. Appendix D shows the Tables and ANCOVA outcomes for individual chapters.

Comparisons involved two-sided tests with alpha levels of .05 for significance. One-sided tests were used for predicted effects (this is mentioned with the p-value). Cohen's (1988) d-statistic was used to indicate the effect size, classified as small for $d = .20$, medium for $d = .50$, and large for $d = .80$.

3. Results

3.1 The effect of practice schedules on self-efficacy

Table 1 shows the findings for the effect of practice schedules on the self-efficacy of the participants. For the initial self-efficacy test an ANCOVA test showed that there is no statistical significant difference in mean score between the two conditions, $F(1, 54) = 1.95, p = .169$. This means that the mean self-efficacy of the two conditions before the start of the training can be considered equal. In short, the two conditions do have the same self-efficacy at the start of the training.

For the final self-efficacy test, after the training, an ANCOVA showed that there is a statistical significant difference between conditions, $F(1, 53) = 8.86, p = .004$. The final self-efficacy is higher for the blocked practice than the mixed practice condition.¹

¹ An additional ANCOVA with initial self-efficacy and age as covariates is executed to double-check the findings. The ANCOVA showed the same outcome, namely that there is a statistical significant difference between conditions, $F(1, 52) = 5.47, p = .02$. This analysis also shows that the mean self-efficacy of the two conditions after the training is higher for the blocked practice condition than for the mixed practice condition.

Table 1. Mean self-efficacy score per condition and test.

Condition	Initial self-efficacy		Final self-efficacy	
	Mean	(SD)	Mean	(SD)
Blocked practice (n = 28, 27) ^a	4.54	(1.54)	6.34	(1.06)
Mixed practice (n = 28, 28) ^a	3.93	(1.20)	5.23	(1.36)
Total (n = 56, 55) ^a	4.24	(1.40)	5.78	(1.33)

^a The number of participants for initial self-efficacy test and final self-efficacy test.

3.2 The effect of practice schedules on flow experience per test

Table 2 shows the findings for the effect of practice schedules on the flow experience of the participants during training, and on the immediate and delayed test. An ANCOVA about the flow during the training showed that there is no statistical significant difference in mean score between the two conditions, $F(1, 53) = 3.20$, $p = .08$. This means that the mean flow of the two conditions during the training of the blocked practice condition and the mixed practice condition can be considered as equal.

An ANCOVA for the flow during the immediate test showed that there is no statistical significant difference in mean score between the two conditions, $F(1, 54) = 3.11$, $p = .084$. This means that the mean flow of the two conditions during the immediate test of the blocked practice condition and the mixed practice condition can be considered as equal.

An ANCOVA for the flow during the delayed test showed that there is a statistical significant difference in mean score between the two conditions, $F(1, 54) = 4.705$, $p = .036$. This means that the mean flow of the two conditions during the delayed test is higher for the blocked practice condition than for the mixed practice condition. In short, the blocked practice condition had a higher flow experience during the delayed test than the mixed practice condition.

Table 2. Mean flow score per condition for the training, immediate test, and delayed test.

Condition	Flow during training		Flow during immediate test		Flow during delayed test	
	Mean	(SD)	Mean	(SD)	Mean	(SD)
Blocked practice (n = 27, 28, 28) ^a	5.70	(1.32)	6.07	(1.45)	6.29	(1.07)
Mixed practice (n = 28, 28, 28) ^a	4.88	(1.48)	5.30	(1.34)	5.36	(1.43)
Total (n = 55, 56, 56) ^a	5.28	(1.45)	5.69	(1.44)	5.83	(1.34)

^a The number of participants for flow during the training, immediate test, and delayed test.

Table 3 shows the findings for the effect of practice schedules on the flow experience of the participants during the transfer test. An ANCOVA test showed that there is no statistical significant difference in mean score between the two conditions, $F(1, 53) = 2.96$, $p = .092$. This means that the mean flow of the two conditions during the transfer test of the blocked practice condition and the mixed practice condition can be considered as equal.

Table 3. Mean flow score per condition for the transfer test.

Condition	Flow during transfer test	
	Mean	(SD)
Blocked practice (n = 28) ^a	5.77	(1.32)
Mixed practice (n = 27) ^a	5.04	(1.37)
Total (n = 55) ^a	5.41	(1.38)

^a The number of participants for flow during the transfer test.

3.3 The effect of practice schedules on task performance

Table 4 shows the findings for the effect of practice schedules on the task performance of the participants during the training, immediate test, and delayed test. An ANCOVA for the task performance during the training showed that there is no statistical significant difference in mean score between the two conditions, $F(1, 54) = 1.46, p = .332$. This means that the mean task performance of the two conditions during the training can be considered as equal.

An ANCOVA for the task performance during the immediate test showed that there is no statistical significant difference in mean score between the two conditions, $F(1, 54) = .36, p = .552$. This means that the mean task performance of the two conditions during the immediate test can be considered as equal.

An ANCOVA for the task performance during the delayed test showed that there is no statistical significant difference in mean score between the two conditions, $F(1, 53) = .77, p = .384$. This means that the mean task performance of the two conditions during the delayed test can be considered as equal.

Table 4. Mean task performance score per condition for the training, immediate test, and delayed test.

Condition	Task performance training		Task performance immediate test		Task performance delayed test	
	Mean	(SD)	Mean	(SD)	Mean	(SD)
Blocked practice (n = 28, 28, 27) ^a	63.61%	(31.55)	66.25%	(33.70)	72.48%	(17.75)
Mixed practice (n = 28, 28, 28) ^a	53.39%	(27.83)	61.82%	(25.01)	65.89%	(21.31)
Total (n = 56, 56, 55) ^a	58.50%	(29.92)	64.04%	(29.49)	69.13%	(19.75)

^a The number of participants for task performance for training, immediate test, and delayed test.

Table 5 shows the findings for the effect of practice schedules on the task performance of the participants during the transfer test. An ANCOVA test showed that there is a statistical significant difference in mean score between the two conditions, $F(1, 54) = 4.90, p = .032$. This means that the mean task performance of the blocked practice condition is higher than the mean task performance score of the mixed practice condition. In short, there is statistical evidence that the blocked practice condition outperformed the mixed practice condition on the transfer test.

Table 5. Mean task performance score per condition for the transfer test.

Condition	Task performance transfer test	
	Mean	(SD)
Blocked practice (n = 28) ^a	59.11%	(26.83)
Mixed practice (n = 28) ^a	42.39%	(23.42)
Total (n = 56) ^a	50.75%	(26.34)

^a The number of participants for task performance for transfer test.

4. Discussion and conclusion

The aim of the present study was to examine the effects of two practice schedules, blocked practice and mixed practice, on students' self-efficacy, flow experience, and task performance. In order to investigate the effectiveness of the schedules of practice, the following research questions are used: (1) *What is the effect of the blocked and mixed practice schedule on the self-efficacy of students during an initial and final self-efficacy questionnaire?* (2) *What is the effect of the blocked and mixed practice condition on flow*

experience of students during the training, and after the training (immediate, delayed, and transfer)? And, (3) What is the effect of the blocked and mixed practice condition on task performance during the training, and after the training (immediate, delayed, and transfer)? The findings for these research questions are discussed below.

CI-research had not before measured the contribution of practice schedule on self-efficacy. The study showed that all participants raised their self-efficacy. In short, the training seemed to be effective in raising this motivational construct. According to Van der Meij (2017), Bandura and Locke (2003), and Bandura (2012) people with a higher level of self-efficacy try harder to complete the task successfully and be more ongoing with comparable task challenges. The results of the immediate and delayed test aligned with the growth in self-efficacy.

In addition, the participants rated their final self-efficacy significantly higher in the blocked than mixed practice condition. In other words, the blocked condition more strongly increased the students' confidence that they had learned to complete the trained tasks. The reason may be that students in the blocked practice condition trained on only one task at a time. Because a task video was immediately followed by task practice these students knew exactly what to do. In contrast, students in the mixed practice condition first viewed two or more videos, and then engaged in task practices. These students therefore had more to remember when they engaged in practice. They were facing more choices, which could make the students less confident.

CI-research on cognitive load was also scarce. The present study measured flow as an indicator of cognitive load. The results of the study found that the blocked practice condition outperformed the mixed practice condition on the delayed test. It seems fair to conclude that the blocked practice condition felt more committed with the learning tasks and felt like they were functioning at their fullest capacity (Shernoff, et al., 2014) more than the mixed practice condition during the delayed test. The flow findings support the contention of Yang and Tao (2015) that cognitive load in the blocked practice is lower for the delayed test than in the mixed practice condition. The tentative conclusion is therefore that a blocked practice condition runs less risk of cognitive overload during the delayed test. For the other three test (practice tasks during training, immediate test, and transfer test) the outcomes of both conditions are equal. No difference was found. Yang and Tao (2015) gave an explanation for this finding, namely that flow experience is stimulated when the teacher is recorded during the demonstration. The used instructional videos did not include a recording of the teacher. The consequence is that, students seemed less focused on what they were doing (Yang & Tao, 2015). Another explanation for the low flow experiences score is that students did not achieve the active state of mind, did not immersed in the learning materials, and missed the enjoyment in learning (Csikszentmihalyi, 1990).

Lastly, the results on task performance are discussed. In CI-research, blocked practice has been predicted (and found) to lead to better performance during training and lesser post training performances and transfer. The opposite effect is predicted (and found) for mixed practice, which will lead to better performance during post training and transfer, and a lesser performance during training (e.g., Helsdingen et al., 2011; Shea & Morgan, 1979; Lee & Simon, 2004; Wulf & Shea, 2002; Paas & Van Merriënboer, 1994; Cross et al., 2007; Guadagnoli & Lee, 2004; Simon, 2007; Welsher & Grierson, 2017; Neville & Trempe, 2017; Jamieson & Rogers, 2000; Broadbent et al., 2017; Schneider et al., 2002; Farrow & Buszard, 2017; De Croock et al., 1998). The findings of the present study did not align well with the CI-research findings. On practice tasks and an immediate and delayed test no significant

difference between conditions emerged. In addition, a contrasting outcome was obtained for the transfer test where the blocked practice condition significantly outperformed the mixed practice condition. An explanation could be that the experiment is obtained in a whole new domain, namely Microsoft Office Word. Due to the task difference between the experiment of this study compared to other studies, it might be that the results deviate from other research' results (e.g. Paas & Van Merriënboer, 1994; Cross et al., 2007; Guadagnoli & Lee, 2004; Simon, 2007; Welsher & Grierson, 2017; Neville & Trempe, 2017; Jamieson & Rogers, 2000; Broadbent et al., 2017; Schneider et al., 2002; Farrow & Buszard, 2017; De Croock et al., 1998). A second explanation could be that the blocked practice condition trained in a more structured manner, than the mixed practice condition, and therefore had more space in the working memory (Mayer, 2008) to acquire and give meaning to the procedural knowledge of the instructional videos, and apply it in other situations. Which will lead to a higher transfer score. Lastly, it might be that self-efficacy may have played a moderator role in this finding. A high self-efficacy level contributes to a person who is determined to complete the task successfully (e.g., Van der Meij, 2017; Bandura & Locke, 2003; Bandura, 2012). Because the self-efficacy level was higher for the blocked practice condition, those students might be more determined and tried harder to complete the transfer test than the students from the mixed practice condition. Besides, the given explanation also states that the construct 'self-efficacy' is an important factor in experimental studies about instructional video and practice schedules, because its inclusion might explain other results. Brar and Van der Meij (2017) also confirmed that self-efficacy has an important role in learning from a video. In short, the study shows that the blocked practice condition outperformed the mixed practice condition, especially on transfer tests.

One limitation of the study is that the experimenter was also the teacher of the children. Despite the fact that the experimenter said that she would not help the children with their task performances during the study, some students asked for help but were not given any. This may have affected their motivation during the experiment. The second limitation is that a relative small sample size is used. A bigger sample size would lead to stronger conclusions.

One recommendation for future research is to register participant actions on the video. Due to log registration the exact process of completing the tasks could be followed. Now, only the product is reviewed. Thereby, the factor of user control also could be registered and measured. Because the registered actions of participants show the process of pause, stop, wind, and rewind during watching the instructional videos, it would give more insight in the learning process of the participants. Besides, the log registration results might explain other results of the study. A second recommendation is to use an experimenter who is unknown by the participants.

In summary, the results are promising and give a whole new insight in DBT and the effectiveness of practice schedules on self-efficacy, flow experience, and task performance. This study does not align well with the CI-research found in other studies, namely the blocked practice condition outperformed the mixed practice condition based on transfer tests. Besides, new findings about self-efficacy during instructional video learning with different practice schedules are discovered. In short, the blocked practice condition outperformed the mixed practice condition on the final self-efficacy measure, the flow experience delayed test, and lastly on the task performance transfer test. Based on the findings from the present study the tentative recommendation is for a blocked practice design. However, more research is needed.

5. References

- Bandura, A. (1986). *Social foundations of thought and actions: A social cognitive theory*. Englewood Cliffs, NJ: Prentice Hall.
- Bandura, A. (1997). *Self-efficacy. The exercise of control*. New York: Freeman and Company.
- Bandura, A. (2012). On the functional properties of perceived self-efficacy revisited. *Journal of Management*, 38(1), 9-44. doi:10.1177/0149206311410606
- Bandura, A., & Locke, E. A. (2003). Negative self-efficacy and goal effects revisited. *Journal of Applied Psychology*, 88(1), 87-99. doi:10.1037/0021-9010.88.1.87
- Boucheix, J. M., & Guignard, H. (2005). What animated illustrations conditions can improve technical document comprehension in young students? Format, signaling and control of the presentation. *European Journal of Psychology of Education*, 20(4), 369-388. doi:10.1007/BF03173563
- Boucheix, J. M., & Lowe, R. (2010). An eye tracking comparison of external pointing cues and internal continuous cues in learning with complex animations. *Learning and Instruction*, 20, 123-135. doi:10.1016/j.learninstruc.2009.02.015
- Boucheix, J. M., Lowe, R. K., & Bugaiska, A. (2015). Age differences in learning from instructional animations. *Applied Cognitive Psychology*, 29, 524-535. doi:10.1002/acp.3131
- Brar, J., & Van der Meij, H. (2017). Complex software training: harnessing and optimizing video instruction. *Computers in Human Behavior*, 70, 475-485. doi:10.1016/j.chb.2017.01.014
- Broadbent, D. P., Causer, J., Williams, M. A., & Ford, P. R. (2017). The role of error processing in the contextual interference effect training of perceptual-cognitive skills. *Journal of experimental psychology: Human perception and performance*, 43(7), 1329-1342. doi:10.1037/xhp0000375
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
- Cross, E. S., Schmitt, P. J., & Grafton, S. T. (2007). Neural substrates of contextual interference during motor learning support a model of active preparation. *Journal of Cognitive Neuroscience*, 19, 1854-1871. doi:10.1162/jocn.2007.19.11.1854
- Csikszentmihalyi, M. (1990). *Flow: The psychology of optimal experience*. New York: Harper Perennial.
- De Croock, M. B. M., Van Merriënboer, J. J. G., & Paas, F. (1998). High vs. low contextual interference in simulation-based training of troubleshooting skills: Effects on transfer performance and invested mental effort. *Computers in Human Behavior*, 14, 249-267. doi:10.1016/S0747-5632(98)00005-3
- Ertelt, A., Renkl, A., & Spada, H. (2006). Making a difference: exploiting the full potential of instructionally designed on-screen videos. In S. Barab, K. Hay, & D. Hickey (Eds.), *Proceedings of the 7th international conference on Learning sciences* (pp. 154-160). Bloomington: International Society of the Learning Sciences.
- 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. Retrieved from <https://pdfs.semanticscholar.org/9d19/11406d46327d81e43c61a99ad019a3041f6f.pdf>
- Farrow, D., & Buszard, T. (2017). Exploring the applicability of the contextual interference effect in sports practice. *Progress in Brain Research*, 234, 69-83.

- doi:10.1016/bs.pbr.2017.07.002
- Fokides, E. (2017). Digital educational games and mathematics. Results of a case study in primary school settings. *Education and Information Technologies*, 23(2), 1-17. doi:10.1007/s1063
- Giannakos, M. N. (2013). Exploring the video-based learning research: A review of the literature. *British Journal of Educational Technology*, 44(6), 191–195. doi:10.1111/bjet.12070.
- Grossmann, R., Salas, E., Pavlas, D., & Rosen, M. A. (2013). Using Instructional Features to Enhance Demonstration-Based Training in Management Education. *Academy of Management Learning & Education*, 12(2), 219-243. doi:10.5465/amle.2011.0527
- Guadagnoli, M. A., & Lee, T. D. (2004). Challenge point: A framework for conceptualizing the effects of various practice conditions in motor learning. *Journal of Motor Behavior*, 36, 212–224. doi:10.3200/JMBR.36.2.212-224
- Guo, P. J., Kim, J., & Rubin, R. (2014). How video production affects student engagement: An empirical study of MOOC videos. In *Proceedings of the first ACM Conference on Learning@Scale Conference* (pp. 41-50). Atlanta, GA.
- Helsdingen, A., Van Gog, T., & Van Merriënboer, J. (2011). The Effects of Practice Schedule and Critical Thinking Prompts on Learning and Transfer of a Complex Judgment Task. *Journal of Educational Psychology*, 103(2), 383–398. Retrieved from: http://dspace.ou.nl/bitstream/1820/2911/6/helsdingen_2009-0424-RRR.pdf
- Hobbs, R. (1998). Teaching with and about film and television: Integrating media literacy concepts into management education. *Journal of Management Development*, 17, 259-272. doi:10.1108/02621719810210136
- Jamieson, B. A., & Rogers, W. A. (2000). Age-related effects of blocked and random practice schedules on learning a new technology. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, 55B(6), 343–353. doi:10.1093/geronb/55.6.P343
- Kim, M. K., Kim, S. O., Khera, O., & Getman, J. (2014). The experience of three flipped classrooms in an urban university: An exploration of design principles. *Internet and Higher Education*, 22, 37-50. doi:10.1016/j.iheduc.2014.04.003
- 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, 317-2007. doi:10.1080/15213260701532880
- Lee, T. D., & Magill, R. A. (1983). The locus of contextual interference in motor-skill acquisition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 9, 730–746. doi:10.1037/0278-7393.9.4.730
- Lee, T. D., & Magill, R. A. (1985). Can forgetting facilitate skill acquisition? In D. Goodman, R. B. Wilberg, & I. M. Franks (Eds.), *Differing perspectives in motor learning, memory, and control* (pp. 3–22). Amsterdam: Elsevier.
- Lee, T. D., & Simon, D. A. (2004). Contextual interference. In A. M. Williams, & N. J. Hodges (Eds.), *Skill acquisition in sport: Research, theory and practice* (pp. 29–44). London: Routledge.
- Lemarié, J., Lorch, R. F., Eyrolle, H., & Virbel, J. (2008). Sara: A text-based and reader-based theory of signaling. *Educational Psychologist*, 43(1), 27-48. doi:10.1080/00461520701756321
- Leppink, J., Paas, F., Van Gog, T., Van der Vleuten, C. P. M., & Van Merriënboer, J. J. G.

- (2014). Effects of pairs of problems and examples on task performance and different types of cognitive load. *Learning and Instruction*, 30, 32-42.
doi:10.1016/j.learninstruc.2013.12.001
- Lin, C-H. J., Fisher, B. E., Winstein, C. J., Wu, A. D., & Gordon, J. (2008). Contextual Interference Effect: Elaborative Processing or Forgetting—Reconstruction? A Post Hoc Analysis of Transcranial Magnetic Stimulation—Induced Effects on Motor Learning. *Journal of Motor Behavior*, 40(6), 578-586.
doi:10.3200/JMBR.40.6.578-586
- Lloyd, S. A., & Robertson, C. L. (2012). Screencast tutorials enhance student learning of statistics. *Teaching of Psychology*, 39(1), 67-71. doi:10.1177/0098628311430640
- Marx, R. D., & Frost, P. J. (1998). Toward optimal use of video in management education: examining the evidence. *Journal of Management Development*, 17(4), 243-150.
doi:10.1108/02621719810210154
- Margulieux, L., Guzdial, M., & Catrambone, R. (2012). Subgoal-labeled instructional material improves performance and transfer in learning to develop mobile applications. In *Paper presented at the 9th international conference on international computing education research (ICER'12)*. Auckland, New Zealand.
- Mayer, R. E. (2008). *Learning and instruction (2nd ed.)*. Upper Saddle River: Prentice Hall.
- Neville, K. M., & Trempe, M. (2017). Serial practice impairs motor skill consolidation. *Experimental Brain Research*, 235(9), 2601-2613. doi: 10.1007/s00221-017-4992-6
- Paas, F., & Van Merriënboer, J. J. G. (1994). Variability of worked examples and transfer of geometrical problem-solving skills: A cognitive-load approach. *Journal of Educational Psychology*, 86, 122–133. doi:10.1037/0022-0663.86.1.122
- Richter, J., Scheiter, K., & Eitel, A. (2015). Signaling text-picture relations in multimedia learning: A comprehensive meta-analysis. *Educational Research Review*, 17, 19-36.
doi:10.1037/edu0000220
- Rittle-Johnson, B., Siegler, R. S., & Alibali, M. W. (2001). Developing conceptual understanding and procedural skill in mathematics: An iterative process. *Journal of Educational Psychology*, 93, 346–362. doi:10.1037/0022-0663.93.2.346
- 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. doi:10.1177/ 0018720810381071.
- Schneider, V. I., Healy, A. F., & Bourne, L. E. (2002). What is learned under difficult conditions is hard to forget: Contextual interference effects in foreign vocabulary acquisition, retention and transfer. *Journal of Memory and Language*, 46, 419–440.
doi:10.1006/jmla.2001.2813
- Shernoff, D.J., Csikszentmihalyi, M., Schneider, B., & Shernoff, E. S. (2014). Engagement in High School Classrooms from the Perspective of Flow Theory. In M. Csikszentmihalyi (Eds.), *Applications of Flow in Human Development and Education* (pp. 475-494). Dordrecht: Springer.
- Simon, D. A. (2007). Contextual interference effects with two tasks. *Perceptual and Motor Skills*, 105, 177–183. doi:10.2466/pms.105.1.177-183
- Skillen, M.A. (2008). Promoting Thinking Skills within the Secondary Classroom Using Digital Media. In M. Kendall, & B. Samways (Eds.), *Learning to Live in the Knowledge Society* (pp. 119-126). Boston: Springer.
- Shea, J. B., & Morgan, R. L. (1979). Contextual interference effects on acquisition, retention

- and transfer of a motor skill. *Journal of Experimental Psychology: Human Learning and Memory*, 5, 179–187. doi: 10.1037/0278-7393.5.2.179
- 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. Retrieved from <http://www.ingentaconnect.com/content/stc/tc/2013/00000060/00000003/art00004>
- 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. doi:10.1016/j.compedu.2014.06.003
- Van der Meij, H. & Van der Meij, J. (2016). Demonstration-based training (DBT) in the design of a video tutorial for software training. *Instructional Science*, 44, 527-542. doi:10.1007/s11251-016-9394-9
- Van der Meij, H. (2017). Reviews in instructional video. *Computers & Education*, 114, 164-174. doi:10.1016/j.compedu.2017.07.002
- 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 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
- Welsher, A., & Grierson, L. E. M. (2017). Enhancing technical skill learning through interleaved mixed-model observational practice. *Advances in Health Sciences Education*, 22(5), 1201-1211. doi:10.1007%2Fs10459-017-9759-0
- Wistia. (2012). *Does length matter?* Retrieved from <http://wistia.com/blog/does-length-matter-it-does-for-video-2k12-edition>.
- Witteman, M. J., & Segers, E. (2010). The modality effect tested in children in a userpaced multimedia environment. *Journal of Computer Assisted Learning*, 26, 132-142. doi:10.1111/j.1365-2729.2009.00335.x
- 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
- Wulf, G., & Shea, C.H. (2002). Principles derived from the study of simple skills do not generalize to complex skill learning. *Psychological Bulletin & Review*, 9, 185–211. doi:10.3758/BF03196276
- Yang, J., & Tao, Y. (2015). Effects of different video types about procedural knowledge on cognitive load, learning flow and performance. *International Conference of Educational Innovation through Technology*, 175-179. doi: 10.1109/EITT.2015.44

6. Appendices

Appendix A

Content instructional videos

H1 Starting and saving a Word file

1.1 Starting an empty Word document

1.2 Open an existing Word file

1.3 Save a Word file

H2 Changing a text

2.1 Delete a part of a text

2.2 Replace a part of a text

2.3 Copy and paste a part of a text

H3 The use of pictures in a Word file

3.1 Add a picture in a Word file

3.2 Change the size of a picture

Appendix B*Design guidelines for DBT (Brar & Van der Meij, 2017; Van der Meij & Van der Meij, 2013)*

Fundamental process	Design guidelines
Attention	- Signaling/cueing - Pace - User Control - Preview
Retention	- Segmentation - Label - Simple-to-complex sequence - Review - pause
Production	- Practice - Arrangement - Practice files
Motivation	- Anchor in task domain - Conversational style - Length - Music

Appendix C

Experimenter script sessie 1:

Uitleg Self-efficacy test 1

Jullie hebben een blad met vragen voor je liggen. Je ziet een voorblad. Bij 'naam' schrijf je jouw voornaam en achternaam op. Bij nummer schrijf je niets op. Je slaat de bladzijde om. Je leest zelfstandig door wat er op dit blad staat. *Daarna lezen we het samen door*. Je gaat nu de vragen beantwoorden. Werk door tot er staat 'Je bent klaar'. Daarna pak je je leesboek en wacht je tot ik zeg dat je mag stoppen. Je hebt vijf minuten de tijd op deze test in te vullen.

Uitleg training

Vandaag wil ik met jullie een training gaan doen om te kijken hoe goed jullie zijn in het mooier maken van verslagen in Microsoft Word. Jullie zitten nu achter een tablet. Er ligt een handleiding voor je. Deze handleiding helpt je bij het maken van de opdrachten. Je moet precies doen wat er in de handleiding staat.

We nemen de handleiding samen door.

Op de website kun je de video's vinden die je nodig hebt om de opdrachten te maken. Je krijgt straks van mij een nummer. Dit nummer vul je in bij 'Gebruikersnaam' en bij 'Wachtwoord'.

De documenten die je nodig hebt voor de opdrachten staan al op de tablet.

Ik mag je niet helpen tijdens de opdrachten, ook als je het niet snapt, mag ik je niet helpen. Je moet het daarom zelfstandig doen. Als het niet lukt, is het niet erg. Daarnaast is het ook belangrijk dat je het zelfstandig doet en niet samen met een klasgenoot. Als je klaar bent met de training, steek je je vinger op en krijg je van mij een puzzelboekje waar je in kunt werken. Hier ga je mee aan de slag tot ik zeg dat je mag stoppen. In totaal heb je 45 minuten de tijd voor de training. Je probeert zo ver mogelijk te komen en alle opdrachten te doen. De tijd kun je zien op de time timer.

Zijn er vragen over de handleiding? Doet de tablet het bij iedereen?

Dan mag je nu beginnen, kijk goed in de handleiding.

De leerlingen gaan bezig en ik controleer of alles individueel gemaakt wordt

Jullie zijn nu klaar met de training. Controleer of je alles hebt ingevuld in de handleiding. De handleiding lever je bij mij in mét naam én nummer. Dan heb je nu tijd voor een kwartier pauze.

Uitleg self-efficacy test 2

Jullie hebben een blad met vragen voor je liggen. Je ziet een voorblad. Bij 'naam' schrijf je jouw voornaam en achternaam op. Bij nummer schrijf je het nummer op dat je van mij gekregen hebt. Je slaat de bladzijde om. Je leest zelfstandig door wat er op dit blad staat.

Daarna lezen we het samen door.

Je gaat nu de vragen beantwoorden. Werk door tot er staat 'Je bent klaar'. Daarna werk je in het puzzelboekje dat je van mij hebt gekregen. Je hebt vijf minuten de tijd voor deze test.

Uitleg immediate test

Je hebt nu een toets voor je liggen. In deze toets staan vragen en oefeningen die je net in de training geoefend hebt. Alle bestanden staan op je tablet. Ook moet je weer in het boekje

een paar vragen invullen aan het einde van elk hoofdstuk. Je slaat het boekje open en leest de eerste bladzijde zelf door. *Daarna nemen we het samen door*. Je hebt 20 minuten de tijd voor deze opdrachten.

Als je eerder klaar bent, ga je aan de slag met het puzzelboekje. Zijn er nog vragen voor je kunt beginnen?

Je bent klaar met de training over Word. Als je het niet af hebt gekregen, is het niet erg. We gaan er een andere keer opnieuw mee aan de slag.

Experimenter script sessie 2:

Uitleg delayed test

Je hebt nu een toets voor je liggen. In deze toets staan vragen en oefeningen die je vorige week in de training geoefend hebt. Alle bestanden staan op je tablet. Ook moet je weer in het boekje een paar vragen invullen aan het einde van elk hoofdstuk. Je slaat het boekje open en leest de eerste bladzijde zelf door. Je hebt 20 minuten de tijd voor deze opdrachten.

Daarna nemen we het samen door.

Als je eerder klaar bent, ga je aan de slag met het puzzelboekje. Zijn er nog vragen voor je kunt beginnen?

Uitleg transfer test

Je hebt net een toets gemaakt met vragen en oefeningen die je geoefend hebt in de training. Nu ga je een toets maken met oefeningen die niet zijn uitgelegd in de training. Deze oefeningen zou je wel kunnen weten, door de training die je gehad hebt. Ook moet je weer in het boekje een paar vragen invullen aan het einde van elk hoofdstuk. Je slaat het boekje open en leest de eerste bladzijde zelf door.

Daarna nemen we het samen door. Je hebt 15 minuten de tijd voor deze opdrachten.

Als je eerder klaar bent, ga je aan de slag met het puzzelboekje. Zijn er nog vragen voor je kunt beginnen?

Je bent klaar met de toetsen over Word.

Appendix D

The effect of practice schedules on flow experience per chapter

Table 6 shows the finding for the effect of practice schedules on the flow experience of the participants during the training for each chapter. An ANCOVA test for flow during the training of chapter 1 showed that there is no statistical significant difference in mean score between the two conditions, $F(1, 46) = 2.18, p = .147$. This means that the mean flow of the two conditions during the training on chapter 1 can be considered as equal.

An ANCOVA test for flow during the training of chapter 2 showed that there is no statistical significant difference in mean score between the two conditions, $F(1, 33) = .309, p = .582$. This means that the mean flow of the two conditions during the training on chapter 2 can be considered as equal.

An ANCOVA test for flow during the training of chapter 3 showed that there is no statistical significant difference in mean score between the two conditions, $F(1, 25) = .671, p = .421$. This means that the mean flow of the two conditions during the training on chapter 3 can be considered as equal.

In short, no ANCOVA test showed a significant difference in flow experience between the condition per chapter for the training.

Table 6. Mean score of flow experience during the training per chapter.

Condition	Chapter 1		Chapter 2		Chapter 3	
	Mean	(SD)	Mean	(SD)	Mean	(SD)
Blocked practice (n = 26, 23, 19) ^a	5.43	(1.23)	6.27	(.87)	6.57	(.51)
Mixed practice (n = 22, 12, 8) ^a	4.82	(1.66)	6.08	(1.10)	6.34	(.91)
Total (n = 48, 35, 27) ^a	5.15	(1.46)	6.21	(.94)	6.50	(.64)

^a The number of participants for flow experience during the training chapter 1, chapter 2, and chapter 3.

Table 7 shows the finding for the effect of practice schedules on the flow experience of the participants during the immediate test for each chapter. An ANCOVA test for flow during the immediate test of chapter 1 showed An ANCOVA test showed that there is no statistical significant difference in mean score between the two conditions, $F(1, 49) = 2.86, p = .097$. This means that the mean flow of the two conditions during the immediate test on chapter 1 can be considered as equal.

An ANCOVA test for flow during the immediate test of chapter 3 showed that there is a statistical significant difference in mean score between the two conditions, $F(1, 45) = 6.31, p = .016$. This means that the mean flow of the blocked practice condition during the immediate test on chapter 2 can be considered as higher than the mean flow of the mixed practice condition.

An ANCOVA test for flow during the immediate test of chapter 3 showed that there is no statistical significant difference in mean score between the two conditions, $F(1, 43) = 2.95, p = .093$. This means that the mean flow of the two conditions during the immediate test on chapter 3 can be considered as equal.

In short, no ANCOVA test showed a significant difference in flow experience between the condition per chapter for the immediate test, except for chapter 2.

Table 7. Mean score of flow experience during the immediate test per chapter.

Condition	Chapter 1		Chapter 2		Chapter 3	
	Mean	(SD)	Mean	(SD)	Mean	(SD)
Blocked practice (n = 26, 26, 25) ^a	6.54	(.71)	6.21	(1.37)	6.43	(.85)
Mixed practice (n = 25, 21, 20) ^a	6.12	(1.06)	5.02	(1.88)	5.90	(1.22)
Total (n = 51, 47, 45) ^a	6.34	(.91)	5.68	(1.70)	6.19	(1.05)

^a The number of participants for flow experience during the training chapter 1, chapter 2, and chapter 3.

Table 8 shows the finding for the effect of practice schedules on the flow experience of the participants during the delayed test for each chapter. An ANCOVA test for flow during the delayed test of chapter 1 showed that there is no statistical significant difference in mean score between the two conditions, $F(1) = .330$, $p = .569$. This means that the mean flow of the two conditions during the delayed test on chapter 1 can be considered as equal.

An ANCOVA test for flow during the delayed test of chapter 2 showed that there is no statistical significant difference in mean score between the two conditions, $F(1) = .659$, $p = .422$. This means that the mean flow of the two conditions during the delayed test on chapter 2 can be considered as equal.

An ANCOVA test for flow during the delayed test of chapter 3 showed that there is no statistical significant difference in mean score between the two conditions, $F(1) = 1.093$, $p = .304$. This means that the mean flow of the two conditions during the delayed test on chapter 3 can be considered as equal.

In short, no ANCOVA test showed a significant difference in flow experience between the condition per chapter for the delayed test.

Table 8. Mean score of flow experience during the delayed test per chapter.

Condition	Chapter 1		Chapter 2		Chapter 3	
	Mean	(SD)	Mean	(SD)	Mean	(SD)
Blocked practice (n = 24, 24, 20) ^a	6.54	(.69)	6.43	(1.26)	6.90	(.26)
Mixed practice (n = 15, 15, 13) ^a	6.67	(.61)	6.01	(1.33)	6.79	(.35)
Total (n = 39, 39, 33) ^a	6.59	(.66)	6.29	(1.28)	6.86	(.30)

^a The number of participants for flow experience during the delayed chapter 1, chapter 2, and chapter 3.