

## MASTER

### Evaluating the Incorporation of Microlearning into an Intelligent Tutoring System in the Work Environment

Philippens, Pam

*Award date:*  
2023

[Link to publication](#)

#### **Disclaimer**

This document contains a student thesis (bachelor's or master's), as authored by a student at Eindhoven University of Technology. Student theses are made available in the TU/e repository upon obtaining the required degree. The grade received is not published on the document as presented in the repository. The required complexity or quality of research of student theses may vary by program, and the required minimum study period may vary in duration.

#### **General rights**

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain

Eindhoven, 13-5-2022

# **Evaluating the Incorporation of Microlearning into an Intelligent Tutoring System in the Work Environment**

by Pam Philippens

Identity number: 0939011

in partial fulfilment of the requirements for the degree of

**Master of Science  
in Human-Technology Interaction**

Supervisors:

Dr. ir. M.A. (Rianne) Conijn

Dr. ir. P.A.M. (Peter) Ruijten - Dodoiu

Eindhoven University of Technology

## Table of Contents

Introduction .....	4
Literature Background .....	5
Intelligent Tutoring Systems (ITS's) .....	5
Learning Pace and Distributed Learning.....	7
Microlearning .....	8
Evaluating Learning Tools .....	10
Method .....	16
Design .....	16
Participants .....	16
Materials .....	17
Procedure .....	21
Measures .....	19
Level 1: Reaction.....	19
Level 2: Learning.....	19
Level 3: Behavior.....	20
Data Analysis.....	24
Results .....	25
Level 1: Reaction.....	25
Level 2: Learning .....	26
Level 3: Behavior .....	28
Discussion .....	31
Main effects of microlearning on evaluation.....	31
Additional predictors of evaluation .....	33
Limitations.....	34
Future research.....	35
Implications.....	35
Conclusion .....	35
Bibliography.....	38
Appendices .....	45
Appendix A.....	45
Appendix B.....	46
Appendix C .....	48

## Abstract

Work-related training, in order to stay up-to-date and relevant, continuously takes up valuable time from busy employees. These employees all have their own learning pace, which makes it difficult to generalize training. Intelligent Tutoring Systems offer a partial solution by providing instant, personal feedback to a learner. A learning method called ‘microlearning’ focuses on solving the problem of time consumption of work-related training. Learners process smaller bits of information in several short learning sessions, instead of large pieces of information at once. Microlearning could facilitate as a solution to lighten workload and improve training appreciation, engagement, knowledge retention, attitude, and desired behavior changes.

Existing research on microlearning showed that there is a gap in how these learning tools are evaluated, especially within a work environment. Often, only (subjective) aspects of the effects of a learning tool were studied. Kirkpatrick’s Evaluation Model (2006) is a renowned method for evaluating learning tools and consists of four levels – Reaction (i.e. appreciation and engagement), Learning (i.e. knowledge and attitude), Behavior, and Results. A literature review on the methods used to evaluate these levels indicated that the Behavior and Results levels are often neglected in research, as well as the usage of objective measures.

The current study investigated the difference in evaluation between learning with incorporated microlearning and learning without, using the first three levels of Kirkpatrick’s Evaluation Model. A microlearning group and a control group interacted with a learning tool for five times ten minutes, respectively for 45 minutes at once. For each evaluation level, both a subjective and objective measure was used. For the Reaction level, appreciation was measured by a subjective appreciation scale and the duration of learning sessions determined participants’ engagement. The Learning level was evaluated by pre- and post-knowledge tests and a subjective attitude scale. Lastly, two subjective behavior scales and objective clicks on fake phishing e-mails measured the Behavior level.

The results showed a small effect of microlearning on the appreciation of the training. The other expected results of increased engagement, knowledge, attitude, and behavior changes were not found. This outcome can probably be partly explained due to limitations in sample size, the used learning tool, and the learned topic. The current study added several aspects to existing research, such as a literature review on methods used for the evaluation of the levels of Kirkpatrick’s model, using objective measures and a control group, and incorporating microlearning into an Intelligent Tutoring System. To conclude, more research is needed to investigate microlearning’s potential for improving training and decreasing training load for every-busy employees.

## Introduction

The constantly developing nature of organizations and markets results in the aim to continuously improve job performances of employers and employees. In many fields, companies therefore expect their employees to be up-to-date at all times, regarding knowledge and skills relevant to their job (Sabitov et al., 2017; Mayer & Solga, 2008). It is critical for organizations to stay competitive within their market (Tumi et al., 2022) and this results in high amounts of training for employees with the goal of developing, improving, or maintaining certain levels of performance (Mak & Sockel, 1999).

Intuitively, every person learns in their own way and at their own pace. Some people easily master a topic, whilst their colleagues need more time to grasp the concepts. E-learning tools are suitable for high amounts of training and learning at one's own pace (Lau et al., 2013). It is therefore important to study these individual learning needs and preferences, when designing new learning tools (OECD, 2006). Intelligent Tutoring Systems (ITSs) are e-learning tools that aim to take these personal differences into account by granting instant feedback and personalized guidance to each individual learner (Tacoma et al., 2021).

Employees often do not have time for numerous long training sessions, next to the actual tasks in their job description (Emerson & Berge, 2018). Due to this, learning in this context should be low in time consumption. A concept within the e-learning domain that aims to tackle the issue of time consumption is microlearning. This way of learning concerns processing smaller bits of information in multiple short sessions, instead of receiving large pieces of information at once (Shail, 2019). In this way, learners are able to distribute their learning over a longer time period, at personally convenient times (Hug et al., 2006). Adding the microlearning method to Intelligent Tutoring Systems within the work field could result in a learning tool that allows employees to be up-to-date on relevant knowledge and skills, at their own pace, with personal feedback, and with low time consumption. This could additionally have the benefit of increasing employees' motivation regarding their learning (Upneja & Ozdemir, 2014).

Due to the importance of training in a work environment, it is equally essential to investigate how people learn, or desire to learn, and how training can be constructed according to these needs. Therefore, training tools should be evaluated in order to secure that they are appreciated by the learners, efficient, and deliver the desired results (Mann, 1996). Currently, a gap exists in research on the evaluation of microlearning tools in a work environment. Scientific papers either do not evaluate learning tools at all, or only on one or two basic aspects, such as employees' response to the training (e.g. Zandbergs et al., 2021; Hesse et al., 2019). With the goal of filling this research gap, the current study aimed to provide an extensive evaluation of a microlearning tool in a work environment. The evaluation was done by using one of the most popular learning tool evaluation techniques: Kirkpatrick's Evaluation Model (Sim, 2017; Kirkpatrick, 2006). This model can be used to evaluate a learning tool on four levels: Reaction, Learning, Behavior, and Results. In order to put the evaluation of microlearning into perspective, a version of the learning tool without the microlearning aspect was evaluated in the same manner.

The goal of the current paper was to use Kirkpatrick's model to investigate whether an Intelligent Tutoring System that includes the concept of microlearning is evaluated differently than

an ITS without microlearning, specifically in a work environment. Therefore, the following research question was formulated:

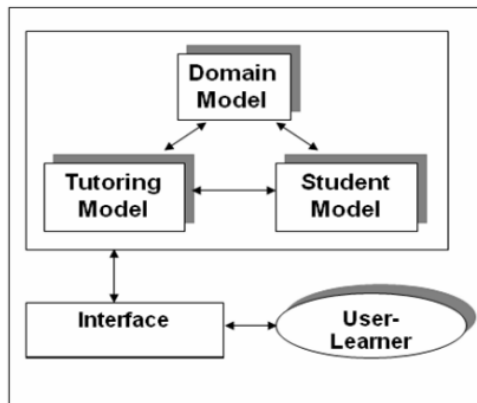
What is the difference in evaluation between an ITS with and without microlearning, in a work environment?

## Literature Background

### Intelligent Tutoring Systems (ITS's)

E-learning tools that aim to take individual preferences and needs into account are Intelligent Tutoring Systems (ITSs). These systems allow the learning pace of individual users to diverge, by granting instant feedback and personalized guidance for each learner (Tacoma et al., 2021). An important additional factor of these systems is that there is no human involved in the learning (e.g. in the role of a teacher or assistant; Ramesh & Rao, 2012). An Intelligent Tutoring System generally consists of four components: the domain/expert model, the tutoring model, the student model, and the user interface. The expert module contains the information to be taught to the learner. It functions as a standard measure for the evaluation of the user's performance (Nkambou et al., 2010). The tutoring model is in charge of the learning by the user. It responds to the user's actions in the learning tool, by providing corrections, suggestions, or checks (Ukamaka et al., 2021). In its turn, the student model stores information about the learner's progress and learning process. Its goal is to create an accurate overview of an individual's knowledge and skills, at any moment. Together with the tutoring model, and whilst comparing to the domain model, the student module determines where the gaps or errors in the user's knowledge are and what actions will be taken to solve these (Nkambou et al., 2010). Finally, the user interface is in charge of the actual interaction between the Intelligent Tutoring System and the learner.

In the schematic overview of an ITS (Figure 1), it becomes clear that the domain model, tutoring model, and student model interact directly with each other. The tutoring model then communicates with the user interface to determine what content will be displayed to the learner. Ultimately, the user of the learning tool (i.e. the learner) directly interacts with the user interface only. Due to this structure of four modules, an ITS is able to take an individual's learning pace into consideration. The modules shape and personalize the learning and the interface, according to the current state of the user's learning. Learners who experience more difficulty with the content of the tool will interact with a different user interface for a different amount of time, than learners who have already mastered the topics.

**Figure 1***Schematic overview of an Intelligent Tutoring System*

Note. Adapted from Nkambou, R., Bourdeau, J., and Psyché, V. (2010).

An additional advantage of Intelligent Tutoring Systems is the fact that there is not necessarily a teacher physically present or actively involved in the learning process. Therefore, employees do not have to be present in an offline classroom for their training, which means that training sessions can be followed from any place, at any time (Fletcher, 2003). Another benefit is that ITSs are a form of one-on-one teaching, which has proven to have relatively high levels of intensity and interactivity. Graesser and Person (1994) showed that teachers and students ask each other considerably more questions in the one-on-one teaching sessions, compared to the classroom setting, which could be linked to higher engagement. Intuitively, it would be too expensive and impractical to provide personal training sessions for each employee of an organization. Therefore, Intelligent Tutoring Systems offer comparable benefits of one-on-one training, without the actual teachers needed for each session (Fletcher, 2003; Stevens et al., 2009).

In the past years, the implementation of Intelligent Tutoring Systems has become increasingly popular, especially within school systems (Naser, 2008). The application of ITSs in elementary or higher education has therefore been a widely researched topic in academic papers, with overall promising results. Examples are the development of an ITS to aid students in learning programming languages (Sykes & Franek, 2003; Naser, 2008; Hooshyar et al., 2018), investigating the effect of voluntary use of or collaboration with an ITS on students' learning (Mitrović & Holland, 2020; Olsen et al., 2016), exploring the relation between a student's behavior within an ITS and their personality traits (Erickson et al., 2019; Sreenivasa Sarma & Ravindran, 2007), and studying how an ITS can improve student's ability to self-asses (Roll et al., 2011). Pane et al. (2014) found that students' test scores significantly increased after time, when their teacher introduced an ITS in a course. Similarly, Koedinger et al. (1997) implemented an ITS in high school classes. They observed that the students who learned with the ITS scored better in 100% of the cases, than the control group. However, both studies did not evaluate to what extent the students were more satisfied with this way of learning and whether they were able to (better) apply the gained knowledge in practice, compared to a control group.

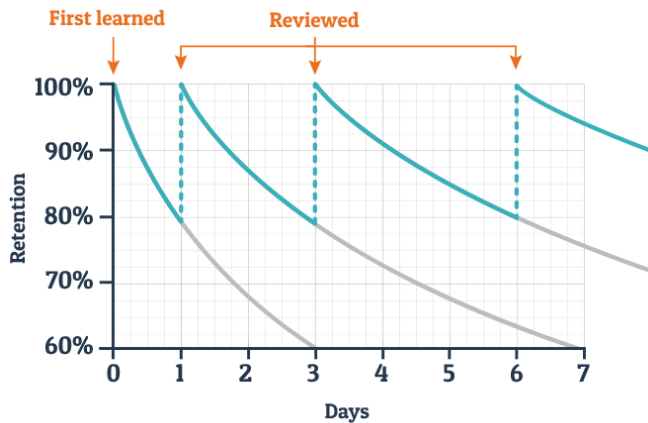
In contrast, the incorporation of Intelligent Tutoring Systems specifically within the work environment has received less attention from research in the past years. Nevertheless, there exist some studies regarding this application. In 2009, for example, Amokrane and Lourdeaux developed an ITS with an additional module that detects causalities between mistakes made by employees and possible risks at industrial sites with a relatively large amount of dangerous substances. However, they did not yet run an experiment to verify their module. Another example is the study by Pokorny et al. (1996), in which aircraft maintenance technicians were subjected to training on intervention and troubleshooting. One group followed instructions from an ITS, whilst the other technicians got informal on-the-job instructions. The results of the study showed that the group of employees that had experienced the ITS was more inclined to ensure the faultiness of parts, before replacing these, compared to the other experimental group. The extent to which the technicians were pleased with being trained by the ITS or whether this resulted in safer aircrafts or lower maintenance costs was not evaluated.

### Learning Pace and Distributed Learning

Regardless of where or when learning takes place, every individual learns in their own way and at a different pace. Varying per topic or skill, each employee understands or masters them at their own pace. Research has aimed investigating to what extent personalized learning paces and distribution of learning moments are able to improve an individual's learning experience. Tullis and Benjamin (2011) found that learners who were able to determine their own pace and timing of learning sessions scored significantly higher on a knowledge retention test, compared to individuals whose learning time was determined by the experimenter. This effect even occurred when the total amount of learning was equal for both the self-paced learning group, as well as the control group. The same effect was found in a study by Mazzoni and Cornoldi (1993), where participants were asked to study words and recall them. The experimental group that was allowed to determine their own learning pace was able to recall more of the studied words than the control group whose pace was determined by an average. In 1885, Ebbinghaus created his renowned 'Forgetting Curve'. He stated that people forget anything they learn after 24 hours, if they do not actively focus on retaining the knowledge. The rate at which one forgets decreases with each repetition of the knowledge (Figure 2). Thus, it seems that personalized, distributed learning paces are able to improve an individual's knowledge retention and, consequently, their learning performance.

When specifically focusing on learning and training within a work environment, it is seen that employees' time schedules are often completely saturated with meetings and tasks. One of the main difficulties that individuals experience with learning is the subjective lack of time for the activity, next to the actual tasks in their job description (Emerson & Berge, 2018). More traditional training can significantly increase employees' workload, which could eventually result in a decrease in job performance and an increase in job-related stress (Andriana et al., 2019; Van Ruysseveldt & Van Dijke, 2011). Due to this, it is argued that learning in a work context should be relatively low in time consumption, in order to minimize extra workload for employees. Reserving larger periods of time for learning can be challenging, while shorter moments can be easier to realize (Hug et al., 2006).



**Figure 2***Ebbinghaus' forgetting curve***Typical Forgetting Curve for Newly Learned Information**

Note. From Chun & Heo (2018).

**Microlearning**

A type of distributed learning that deals with time consumption is microlearning. Within literature, there seem to be a number of different definitions of this term. Many papers state that this way of teaching provides learners with smaller pieces of information or examinations over a longer period of time, compared to more traditional learning methods where learners get large chunks of information at once, during a (longer) training session or lecture (Shail, 2019). Other studies do not emphasize the longer time period, but argue that e.g. watching one short video can already be considered microlearning (Boring, 2020). Since the former seems to be used in most literature on microlearning, the current study uses this first definition. Hug et al. (2006) emphasize three main characteristics and advantages of the incorporation of microlearning in any field. The first aspect is to reduce large and complex pieces of information to relatively small and comprehensible bits. Then, microlearning enables individuals to deal with extensive amounts of new knowledge in a manageable, motivating manner. Finally, this way of learning allows persons to be in charge of their learning process. They are able to decide when and where they will engage in learning activities. This complements the goal of an Intelligent Tutoring System: to allow learners to learn at their own pace.

Since the essence of microlearning is that learning sessions are short, this approach allows learners to learn in multiple short time frames, whenever it is convenient to them (Hug et al., 2006). Regarding the length of the learning sessions, research is not yet in agreement. Some articles suggest that a microlearning session should take between three and five minutes (Dixit et al., 2021; Jahnke et al., 2020), whereas other papers define time frames of approximately fifteen minutes (Emerson & Berge, 2018).

Another important and relevant application of microlearning is assisting lifelong learning. As employees are expected to be continuously increasing, retraining, and updating their knowledge and skills, there exists no end date for their learning. As long as they are employed (i.e. usually a large part of their life), they will have to participate in training and courses. Microlearning can be an

ideal tool for lifelong learning, as the short sessions with small pieces of information can be incorporated into one's weekly job routine without too much interference for long periods of time (Chai-Arayalert & Puttinaovarat, 2020).

The incorporation of microlearning specifically within work environments and corporate organizations has only been investigated in a modest amount of scientific research. Puah et al. (2021) studied employees' likelihood of participating in microlearning activities. They found that positive opinions of colleagues regarding this type of learning and one's own feeling of being able to successfully engage in microlearning were important factors contributing to the extent to which employees' engaged in the training. The researchers therefore advise to consider community elements when encouraging employees to participate in microlearning. Zandbergs et al. (2021) investigated to what extent learning analytics could be a tool for predicting results of microlearning at work. By analyzing learning data, they aimed to predict the outcomes of microlearning training. They showed that participating in microlearning training, and positively evaluating this activity, could lead to higher test scores. Even though the study was a pilot with few participants ( $N=16$ ) and the data was not statistically analyzed, the researchers concluded that analyzing learners' data could be interesting to use for improvements in designing microlearning training at work. They did not evaluate to what extent these higher test scores would lead to relevant behavior change in the participants. In the context of working in the dairy industry, Hesse et al. (2019) invited employees to participate in a number of microlearning sessions to optimize job performance. Afterwards, approximately eighty percent of the learners reported that the accuracy with which they conducted their work tasks had increased, as well as their confidence to successfully complete these tasks. Additionally, they found subjective indications of high engagement in most participants. Besides these subjective remarks, there were no objective measures used to evaluate the microlearning sessions, nor a control group. Orwoll et al. (2017) asked nurses to use an online application for self-assessment. The study focused on a specific type of bloodstream infection and it was stated that these can be prevented by following specific practice standards. The application used gamification and microlearning to instruct and inform the nurses on the prevention practices. The results showed that the total amount of infections for the nurse group who used the application had decreased by 48%, while the control group did not show significant changes. The extent to which the nurses were satisfied with the tool or whether they had actually retained the knowledge of the prevention methods was not investigated. Similarly, the paper discussed group results only, instead of also looking at nurses' individual infection rate changes.

Besides the work floor, the microlearning approach is occasionally found to be incorporated in a school environment. This application has recently become more popular due to the COVID-19 pandemic, because the amount of online, remote education has rapidly increased. Research on microlearning used in courses at schools states that students' reaction to this type of learning is mostly positive (Dixit et al., 2020; Tolstikh et al., 2021). A study by Javorcik and Polasek (2019) investigated the application of microlearning at a school for future teachers and compared this to another e-learning tool. The researchers found that the microlearning students had more efficient learning sessions, even though they had studied fewer hours in total, compared to the other e-

learning group. Similar to the paper by Zandbergs et al. (2021), it was not investigated to what extent the students would be able to apply the learned knowledge in real life.

Reviewing existing literature, regarding both work and school context, it was concluded that no papers were found that did an extensive investigation of the incorporation of microlearning. First of all, the majority either asked learners for their opinion on this way of learning or whether the learning goals were met. However, a combination of these two evaluations or whether behavior changes occurred, as a result of the microlearning training, have mostly been neglected in research. Secondly, most papers used only subjective methods for evaluation, while neglecting objective methods. Finally, a control group for comparison of results was missing in most studies. In order to determine whether microlearning tools can be more beneficial for work training than other (more traditional) ways of learning, we argue that it is of importance that both methods are evaluated on multiple aspects, using sub- and objective measures, and are sequentially compared.

### Evaluating Learning Tools

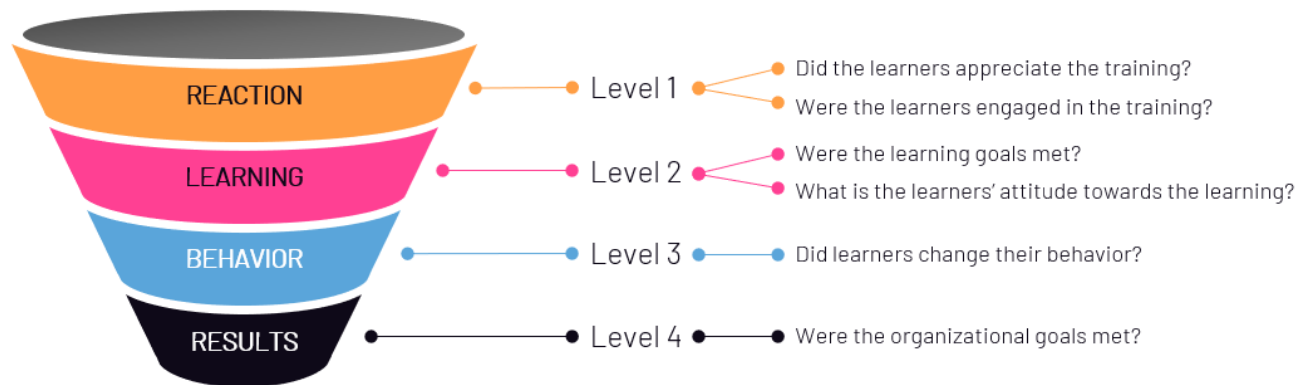
For decades, researchers have studied the evaluation of learning tools. The question aimed to be answered was how one could best prove or discuss the effectiveness and appreciation of a (new) way of learning. One of the most used methods for evaluating learning tools is Kirkpatrick's Four-Level Evaluation Model (1976, Figure 3). In short, this model describes four levels on which a learning tool should be evaluated; namely Reaction, Learning, Behavior, and Results. The Reaction level concerns the learners' appreciation of and engagement with the learning tool. The Learning level evaluates whether the learning goals have been met and the learners' attitude toward the learning. To what extent behavior change occurs after the learning is measured by the Behavior level. Lastly, the Results level focuses on organizational-wide results as an effect of the implemented learning tool.

Kirkpatrick's model is recognized as the model that is used most often by instances with renowned training programs, in the United States (United States Office of Personnel Management, 2008). The reason for this popularity is the fact that the model does not only consider the direct learning of a participant, but also what happens after the training or course, in terms of behavior changes and broader impact (Sim, 2017). Bates (2004) praises the model for being able to simplify the usually complex process of evaluation. He explains that it enables organizations to generate concrete points for improvement, optimization, or additions that the learning tool needs.

There exists an extensive amount of scientific papers that have made use of this model, for a large variety of learning tools and forms. In order to get a general idea of which levels of the model are used in research and how they were measured, a literature review was conducted. In total, 25 papers were reviewed that evaluated a learning tool (on multiple aspects), which was not necessarily a microlearning tool. In some papers, Kirkpatrick was not explicitly mentioned, but (some of) the four levels could be intuitively distinguished. For each level, the used method has been noted in an overview in Table 1. It is indicated whether the measure occurred before (pre) or after (post) the learning and whether the measure was subjective or objective.

**Figure 3**

*Illustration of Kirkpatrick's Four-Level Evaluation Model*



# Evaluating the Incorporation of Microlearning into an Intelligent Tutoring System in the Work Environment

**Table 1**

Overview of methods used per level of Kirkpatrick's Evaluation Model

	Level 1						Level 2				Level 3						Level 4					
	Survey by trainee	Survey by trainee	Survey by trainer	Daily reaction sheets	Learning Activity (engagement)	Interview/Focus group	Knowledge test	Knowledge test	Subjective survey (attitude)	Interview	Self-report survey	Interview trainee	Supervisor focus groups	Incident report	Observations by experimenter	Survey Supervisor	Survey supervisor	Financial results observation	Self-report survey	Interview trainee	Interview supervisor	Focus group with organization
	Subj.	Subj.	Subj.	Subj.	Obj.	Subj.	Subj./Obj.	Subj./Obj.	Subj.	Subj.	Subj.	Subj.	Subj.	Obj.	Obj.	Subj.	Subj.	Obj.	Subj.	Subj.	Subj.	Subj.
	Post	Pre + post	Post	During		Post	Pre + post	Post			Post	Post		Pre + post		Post	Post		Post	Post	Post	Post
Attia (1998)	X		X						X		X					X	X		X			
Bledsoe (1999)	X						X				X						X					
Larsen (1985)				X			X					X									X	
Lockwood (2001)	X							X								X		X				
Yaw (2005)	X						X						X	X								
Chen et al. (2015)	X							X														
Konstan et al. (2015)		X					X		X		X											
Jacquet et al. (2018)	X						X															
Liang et al. (2014)	X				X			X														
Cross (2013)		X			X		X		X		X				X							
Colvin et al. (2014)							X															
MacKay et al. (2016)	X				X		X		X													
Brunton et al. (2017)	X							X	X													
Rubio (2015)					X		X															
Stephens et al. (2014)		X							X													
Alturkistani et al. (2018)						X				X												
Hossain et al. (2015)	X*						X		X													
Mackness et al. (2013)	X					X**																
Milligan et al. (2014)						X				X		X										
Hsu et al. (2021)						X				X		X								X		
Naal et al. (2021)	X						X			X		X										X
Alsuwaidi et al. (2021)	X						X															
Tamani et al. (2021)	X		X						X		X											
Butler-Warke et al. (2021)	X					X			X	X	X	X							X			
Alsalamah et al. (2021)	X								X		X									X		
Total (25):	16	3	2	1	4	5	12	4	10	5	7	5	1	1	1	2	2	1	2	2	1	1

Note. No cross in any box for a certain level means that this level was not evaluated in that paper. \* Only one question referred to level 1 (in this case, the relevance of the learning tool). \*\* Both one-on-one interviews and focus groups were conducted.

It is notable that there is a wide variety of methods that have been used for each level. However, not every level was evaluated in each paper. Except for one paper each, levels one and two were always considered, but level four was most often neglected. Reasons for experimenters to skip the evaluation of the Behavior and Results levels are usually because of misconceptions and lack of expertise regarding the evaluation of learning tools (Bomberger, 2003; Phillips, 2000). Staff members assume that higher-level evaluations are too complex and therefore are often not considered, as long as employees react positively to the training and the organization remains financially healthy (Reio et al., 2017).

Below, the results of the literature review on evaluation methods are discussed per level. Subsequently, it is shortly discussed what we currently know about the evaluation of microlearning on each level, based on the discussed existing literature. This information was used to form the hypotheses of the current study.

### *Reaction Level*

#### Methods used in general

The Reaction level regards learners' appreciation of aspects such as the quality and relevance of the tool. In 2006, the aspect of engagement of the learner was added to this level (Kirkpatrick & Kirkpatrick, 2006). For this level, one method was clearly shown to be the most popular (used in 16/25 papers): a post-experiment survey with subjective questions, to be filled in by the trainee/learner (e.g. "Did you find the trainer capable of properly explaining the material?"). Sometimes in addition to the survey, interviews and focus groups were used (five papers). Engagement was measured in only four papers (but five papers were published before this aspect was added to the Reaction level).

#### Microlearning and the Reaction level

Existing literature on appreciation of microlearning tools has shown that microlearning was appreciated in multiple contexts, compared to more traditional teaching methods (Zandbergs et al., 2021; Dixit et al., 2020; Tolstikh et al., 2021). The effects of microlearning on learners' engagement in a work environment seem to be understudied in existing research. Nevertheless, in other contexts, microlearning was found to positively impact how engaged students were in courses (Hesse et al., 2019). This leads to the first two hypotheses of the current study:

*H1a: The Intelligent Tutoring System with incorporated microlearning approach is more appreciated, compared to the ITS without microlearning.*

*H1b: The Intelligent Tutoring System with incorporated microlearning approach results in more engagement, compared to the ITS without microlearning.*

### *Learning Level*

#### Methods used in general

The Learning level concerns the effectiveness of the learning that has taken place during the interaction with the learning tool, in terms of knowledge and skills. Later, a learner's attitude and

motivation towards performing the desired behavior change became an additional part of this level (Kirkpatrick & Kirkpatrick, 2006). This means that this level does not only consider the extent to which the learning goals have been met, but also whether learners are planning to convert this knowledge into action. Regarding the Learning level, there are two methods that were used most frequently. Approximately half of the studies (12/25) subjected their participants to a pre- and post-knowledge test with either objective (e.g. “What is the correct definition of this term?”) or subjective (e.g. “Do you think you have gained knowledge during this training?”) questions. Ten studies (also) made use of a subjective survey on the learners’ attitude toward applying their gained knowledge in practice (e.g. “Are you motivated to use the skills you have learned on a daily basis?”). Less popular methods were post-knowledge tests only (four papers) and interviews (five papers).

#### Microlearning and the Learning level

Research on the Learning level of microlearning suggests that microlearning can result in higher test scores (Zandbergs et al., 2021) and more efficient learning sessions (Javorcik & Polasek, 2019). Additionally, microlearning has been found to lead to a more confident attitude regarding the execution of trained tasks (Hesse et al., 2019). Therefore, the following hypotheses have been formulated:

*H2a: The Intelligent Tutoring System with incorporated microlearning approach better achieves its learning goals, compared to the ITS without microlearning.*

*H2b: The Intelligent Tutoring System with incorporated microlearning approach results in a more positive learning attitude, compared to the ITS without microlearning.*

#### Behavior Level

##### Methods used in general

The level of Behavior assesses the extent to which change in behavior has occurred, as a result of interaction with the learning tool. In other words, this level evaluates whether the knowledge and skills of level two have been applied in practice. In the literature review, there was less consensus on which method to use for evaluation of this level. First of all, only 13 of the 25 papers evaluated their learning tool on the Behavior level. Seven studies made use of a self-report survey after the experiment (e.g. “I believe that I have become more efficient in completing my tasks”) and five studies asked similar questions, but in an interview setting. Other methods varied from observations by the experimenter (one paper) to subjective supervisor surveys (two papers).

#### Microlearning and the Behavior level

The limited amount of existing research on behavior change as a result of microlearning shows that when microlearning is incorporated into work-related training, this can result in desired (subjective) changes in trainees’ behavior, as a result of the learning (Hesse et al., 2019). This leads to the final hypothesis of the current study:

*H3: The Intelligent Tutoring System with incorporated microlearning approach better achieves the intended behavior changes, compared to the ITS without microlearning.*

### *Results Level*

#### Methods used in general

Finally, the fourth level investigates the Results of a tool. It considers the impact of individuals' changed behavior (i.e. level three) on organizational goals. Examples of those goals are workplace productivity or cost reduction. The level is usually used to assess the total value and worth of the learning tool for organizations (Sim, 2017). The Results level was considered in only 9 of the 25 papers. For this level, there were no methods that stood out in terms of popularity. Some studies subjected supervisors to a survey or interview (e.g. "Do you believe that the general job performance of your team has increased?"), while others asked the participants to self-report (e.g. "Do you believe that you and your colleagues have managed to achieve the desired cost reduction?")'.

#### Microlearning and the Results level

The minimal amount of microlearning studies that considered this level showed that microlearning tools can have the ability to change behavior in groups of nurses and, therefore, reach the organization-wide goal of lower infection rates (Orwoll et al., 2017). Nevertheless, due to time constraints, the fourth level was not considered in the experiment of this study.



## Method

### Design

To answer the research question, an online experiment was conducted. The experiment followed a between-subject design with a microlearning condition and control condition. The control condition consisted of one learning session of 45 minutes, whilst the microlearning condition contained five (micro)learning sessions of ten minutes each (Table 2). Therefore, all participants interacted with the same learning tool, but the people in the microlearning condition experienced microlearning and the control group did not. The reason for this was to compare how a group of participants evaluated the learning tool, with or without the microlearning aspect. Regarding the evaluation, the same measures were used for both groups. For each of the first three levels of Kirkpatrick's Evaluation Model (i.e. Reaction, Learning, and Behavior), one subjective and one objective measure was used. As explained, the independent between-factor was the microlearning component. The dependent factors were the participants' appreciation, engagement, scores on the knowledge tests, attitude, and behavior (changes). The knowledge tests and two behavior change measures were repeated measures and within factors.

**Table 2**

*Overview of the two experimental conditions*

	Control group	Microlearning group
Learning tool	PowerApp	PowerApp
Number of knowledge tests	3	3
Number of questionnaires	3	3
Number of learning sessions	1	5
Length of one learning session	45 minutes	10 minutes
Number of comeback sessions	1	1
Total duration of the experiment	Two weeks	Four weeks

### Participants

The participants of this experiment were either students from Eindhoven University of Technology (TU/e) or employees of the company Conclusion B.V. in Utrecht. The students were recruited through the university's online experiment participation database, while the employees responded to an invite post on an internal company platform. In total, 52 participants were recruited. Each participant was rewarded with €10,- if they completed the first part of the experiment, and an additional €5,- when the final part was completed as well. The students could also choose course credits, which are needed to pass a compulsory course on psychological research methods of a major program at the TU/e, instead of the money.

After the experiment, the data of 47 participants could be used for further analysis, because five people had already withdrawn their participation before the experiment had started. Of those

47, 43 participants completed the entire experiment and four dropped out at some point during the experiment. Twelve participants were Conclusion B.V. employees and 31 were TU/e students. Twenty-eight participants were between 17 and 25 years old. Sixteen people identified as female, while 27 participants identified as male (Table 3).

**Table 3**

*Descriptive statistics of the participants' demographics, per condition*

Variable		Control	Microlearning	Total
Age	17-25 yrs	14	14	28
	25-35 yrs	4	5	9
	35-45 yrs	1	-	1
	45-60 yrs	3	2	5
Gender	Female	7	9	16
	Male	15	12	27
Working status	Student	15	14	29
	Employee	7	7	14

Due to the fact that we used a Dutch learning tool and therefore Dutch knowledge tests, but English questionnaires, one of the selection criteria was that the participants had to be fluent in both Dutch and English. In the TU/e's database, a few other inclusion criteria were listed. First of all, the participant's age had to be maximum 60 years old, in order to avoid digital illiteracy and ensure that they were not retired. Similarly, only participants were invited who had self-evaluated their 'computer experience' with four out of five. Finally, no dyslexic participants were invited, due to the relatively large amount of texts in the learning tool, knowledge tests, and questionnaires.

An a-priori power analysis was performed using the G\*Power application (version 3.1.9.4). Nikou and Economides (2018) investigated the effect of Mobile-Based micro-Learning and Assessment (MBmLA) on learning performance, compared to conventional approaches in high-school students. They found a statistically significant difference in factual knowledge recall between students who followed an MBmLA trajectory and the control group with  $\eta^2 = 0,07$ , which indicates a moderate effect size. Entering this effect size (transformed to Cohen's  $d = 0,55$ ) into G\*Power, with  $\alpha = 0,05$  and power = 0,90, gave a total amount of 116 participants divided into two equal groups. Based on feasibility (i.e. time and money constraints), we decided upon recruiting 50 participants divided into two equal groups. This number of participants, combined with the expected effect size of  $d = 0,55$  and  $\alpha = 0,05$ , led to a power of 0,61 (according to G\*Power). Even though a power of 61% is less than generally desired, we argued that this was a risk worth taking.

## Materials

### PowerApp

The learning tool that the participants interacted with is the PowerApp (Appendix A, Figure A1). This application is made and owned by the company Bright Alley, part of Conclusion B.V. The goal of the PowerApp is to assist in work-related learning for employees. Users can learn with the application in short sessions at times that are convenient to them and about content that their employers have

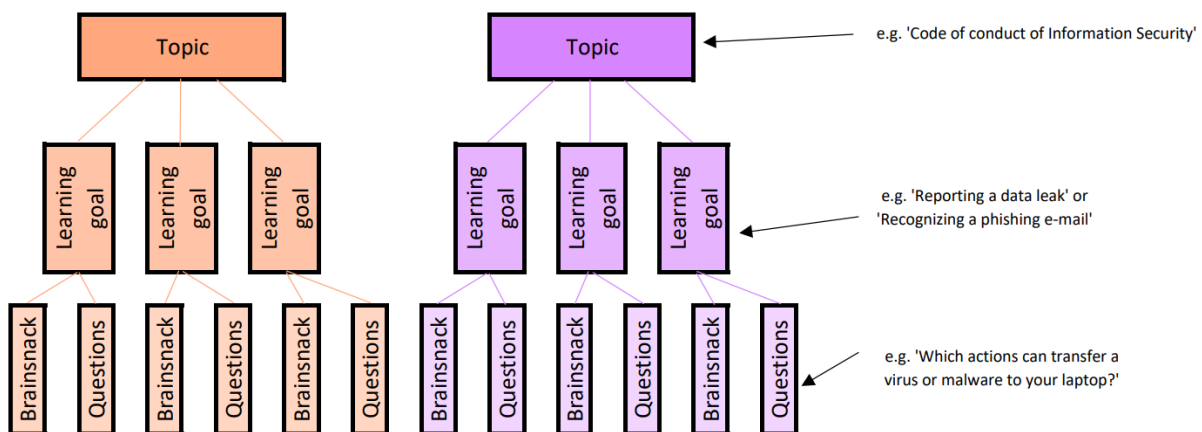
chosen to be relevant. The PowerApp can be labeled as an Intelligent Tutoring System, due to the fact that it provides instant, personal feedback to its users. The algorithm behind the PowerApp is based on Ebbinghaus' forgetting curve (1885). Because the rate at which a learner forgets knowledge decreases with each repetition, sessions can become less frequent as time progresses (Figure 2).

Any content entered in the PowerApp is normally categorized into several topics. Each topic is then divided into a number of learning goals. If a user meets all learning goals, they have 'mastered' the corresponding topic. Finally, a learning goal is taught by means of short texts of information (called 'brain snacks') and tested by short questions in different forms (e.g. multiple-choice, multiple select, slider). Figure 4 presents a general overview of content in the PowerApp. Answering a question correctly provides a notification, indicating with what percentage a user's progress has increased. This screen is alternated with the percentage of how up-to-date a user is in total.

In the experiment, the participants learned about information security. The content existed out of two topics that have been created and checked by the learning consultants of Bright Alley. For this experiment, we selected 25 learning goals with 65 items in total (either a brain snack or a question). Some questions specifically considered a work/office context, so the student participants were told to imagine themselves working at a company. Examples of questions are (roughly translated from Dutch) "You receive an e-mail that you expect to be a phishing mail. What action should you take?" and "A so-called 'information security incident' concerns the disruption of the availability, confidentiality, and integrity of personal data. True or false?".

**Figure 4**

*Schematic overview of PowerApp content (incl. information security examples)*



### Phishing mails

Participants received a fake phishing e-mail before the first and last session of the experiment as a measure of behavior change. In order to check whether the participants clicked on the links in the phishing mails, the website LinkClickCounter.com (n.d.) was used. For each participant, two unique links were created, one per phishing mail. This link directed to either the Wikipedia page of HTTP 404

(Wikipedia contributors, 2022) for the first phishing mail, or to the Error 404 page of the website Magnt.com (“404”, n.d.) for the second mail. Each personal link was added to the phishing mail meant for the corresponding participant. LinkClickCounter.com (n.d.) counted the number of times that the link was clicked upon and displayed this on our account on their website next to each unique link.

## Measures

### Level 1: Reaction

#### *Appreciation survey (subjective)*

The participants’ subjective appreciation of the learning tool was measured by the use of a scale on participants’ reactions on training delivery by Alsalamah and Callinan (2021). The questionnaire consisted of a scale of ten 5-point Likert scale items, ranging from strongly disagree (=1) to strongly agree (=5). In order to ensure that the questionnaire fit the context of the study, we made some changes to the items in the scale (Appendix B – Table B1). The words ‘training’ and ‘training program’ were changed to ‘learning session’ or ‘learning tool’, since the PowerApp was referred to as a learning tool to the participants. The word ‘future’ (job) was added to items 2 and 9, because of the many students who participated and are not yet working (full-time). Finally, item 8 was removed from the scale, as there were no handouts provided in the experiment. This resulted in a reliable scale ( $\alpha = .77$ ).

#### *PowerApp statistics (objective)*

To objectively measure the engagement aspect of the Reaction level, we have looked into the duration of participants’ learning sessions in the PowerApp. This statistic provided an indication of whether participants indeed learned for the amount of time that they were instructed to, or quit earlier. We argued that the longer the duration of a session, the more engaged a participant was. After data collection, the participants belonged to one of three engagement categories (coded as ‘0’, ‘1’, or ‘2’), based on the duration of their learning session(s). Participants in the control condition were assigned to category ‘0’ if the duration was shorter than the instructed 45 minutes minus one SD. They were added to category ‘1’ if it was between 45 minutes minus one SD and 43 minutes and finally to category ‘2’ for all durations longer than 43 minutes. For the microlearning group, the average duration was calculated from the five learning sessions. The categories were assigned similarly as for the control condition, but with the instructed ten minutes instead of 45 and nine minutes instead of 43. Participants in the control group who had reached a score of 100% in their session were automatically assigned to category ‘2’.

### Level 2: Learning

#### *Knowledge tests (objective)*

The subjective measure of the Learning level were knowledge tests on the information security content the participants learned about in the PowerApp. The participants filled in a pre-knowledge test before the start of their (first) learning session. Then, after their (last) learning session, they filled in the first post-knowledge test. Finally, ten days later, they completed a second post-knowledge test. The reason for using a pre-knowledge test was to get data on participants’ pre-knowledge of information security before learning with the PowerApp, as a baseline. The goal of the first post-

knowledge test was to measure to what extent the participants had learned from their learning session(s). The final test was used to measure knowledge retention after ten days of no interaction with the content in the PowerApp.

The questions in the knowledge tests were items from existing security content in the PowerApp. Each learning goal existed out of five items, either ‘brain snacks’ or questions. From each learning goal, one item was used in the pre-knowledge test and one in the post-knowledge tests. We chose to differentiate between the pre- and post-knowledge tests, but the two post-knowledge tests were identical. The reason for this was that the control group would see the pre-knowledge and first post-test in one session of approximately 75 minutes and could possibly remember the questions and answers, instead of understanding the learned content. Since the knowledge tests were relatively long (25 questions each), it was decided to remove some questions whose content marginally overlapped with each other. Finally, all three tests consisted of sixteen questions.

#### *Attitude survey (subjective)*

The participants’ subjective attitude towards the learning was measured by the use of another scale by Alsalamah and Callinan (2021). For the Learning level of an adapted version of Kirkpatrick’s model, they created a scale on the trainees’ views on the impact of the training on their knowledge and learning. The questionnaire consisted of a scale of seven 5-point Likert scale items, ranging from strongly disagree (=1) to strongly agree (=5). In order to ensure that the questionnaire fit the context of our study, we made some minor changes to the items in the scale (Appendix B – Table B2). The words ‘training’ and ‘training program’ were changed to ‘learning/interaction with the learning tool’ or ‘learning tool’, since the PowerApp was referred to as a learning tool to the participants. The word ‘participants’ was changed to ‘(future) colleagues’ in item 4 and the word ‘future’ was added to item 5. No items on the scale were removed and we reached a reliable scale ( $\alpha = .85$ ).

### **Level 3: Behavior**

#### *Behavior surveys (subjective)*

The participants’ own perception of whether their (information security) behavior had changed after learning with the PowerApp was measured by two scales. The first was a scale by Alsalamah and Callinan (2021) on information security behavior. For the Behavior level of an adapted version of Kirkpatrick’s model, they created a scale on trainees’ perspectives on the impact of the training on their behavior. The questionnaire consisted of a scale of seven 5-point Likert scale items, ranging from strongly disagree (=1) to strongly agree (=5). In order to ensure that the questionnaire fit the context of our study, we made some changes to the items in the scale (Appendix B – Table B3). The words ‘(training) program(s)’ were changed to ‘learning’ or ‘learning tool’, since the PowerApp was referred to as a learning tool to the participants. The word ‘future’ was added to items 3 and 6. In item 4 and 7, ‘leadership’, respectively ‘job’, was changed to ‘information security’, and similarly, the words ‘as head teacher’ were removed from item 6 in order to match the context of the current study. Finally, the first item was removed, since we argued that it was too context-specific concerning headteachers. The final scale turned out to be reliable ( $\alpha = .89$ ).

The second behavior scale used to measure level 3 of Kirkpatrick's model concerned behavior change and originated from Egelman and Peer (2015). We used a 13-item 5-point Likert scale with three sub-categories: device securement, password generation, and proactive awareness. The items matched the PowerApp content enough to be added to the behavior survey and were not altered, which lead to a reliable scale ( $\alpha = .79$ , Appendix B – Table B4).

### *Fake phishing mails (objective)*

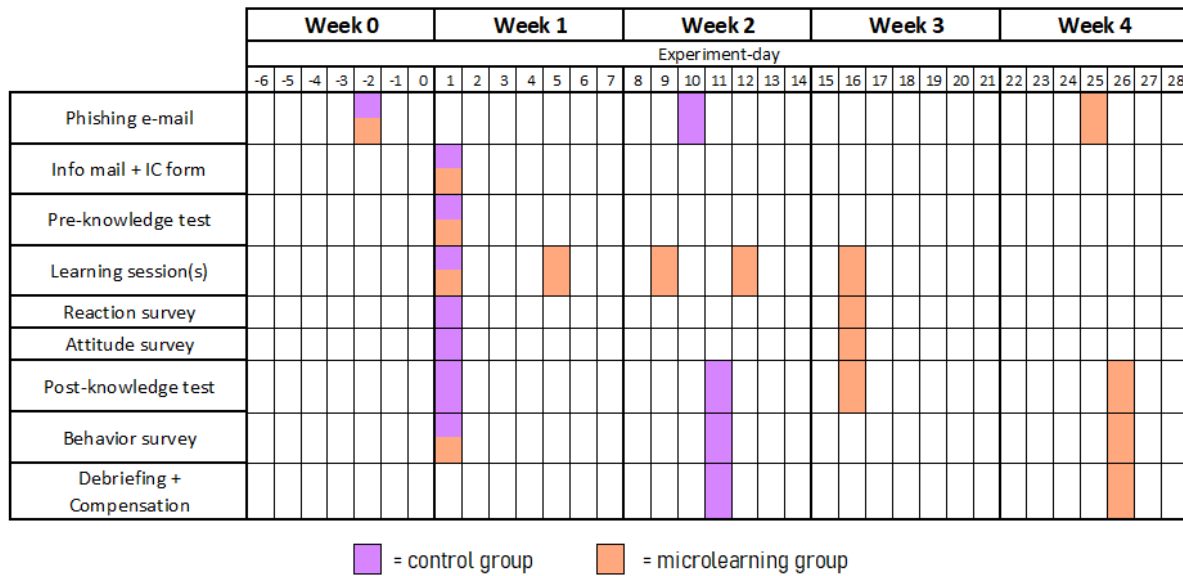
We wanted to objectively measure whether participants had not only learned how to recognize phishing e-mails and how to act on them, but were also able to apply this new knowledge in practice. Therefore, we sent fake phishing mails to the participants, the day before they received the first e-mail about the start of the experiment and another phishing mail the day before the final session. The participants were not told that the phishing mails were part of the experiment, nor were these phishing mails mentioned at any moment until the final debriefing. The first fake phishing mail was used as a baseline for the participants' behavior.

The goal was to create a phishing mail that was realistic, in order to tempt the participants to believe they were real and click on the link in the mail. Additionally, it had to be ensured that the participants would not block the sender's e-mail address from the first e-mail and therefore not receive the second one. Finally, the participants could be divided into two categories – employees of Conclusion b.v. and students. Because of these circumstances, there were four types of e-mails created (Appendix C – Figures C1-C4) and four new e-mail addresses. The first e-mail's subject was 'A new sign-in on Windows', stating that someone had logged into their account and asking to click upon a link in the case that they had not logged in themselves. The subject of the second e-mail was 'Urgent: mandatory password recovery'. The content stated that there had been a suspicious hacking attempt and the participant had to renew their password for the safety of their account. Both e-mails were created in a layout for both the company and the university, in order to be believable to each target group.

The e-mail addresses from which the e-mails were sent were created in Gmail, specifically for this experiment. The addresses were chosen in such a manner that they could potentially be from the company or university of the participants. However, the e-mail addresses all ended with gmail.com, which was one of the red flags that the participants had to learn about in the PowerApp - to notice a sender's suspicious e-mail address. A link click counter was used to track whether participants clicked upon their unique link in the fake phishing mails.

### *Procedure*

The procedure of the experiment was similar for both conditions. The only difference was the number and length of the learning sessions. The control group had one learning session of 45 minutes and one come-back session after ten days. The microlearning group, on the other hand, had five learning sessions of ten minutes and also a come-back session after ten days. Table 4 provides a schematic overview of the experimental procedure.

**Table 4***Overview experimental procedure*

### Session 1

Three days before the first session, the participants received the first phishing mail. On the morning of the first session, the participants of both groups received an extensive e-mail with the instructions. The fact that they were in one of the two conditions and the existence of the other condition was not mentioned. Each participant received their own unique participant number that they were asked to provide in the first question of each survey. The first session consisted of the following steps:

- 1) The participants were asked to read the informed consent form. After signing this form, they completed the pre-knowledge test. Lastly, they completed a scale on behavior change.
- 2) Simultaneously with the first mail, the participants received an invitation to create a PowerApp account with their e-mail address and participant number. The participants then started to interact with the PowerApp. The control group was instructed to stop when they had achieved a score of 100% or after 45 minutes had passed. The microlearning group was told to stop when they had received an 'up-to-date'-notification (Figure 5) or after ten minutes had passed. The timing of this notification was based on the aforementioned theory of Ebbinghaus (1885).
- 3) For the microlearning group, these two steps concluded the first session. The control group had a third step to complete. After the learning session, these participants completed the reaction and attitude survey. Subsequently, they completed the first post-knowledge test.



**Figure 5**

*PowerApp notification of being up-to-date*



Note. From Bright Alley (2020).

#### *Session 2-5 (microlearning group only)*

Each morning of sessions two, three, four, and five for the microlearning group, these participants received a short e-mail that it was time to learn with the PowerApp again, for ten minutes. After learning, they did not have to fill in any tests or surveys. After the fifth learning session, the participants of the microlearning group were asked to complete the reaction and attitude survey. Subsequently, they completed the first post-knowledge test.

#### *Come-back session*

The day before the come-back session, the participants received the second phishing mail. On the morning of the come-back session, the participants of both groups received another e-mail with instructions. It was highlighted that the experimenters relied on them completing this final session, in order to have useful data. The session consisted of two steps:

- 1) The participants completed the second post-knowledge test and the two behavior surveys. Additionally, some demographic questions were asked.
- 2) Then, the participants were debriefed. The goal of the experiment was explained, some theoretical background was provided, the two conditions were mentioned, and the phishing mails were finally connected to the experiment. A few questions followed, asking the participants whether they had suspected that the phishing mails were part of the experiment and whether they thought that they had clicked upon the link(s). In the end, the participants could enter anything they wanted to share, in an open question. The participants entered their bank account information for the compensation and were thanked for their participation. After the compensation had been transferred, the bank account information was deleted immediately.



received a reminder e-mail. The participants were urgently requested to complete the sessions on the specific day that they received an e-mail, but were given the freedom to decide at precisely what time during that day.

### Data Analysis

The effects of microlearning on appreciation, attitude, and behavior change were investigated by means of multiple regression models. Due to the fact that the participants completed multiple knowledge tests, filled in the information security behavior scale twice, and received two fake phishing mails, there were repeated measures per student. Therefore, multi-level regression models were run to determine the effects of microlearning on knowledge, information security behavior, and phishing click behavior. If the assumptions of either regression model were violated, a robust regression was run. If no evidence was found for clustering on the student level, either a multiple regression or chi-square test of independence was run, depending on the type of data. The effect of microlearning on engagement was investigated using a Fisher's Exact test. Demographic variables (i.e. age, gender, and working status) were added to the regressions for exploratory reasons. All statistical analyses were run with Stata 17.0 software (2021).

## Results

In this study, the effect of microlearning on the evaluation of an Intelligent Tutoring System was divided into three facets, following the first three levels of Kirkpatrick's Evaluation Model (2006). These facets were (a) appreciation of and engagement with the learning tool (level 1: Reaction), (b) achievement of learning goals and attitude (level 2: Learning), and (c) behavior changes (level 3: Behavior). The effect of microlearning on each level of evaluation was analyzed in that order.

### Level 1: Reaction

**Table 5**

*Descriptive statistics of appreciation and engagement, per condition*

Variable	Scale	Control			Microlearning			Total		
		M	SD	N	M	SD	N	M	SD	N
Appreciation	(1 - 5)	3.38	.63	22	3.74	.38	21	3.55	.55	43
Engagement	Not engaged			1			1			2
	Slightly engaged			4			2			6
	Engaged			17			18			35

#### *Subjective measure: appreciation*

On average, the participants reported a moderate to good appreciation of the interaction with the PowerApp ( $M = 3.55$ , Table 5). To investigate whether microlearning predicts the participants' appreciation of the learning tool, a linear regression was conducted. The assumptions of homoscedasticity and normally distributed errors were violated, thus a robust regression was run (M1, table 8). A significant positive association was found,  $F(1, 41) = 4.94$ ,  $\beta = .33$ ,  $p = .03$ , with an  $R^2$  of .11, suggesting that participants in the microlearning condition appreciated the learning tool more than the participants in the control condition. On a scale of 1-5, they reported a slightly higher appreciation of .36 than the control group.

To test whether other potentially relevant variables were significant predictors of knowledge, these were added to the regression (M2, table 8). More (pre-)knowledge and a higher attitude were expected to positively affect how enjoyable participants found the learning, and thus lead to a more appreciation. A significant regression equation was found,  $F(6,36) = 11.36$ ,  $p < .01$ , with an  $R^2$  of .65, which indicates that a strong relation exists between appreciation and the independent variables. Appreciation can be predicted absolutely well, based on these variables. However, microlearning was no longer a significant predictor of appreciation ( $p = .10$ ), whilst attitude ( $\beta = .76$ ,  $p < .01$ ) and being male ( $\beta = .33$ ,  $p < .01$ ) showed significant associations. This means that both a more positive attitude towards putting the learned knowledge into practice and being male predict a higher appreciation of interaction with the learning tool, but microlearning does not. Thus, hypothesis H1A was not supported.

### Objective measure: engagement

Two participants were considered not engaged, six were slightly engaged, and 35 participants were categorized as engaged (Table 5). The average learning time of the control condition was approximately 43 minutes, compared to 11 minutes for the microlearning condition. This indicates that a large majority was engaged and interacted with the learning tool for the instructed time period or longer.

The hypothesis was that microlearning would lead to higher engagement in participants. In order to test this, a Fisher's exact test was conducted. No significant relation was found ( $p = .83$ ), indicating that there was no difference in engagement between participants in the microlearning condition, compared to the control condition. Thus, hypothesis H1B was not supported.

## Level 2: Learning

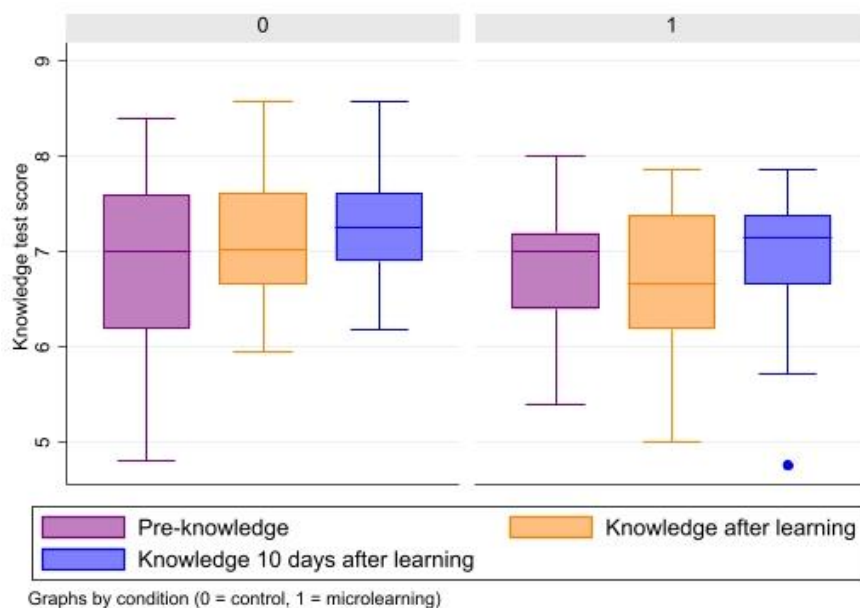
**Table 6**

*Descriptive statistics of knowledge and attitude, per condition*

Variable		Scale	Control			Microlearning			Total		
			M	SD	N	M	SD	N	M	SD	N
Knowledge	Pre-Knowledge	(0 - 10)	6.91	.95	22	6.80	.72	25	6.85	.83	47
	Immediately after learning	(0 - 10)	7.19	.74	22	6.69	.86	21	6.94	.83	43
	10 days after learning	(0 - 10)	7.23	.59	22	6.90	.77	21	7.07	.70	43
Attitude		(1 - 5)	3.06	.75	22	3.31	.58	21	3.18	.67	43

**Figure 6**

*The three knowledge test scores per condition*



*Objective measure: knowledge*

Independent of the experimental condition, participants scored an average of 6.85 on the pre-knowledge test, which implies that most participants were already relatively knowledgeable regarding information security. Ten days after the learning had taken place, this average score had slightly increased to 7.07. This shows that there was, on average, little knowledge gained during the experiment (Table 6, Figure 6).

An empty multilevel regression model for predicting knowledge, clustered at the participants, showed that 36% of the variance was accounted for by the participant level ( $p < .001$ ). This means that a random effect of participant was found. Therefore, the multilevel regression was run again, with microlearning included as a fixed effect (M1, table 9). Since the assumption of independent errors was violated, a robust regression was run. A negative effect of microlearning on knowledge was found, but this was not significant ( $p = .06$ ). Thus, no evidence was found that microlearning leads to more knowledge (i.e. higher knowledge test scores) and hypothesis H2A was not supported.

To test whether other potentially relevant variables were significant predictors of knowledge, these were added to the multilevel regression as random effects (M2, table 9). The moment in the experiment at which the knowledge test was completed was added with the expectation that the knowledge test scores would increase after the learning, compared to the pre-knowledge. No significant regression was found ( $p = .10$ ) and no variables showed to be significant predictors of knowledge. This means that we were not able to predict knowledge, based on the gathered data.

*Subjective measure: attitude*

The average self-reported attitude score was 3.18 (out of five). This indicates a moderately positive attitude towards the interaction with the learning tool and implementing the gained knowledge in practice, regardless of the experimental condition (Table 6). To investigate the effect of microlearning on attitude, a linear regression was run. The assumption of normally distributed errors was violated, thus a robust regression was run (M1, table 8). No significant association was found,  $F(1,41) = 1.39$ ,  $p = .25$ , with an  $R^2$  of .03, indicating that the participants in the microlearning condition did not show a more positive attitude after learning than the people in the control condition.

Other potentially relevant predictors of knowledge were added to the regression (M2, table 8). More (pre-)knowledge and more appreciation and a more positive attitude could again positively relate to how enjoyable participants found the learning, and thus lead to a higher attitude toward applying the knowledge in practice. A significant regression equation was found,  $F(6,36) = 10.91$ ,  $p < .01$ , with an  $R^2$  of .65. The latter value indicates that a strong relation exists between attitude and the independent variables. Attitude can be predicted absolutely well, based on these variables. Microlearning was still not a significant predictor of attitude ( $p = .63$ ), whilst appreciation ( $\beta = .77$ ,  $p < .01$ ) and being male ( $\beta = -.37$ ,  $p < .01$ ) showed significant associations. This suggests that both more appreciation of the learning tool and being female leads to a more positive attitude towards putting the learned knowledge into practice. However, no evidence was found for an effect of microlearning on attitude. Therefore, hypothesis H2B was also not supported.

### Level 3: Behavior

**Table 7**

*Descriptive statistics of current behavior, behavior change, and phishing link clicks, per condition*

Variable		Scale	Control			Microlearning			Total		
			M	SD	N	M	SD	N	M	SD	N
Current behavior	Before learning	(1 - 5)	3.63	.50	22	3.45	.48	25	3.53	.49	47
	10 days after learning	(1 - 5)	3.33	.49	22	3.32	.49	21	3.32	.49	43
Behavior change		(1 - 5)	3.12	.80	22	3.48	.39	21	3.29	.65	43
Phishing link clicks	Before learning				11			7			18
	9 days after learning				3			3			6

#### *Subjective measure: behavior (changes)*

##### Information security behavior

Before interacting with the PowerApp, participants in both conditions rated their information security (IS) behavior with 3.53 out of five on average. Ten days after the learning took place, this average had decreased to 3.32. Thus, both before and after learning, participants reported moderate to good information security behavior, independent of their experimental condition (Table 7).

The empty multi-level regression model for predicting information security behavior, with a random intercept for participant, showed that as much as 71% of the variance was accounted for by the participant level ( $p < .001$ ). This means that a random effect of participant was found and that mostly differences between participants (i.e. personal characteristics) cause differences in IS behavior. Thus, the multilevel regression was run with microlearning as a fixed effect (M1, table 9). No significant association between microlearning and IS behavior was found ( $p = .54$ ). Hence, there was no evidence that microlearning leads to more self-reported IS behavior and thus hypothesis H3A was not supported for information security behavior.

To test whether other potentially relevant variables were significant predictors of information security behavior, these were added to the multilevel regression as random factors (M2, table 9). The reason to add knowledge was that more gained knowledge could lead to more awareness, which would lead to more IS behavior. Similarly, time was added with the expectation that the self-reported IS behavior would increase after the learning, compared to before. A significant regression was found ( $p < .01$ ). Microlearning was still not a significant predictor ( $p = .90$ ), whilst being male was ( $b = .25$ ,  $SE = .11$ ,  $p = .03$ ). This implies that being male leads to more (subjective) IS behavior. Secondly, being in the second ( $b = .63$ ,  $SE = .20$ ,  $p < .01$ ), third ( $b = 1.26$ ,  $SE = .41$ ,  $p < .01$ ), or fourth age group ( $b = .61$ ,  $SE = .27$ ,  $p = .03$ ) was a significant predictor. It seems that the youngest group of participants reported less change in IS behavior than the three older age groups. Finally, completing the behavior scale ten days after the learning ( $b = -.22$ ,  $SE = .05$ ,  $p < .01$ ) showed to be significant. Thus, less information security behavior was reported after ten days of interacting with the PowerApp, compared to before this interaction.

### Behavior changes after interaction with the learning tool

Participants' subjective behavior change as a result of interacting with the PowerApp was on average 3.29 out of five. This implies moderate behavior change, regardless of the experimental condition. To investigate whether microlearning predicts participants' subjective change in behavior, a linear regression analysis on behavior change was run, with microlearning as independent variable. The assumptions of homoscedasticity and normally distributed errors were violated, thus a robust regression was run (M1, table 8). A small, positive effect of microlearning on behavior change was found, but it was not significant,  $F(1,41) = 3.37$ ,  $p = .07$ , with an  $R^2$  of .08. Hence, no evidence was found that microlearning leads to an increase in subjective behavior change and thus hypothesis H3A was not supported for behavior change.

To test whether other potentially relevant variables were significant predictors of behavior change, these were added to the regression (M2, table 8). Appreciation and attitude were included because it was expected that an increase in these variables would positively affect whether participants would implement the gained knowledge in practice and thus change their behavior. A significant regression equation was found,  $F(7,35) = 15.42$ ,  $p < .01$ , with an  $R^2$  of .76. The  $R^2$  indicates that a strong relation exists between behavior change and the independent variables. Behavior change can be predicted absolutely well, based on these variables. Microlearning was again not a significant predictor of attitude ( $p = .58$ ), whilst appreciation ( $\beta = .39$ ,  $p = .03$ ) and attitude ( $\beta = .55$ ,  $p < .01$ ) showed significant associations. This indicates that both more appreciation of the interaction with the learning tool and a more positive attitude toward putting the learned knowledge into practice led to an increase in subjective behavior changes.

### Objective measure: phishing link clicks

Before the learning took place, eighteen participants clicked on the link in the first fake phishing mail. Afterward, only six participants clicked on the link in the second mail (Table 7). To check whether the phishing mail manipulation performed as expected, a chi-square test of independence was conducted on the relation between time and link clicks. This showed a significant effect ( $p = .01$ ). Together with the fact that there were more clicks before the learning than afterward (independent of the experimental condition), it could be concluded that the learning interaction indeed resulted in fewer phishing link clicks, regardless of condition, as was expected.

The empty multilevel regression model for predicting clicks, with a random intercept for participant, showed that 0% of the variance was accounted for by the participant level ( $p = 1.00$ ). This means that no evidence for a random effect of participant was found. Due to the fact that phishing link clicks and microlearning were both binary variables, a chi-square test of independence was run on this relation. No significant relation was found,  $\chi^2(1, N = 47) = 1.17$ ,  $p = .28$ , indicating that microlearning did not influence the total amount of phishing link clicks.

Finally, the hypothesis was that participants in the microlearning condition who clicked on the link in the first phishing mail would less often click on the link in the second mail, compared to participants in the control condition. In other words: they would have gained better knowledge and understanding of information security because of the microlearning aspect and thus show more behavior change (i.e. recognizing the mail as phishing and therefore not clicking anymore). The data tells that fourteen participants clicked on the first link, but not on the second, and thus showed the

expected behavior change. 25 participants did not click at all or clicked on both links, and four participants did not click on the first, but did click on the second. An independent *T*-test of the relation between this behavior change and condition showed no significant result ( $p = .15$ ). Therefore, no evidence was found that microlearning had an effect on changes in phishing link click behavior and thus hypothesis H3B was not supported.

**Table 8**

*Multiple regression models on appreciation, attitude, and behavior change*

		Appreciation		Attitude		Behavior change	
		M1	M2	M1	M2	M1	M2
Microlearning		.36 (.16)*	.20 (.12)	.24 (.20)	-.07 (.15)	.36 (.19)	.07 (.12)
Appreciation					.95 (.13)***		.39 (.17)*
Attitude			.62 (.09)***				.55 (.13)***
Knowledge	Pre-knowledge		-.06 (.07)		.04 (.09)		-.06 (.07)
	Immediately after learning		-.05 (.08)		.13 (.09)		
	10 days after learning						-.00 (.08)
Male			.37 (.12)**		-.51 (.14)**		.12 (.13)
Employee			.06 (.08)		-.05 (.10)		.08 (.08)
<i>N</i>		43	43	43	43	43	43
<i>R</i> <sup>2</sup>		.11	.65	.03	.65	.08	.76

Note. Coefficients with SE in parentheses. \* $p < .05$ , \*\* $p < .01$ , \*\*\* $p < .001$

**Table 9**

*Multi-level models on knowledge and behavior*

		Knowledge		Behavior	
		M1	M2	M1	M2
Microlearning		-.32 (.17)	-.25 (.17)	-.09 (.14)	-.02 (.10)
Male			.15 (.17)		.25 (.11)*
Employee			-.12 (.39)		-.09 (.21)
Age	17-25 yrs				
	25-35 yrs		.36 (.36)		.63 (.20)**
	35-45 yrs		.25 (.41)		1.27 (.41)**
	45-60 yrs		.64 (.53)		.63 (.27)*
Moment	Before learning				
	Immediately after learning		.06 (.16)		
	10 days after learning		.18 (.16)		-.21 (.05)***
<i>N</i> - observations		133	129	90	86
<i>N</i> - groups		47	43	47	43

Note. Coefficients with SE in parentheses. \* $p < .05$ , \*\* $p < .01$ , \*\*\* $p < .001$

## Discussion

The aim of this study was to investigate the difference in evaluation between a learning tool with and without incorporated microlearning. This section discusses the results of the current study, explains its limitations, and provides implications and recommendations for future research.

### Main effects of microlearning on evaluation

#### *Appreciation and Engagement*

A significant effect of microlearning on participants' appreciation of the learning tool was found. People who had been learning in five short microlearning sessions were more appreciative than the ones who had learned in one, longer session. This is in line with the expectations of this study. Even though there is little research on the comparison of appreciation between microlearning and more traditional learning methods, previous studies do indicate that microlearning elicits positive reactions from most learners (Dixit et al., 2020; Tolstikh et al., 2021; Hegerius et al., 2020).

No evidence was found for the hypothesis that participants in the microlearning condition were more engaged than the participants in the control condition. This was presumably partly due to highly skewed data. Engagement was measured by taking the duration of participants' learning sessions and categorizing these. For sessions that were shorter than the instructed time, participants received a low engagement score (0 for truly shorter, 1 for slightly shorter). Sessions as long as instructed or longer, received a score of 2. Overall, the participants scored high on engagement, indicating that they generally followed the instructions on the interaction with the learning tool provided by the experimenter. This was favorable for the study, because it led to the desired difference in duration of the learning sessions between the two conditions, which were either several short sessions or one longer session. A possible explanation for the high engagement is that participants felt motivated to do well on the post-knowledge tests and therefore wanted to interact with the learning tool for the instructed amount of time, or even longer. Another reason could be that participants wanted to do well on the experiment in the interest of the researcher or research in general, i.e. participant bias (Dell et al., 2012; Vashistha et al., 2018). In the current research, engagement was measured by the duration of participants' learning session(s). However, engagement is likely more complex. If possible, it could be favorable to gather more user data, such as number of clicks or mouse tracking. Therefore, an addition could be to let participants self-report on their engagement, besides objectively measuring engagement.

#### *Knowledge and Attitude*

Contradicting our expectations and earlier research, microlearning showed not to be a significant predictor of knowledge. The found result could be explained by the fact that participants in both conditions were already knowledgeable regarding the topic, as seen by the relatively high average score of the pre-knowledge test. This could mean that they were not able to gain high amounts of new knowledge and significantly improve their test scores, explaining the lack of growth in scores in both experimental conditions. Conclusion B.V. and the TU/e both frequently inform their employees and students about the topic of information security. Also, Conclusion B.V. and the university are technology-oriented organizations, which implies that their employees and students have a



generally high interest in this field. Finally, the questions could have been too easy for participants with technology-oriented professions and/or studies, as they have been created for an average pre-understanding of information security.

The fact that the difference in test scores between the pre-knowledge test and the first post-knowledge test is small also explains the absence of a difference in score between the two post-knowledge tests. Intuitively, if the amount of gained knowledge immediately after the learning is negligible, the knowledge retention after ten days will be negligible as well. Another possible explanation for the lack of significant difference between the two post-knowledge tests is the fact that these tests contained the same questions. The participants were not told whether their answers to the first post-test were correct, but recognition of the questions and remembering their own answers could still have influenced their test score for the final post-knowledge test.

Likewise, no significant effect was found of microlearning on participants' attitude towards implementing the gained knowledge in practice, which contradicts the hypothesis of a positive effect. The contradicting findings of the current study and those in literature could be explained by the similar test scores of the experimental conditions. Intuitively, gaining more knowledge by learning could result in a more positive attitude towards implementing these new skills and pieces of information, regardless of the used tool. Therefore, since the participants in both groups did approximately equally well on the post-knowledge tests, it seems understandable that there was no significant difference in attitude found between the conditions.

### *Information Security Behavior and Behavior Changes*

It was found that microlearning did not have a significant effect on information security behavior, as well as behavior change. However, for behavior change, a slightly positive trend was observed. The insignificant effects contradict the expectations of the current study. The absence of an effect could possibly be due to the relatively short period of time between the (last) learning session and the final questionnaires. Lally et al. (2009) argued that it takes on average 66 days to form a new behavior or habit, while Ronis et al. (1988) stated that behavior should be performed a minimum of twice per month and ten times in total to be considered structural behavior change. In the current study, only ten days had passed when participants were asked to evaluate their own behavior, which could arguably be a too short time frame for (permanent) behavior change to occur.

Additionally, the results show that 71% of the variance in behavior was accounted for by the participant level. This implies that it depends on personal characteristics whether people actually showed changes in subjective information security behavior, rather than their experimental condition. It could be that some people were more prone to taking action, more interested in and committed to information security, or more sensitive to changing habits. Therefore, taking personal characteristics into account could provide more insight into the found differences in behavior changes between participants. It could help determine whether the suitability of microlearning as a learning method is dependent on differences between personality traits.

No evidence was found for an effect of microlearning on phishing link click behavior. A possible explanation for the lack of this effect is that participants could have understood that the first mail was phishing, after clicking on the link. Then, when receiving the second mail, they could have recognized the layout (even though the content was different) and judged the mail for being

phishing without clicking the link. Additionally, the content of the first e-mail was different than the second mail, which was not randomized. Therefore, the e-mails could have differed in how convincing they appeared to be. For example, one participant explained that their organization takes care of password changes automatically, which made the second e-mail with the subject of urgently changing their password less believable to them. Nevertheless, several participants stated that they were unaware that the mails were part of the experiment. Only after being debriefed, they understood the connection. Additionally, eighteen out of 43 participants clicked on the link in the first phishing mail. Together, this could indicate that the phishing mails were perceived as relatively realistic and believable.

## Additional predictors of evaluation

### *Appreciation and attitude*

Attitude appeared to positively affect the participants' appreciation of the learning session(s) and the same effect was found of appreciation on attitude. It seems intuitive that participants who were more motivated to put the gained knowledge into practice were also more appreciative of the learning tool and the other way around. For the scale on self-reported behavior change, appreciation and attitude also showed to be significant predictors. More appreciation of and a more positive attitude towards the learning tool led to more self-reported behavior change. In 1975, Fishbein and Ajzen already described the positive relation between attitude (or intention) and behavior change. In more recent literature, this positive effect is still found (Glasman & Albarracín, 2006; Bechler et al., 2021). Therefore, the found effect is in line with existing research.

### *Gender, age, and working status*

Males showed to be more appreciative of the PowerApp than females, regardless of their experimental condition, but had a less positive attitude towards the learning. Ozturan and Kutlu (2010) used Kirkpatrick's model on e-learning in corporate training and, while they did not report a direction, they found a significant effect of gender on participants' reactions. Other research on appreciation of technological (learning) tools often either not considered the effects of gender or did not report any relevant results. Regarding attitude, Iqbal et al. (2021) found that females were significantly more confident in their increased knowledge and skills as a result of training. Rettger (2017) investigated differences in attitude towards using mobile learning between males and females, but found no statistically significant effects. It can be concluded that an effect of microlearning on appreciation of the learning tool seems to exist. Nevertheless, the effect is small, since part of the variance can be explained by a participant's attitude and gender. Similarly, an effect of gender and appreciation on attitude towards the learning tool exists. Still, the explanations behind these effects remain uncertain. Finally, it was found that males reported significantly higher levels of information security behavior than females, regardless of the phase of the experiment. Literature shows that men tend to be more self-confident regarding knowledge than women (Ross et al., 2012; Harrington et al., 2018), which could explain these higher self-reported behavior scores.

Besides being male, significant predictors of information security behavior were being in one of the three oldest age groups of the experiment and completing the behavior survey ten days after the learning (compared to before the learning). The youngest age group (17-25 years) reported the

least IS behavior of all age groups. The first potential explanation for this found effect is that the data is relatively skewed, since 28 out of 43 participants belonged to the youngest group. Additionally, these participants were mostly students (27 out of 28), while the older age groups consisted more of employees. It seems possible that employees have received more training and information in their work regarding the security topic than the students did at the university and, therefore, feel more confident about their IS behavior. Subsequently, an effect was found of the moment in the experiment on behavior. Participants showed a decrease in IS behavior scores in the final questionnaire, compared to the first one. A reason for this could be that they were relatively overconfident before the start of the experiment. Then, when they learned and during the period after the interaction with the learning tool, they might have become more self-conscious about and aware of their information security behavior.

Knowledge and whether the participant was an employee or a student had no significant effects on any of the evaluated levels. The lack of effect of knowledge could be explained by the insignificant gains in knowledge of the participants. A potential reason why the working status of participants did not influence the measures could be that students were told to imagine that they were employees, while interacting with the PowerApp.

## Limitations

One of the main limitations of the current study was the time frame in which the experiment took place, which was four weeks. This constraint especially affected the evaluation of levels two and three. For knowledge and behavior, it has been shown that several months of (learning) time are likely needed for more knowledge retention and actual, structural changes in behavior. Another aspect that presumably influenced the outcomes was the sample size. As explained in the Method section, the a priori power analysis recommended a sample size of 116 participants for a power of 90%. However, due to time and money constraints and dropouts, only 43 participants completed the entire experiment. A post hoc power analysis showed that the power of the study was only 55%. This means that, theoretically, there was a 55% chance of detecting an effect, if it exists. However, the current study shows that, even for a short learning period and with a suboptimal sample size, there already seems to exist an effect of microlearning on how the learning method is appreciated. Future research could investigate to what extent longer learning time results in effects on all levels of Kirkpatrick's model, preferably with a larger sample size. Also, due to recruitment constraints, the people in the sample of the current study consisted of both employees and students. This was despite the fact that the study focused on microlearning within a work environment. Participants who were students were told that they had to imagine that they were working at an office, while learning with the PowerApp. This was because the information and questions considered an employee perspective, rather than a student perspective. Still, students could possibly not have identified with the illustrated problems and situations, which could have influenced their answers to the questions. Nevertheless, the results did not show differences between employees and students on any evaluation level, so the participants' working status seemed not to have influenced the current study.

Regarding the PowerApp, there were some additional limitations. The application has been designed with the purpose of using it in a microlearning manner. However, the control group in this

experiment learned with the learning tool in one, longer session. An issue that resulted from this was that participants received the same questions multiple times in their learning session, as long as they had not yet achieved a 100% score. In the microlearning condition, questions were also repeated, but this occurred across the several sessions, days apart. Several participants in the control condition reported that they were annoyed by the reoccurring questions, even when they answered them correctly each time. This could have influenced their appreciation of and engagement with the learning tool. Thus, the PowerApp's algorithm was arguably not completely suitable for the experimental set-up for the control group. Finally, due to the high pre-knowledge test scores, it could be that the participants were already reasonably familiar with the topic. This could have affected the extent to which the information was new to them. However, the topic of information security was chosen because it was argued to be relevant to both students and employees. The high knowledge test scores seem to confirm that information security was already of (practical) interest to these groups, before the experiment.

### Implications

It was found that microlearning had a small positive effect on how the learning interaction was appreciated. Therefore, employees could consider using a microlearning method for the training of their employees in order to increase training satisfaction. Future research needs to investigate whether greater learning and more behavior changes can occur as well as a result of microlearning training.

The current study evaluated a learning tool using both subjective and objective measures, compared to most literature that uses only subjective ones. For example, the effect of microlearning on behavior change was not significant, but we saw that people in general thought that they had somewhat changed their behavior as a result of the learning tool interaction (3.29 out of 5). Indeed, we observed that fewer people clicked on the phishing link after the interaction. This shows that subjective measures allow for the understanding of a participant's perspective and how they view the world or themselves. Objective measures aim to investigate whether their perspective is true to what actually happens. Using both measures in research allows researchers to find underlying, subjective, reasons for why certain actions are objectively observed and to what extent participants are aware of these actions.

Microlearning as a learning method could not only be relevant for training in a work environment, but possibly also in schools and other types of education. For short courses of a few weeks, short microlearning sessions are likely not a good fit, due to the limited amount of time for mastering the material. However, for year-long courses or part-time courses for adults, microlearning could aid in the distribution of the workload.

### Future research

The specific field of microlearning is still in its early stages and more research is needed to understand its effects and potential benefits. Based on the current study, we propose a number of recommendations for future research on the evaluation of microlearning.

Microlearning is a type of distributed learning. The latter has been researched more often and has shown to benefit learners in terms of knowledge retention. Even though the current study has found some, small, effects of microlearning, future research should aid in confirming whether the same positive effects of distributed learning in general are found for microlearning specifically as well.

Kirkpatrick (1994) argued that there exist causal relations between the four levels of his model. This implies that positive reactions result in more or better learning, which in its turn leads to greater behavior changes, and thus more desired organizational outcomes. However, a number of meta-analyses on applications of Kirkpatrick's work have found a lack of evidence for these links (Alliger & Janak, 1989; Alliger et al., 1997). We argue that learners could show high amounts of learning, whilst not reacting positively to the learning tool. Or, positive reactions could lead to more behavior change, without greater learning. Therefore, when using Kirkpatrick's evaluation model, it could be favorable to evaluate a learning tool on all four levels, in order to generate the most complete evaluation of the tool and the relations between the levels. Additionally, studies should preferably be conducted over longer periods of time (e.g. a year) in order to achieve and measure structural changes on all four levels.

It could be hypothesized that a microlearning tool is evaluated differently when participants can determine when they want to learn and when they have time for it. In the current study, participants in the microlearning condition were told on what day they were expected to learn. Perhaps, the outcome of the evaluations would have been different if the participants were allowed to decide when they wanted to learn themselves. It can differ between persons at what moments or days of the week they tend to be most productive or focused. For future studies, it could be interesting to take these personal preferences into account and investigate the impact of allowing participants to decide when they want to learn on how the learning tool is evaluated. In the case that large differences are found, companies could adapt the flexibility or sturdiness of their training schedules.

Finally, the current study used an Intelligent Tutoring System as a learning tool for the experiment. However, both experimental conditions interacted with the ITS. This means that the influence of learning with an ITS was not investigated. For future research, it could be relevant to research this influence and additionally the interaction effect of an ITS and microlearning. This could for example be realized by having four experimental conditions: a control group that interacts with a non-ITS learning tool without microlearning, a non-ITS group with microlearning, an ITS group without microlearning, and a group that interacts with an ITS with microlearning incorporated.

## Conclusion

Work-related training, in order to stay up-to-date and relevant, continuously takes up valuable time from busy employees. We have theorized that microlearning can facilitate as a solution to lighten workload and improve training appreciation, engagement, knowledge retention, attitude, and desired behavior changes. The current study aimed to add several new aspects to existing research on microlearning as a method and the evaluation of microlearning tools. The first aspect is that an extensive literature review was conducted on the methods that existing studies on learning tools in general (i.e. not specifically microlearning tools) used to measure each level of Kirkpatrick's Evaluation Model. The second aspect is that, in contrast to the majority of studies on microlearning, we used three levels of Kirkpatrick's Model for evaluation of this learning method. The third aspect is the usage of objective measures. The literature review showed that, regarding Kirkpatrick's model, most often only subjective measures are used for evaluation at each level, especially levels one and three. The current study had an objective measure, besides subjective ones, for all three levels: objective engagement in the learning tool for level one, knowledge tests for level two, and the phishing e-mail clicks for level three. Another new aspect was the incorporation of microlearning into an Intelligent Tutoring System, instead of a learning tool without personalized, instant feedback. Finally, we used a control group in our experiment, in contrast to many existing studies on microlearning.

The current study investigated the difference in evaluation between learning with incorporated microlearning and learning without, in a work environment. The results showed a small effect of microlearning on the subjective appreciation of the training. The other expected results of greater learning and behavior changes were not found. This outcome can probably be partly explained due to limitations in time frame, sample size, and the used learning tool. Nevertheless, as long as employees' schedules remain congested, there is a need for a learning method that is able to improve training on multiple levels. Future research should determine whether microlearning is indeed the best candidate for this job.

## Bibliography

404. (n.d.). Retrieved from <https://www.magnt.com/404>.

Alliger, G.M., & Janak, E.A. (1989). Kirkpatrick's levels of training criteria: thirty years later. *Personnel Psychology*, 42(2), 331–342. <https://doi.org/10.1111/j.1744-6570.1989.tb00661>

Alliger, G.M., Tannenbaum, S.I., Bennett, W., Traver, H., & Shotland, A. (1997). A meta-analysis of the relations among training criteria. *Personnel Psychology*, 50(2), 341–358. <https://doi.org/10.1111/j.1744-6570.1997.tb00911>

Alsalamah, A., & Callinan, C. (2021). Adaptation of Kirkpatrick's Four-Level Model of Training Criteria to Evaluate Training Programmes for Head Teachers. *Education Sciences*, 11(3), 116. <https://doi.org/10.3390/educsci11030116>

Amokrane, K., & Lourdeaux, D. (2009, July 13-16). *Virtual reality contribution to training and risk prevention* [Conference]. International Conference on Artificial Intelligence, Las Vegas, Nevada.

Andriana, I., Riyanto, D., & Darmawan, D. (2019, November). *Workload and motivation on employees performance analyzed by information technology* [Conference]. IOP Conference Series: Materials Science and Engineering, 662(2) doi:10.1088/1757-899X/662/2/022120.

Bates, R. (2004). A critical analysis of evaluation practice: the Kirkpatrick model and the principle of beneficence. *Evaluation and Program Planning*, 27(3), 341–347. <https://doi.org/10.1016/j.evalprogplan.2004.04.011>

Bechler, C. J., Tormala, Z. L., & Rucker, D. D. (2021). The Attitude–Behavior Relationship Revisited. *Psychological Science*, 32(8), 1285–1297. <https://doi.org/10.1177/0956797621995206>

Boring, C. (2020). *Microlearning: An andragogical case study on knowledge retention, learner satisfaction, and behavior change* (Order No. 27744159). Available from Publicly Available Content Database. (2387259958). Retrieved from <https://www.proquest.com/dissertations-theses/microlearning-andragogical-case-study-on/docview/2387259958/se-2>

Chai-Arayalert, S., & Puttinaovaratt, S. (2020). Designing mangrove ecology self-learning application based on a micro-learning approach. *International Journal of Emerging Technologies in Learning*, 15(11), 29-41. Doi:10.3991/IJET.V15I11.12585

Chun, B. A., & Heo, H. J. (2018, January 6-8). *The effect of flipped learning on academic performance as an innovative method for overcoming ebbinghaus' forgetting curve* [Conference]. 6th



- International Conference on Information and Education Technology, Osaka, Japan.  
<https://doi.org/10.1145/3178158.3178206>
- Dell, N., Vaidyanathan, V., Medhi, I., Cutrell, E., & Thies, W. (2012, May). “Yours is better!” [Conference]. SIGCHI Conference on Human Factors in Computing Systems, Austin, Texas, U.S.A. <https://doi.org/10.1145/2207676.2208589>
- Dixit, R. K., Yalagi, P. S., & Nirgude, M. A. (2021). Breaking the walls of classroom through Micro learning: Short burst of learning. *Journal of Physics: Conference Series*, 1854(1). <https://doi.org/10.1088/1742-6596/1854/1/012018>
- Ebbinghaus H. (1885). *Über das Gedächtnis*. Leipzig: Dunker.
- Egelman, S., & Peer, E. (2015, April 18-23). *Scaling the Security Wall* [Conference]. ACM Conference on Human Factors in Computing Systems, Seoul, Republic of Korea. <https://doi.org/10.1145/2702123.2702249>
- Emerson, L. C., & Berge, Z. L. (2018). Microlearning: Knowledge management applications and competency-based training in the workplace. *Knowledge Management & E-Learning*, 10(2), 125–132
- Erickson, E., Arroyo, I., & Woolf, B. (2018). Exploring Gritty Students’ Behavior in an Intelligent Tutoring System. *Lecture Notes in Computer Science*, 79–83. [https://doi.org/10.1007/978-3-319-93846-2\\_15](https://doi.org/10.1007/978-3-319-93846-2_15)
- Fishbein, M., & Ajzen, I. (1975). *Belief, Attitude, Intention and Behavior: An Introduction to Theory and Research*. Addison-Wesley.
- Fletcher, J. D. (2003). Evidence for learning from technology-assisted instruction. In H. F. O’Neil, Jr. & R.S. Perez (Ed.), *Technology applications in education: A learning view* (pp. 79–99). Lawrence Erlbaum Associates Publishers.
- Glasman, L. R., & Albarracín, D. (2006). Forming attitudes that predict future behavior: A meta-analysis of the attitude-behavior relation. *Psychological Bulletin*, 132(5), 778–822. <https://doi.org/10.1037/0033-2909.132.5.778>
- Griesser, J. W. (1993). Motivation and Information System Professionals. *Journal of Managerial Psychology*, 8(3), 21–30. <https://doi.org/10.1108/02683949310027772>
- Hamtini, T. M. (2008). Evaluating E-learning Programs: An Adaptation of Kirkpatrick’s Model to Accommodate E-learning Environments. *Journal of Computer Science*, 4(8), 693–698. <https://doi.org/10.3844/jcssp.2008.693.698>



- Harrington, B., Peng, S., Jin, X., & Khan, M. (2018, July 2-4). *Gender, confidence, and mark prediction in CS examinations* [Conference]. ACM Conference on Innovation and Technology in Computer Science Education, Larnaca, Cyprus. <https://doi.org/10.1145/3197091.3197116>
- Hegerius, A., Caduff-Janosa, P., Savage, R., & Ellenius, J. (2020). E-Learning in Pharmacovigilance: An Evaluation of Microlearning-Based Modules Developed by Uppsala Monitoring Centre. *Drug Safety*, 43(11), 1171–1180. <https://doi.org/10.1007/s40264-020-00981-w>
- Hesse, A., Ospina, P., Wieland, M., Yepes, F. A. L., Nguyen, B., & Heuwieser, W. (2019). Short communication: Microlearning courses are effective at increasing the feelings of confidence and accuracy in the work of dairy personnel. *Journal of Dairy Science*, 102(10), 9505–9511. <https://doi.org/10.3168/jds.2018-15927>
- Hooshyar, D., Ahmad, R. B., Yousefi, M., Fathi, M., Horng, S. J., & Lim, H. (2016). SITS: A solution-based intelligent tutoring system for students' acquisition of problem-solving skills in computer programming. *Innovations in Education and Teaching International*, 55(3), 325–335. <https://doi.org/10.1080/14703297.2016.1189346>
- Hug, T., Lindner, M., & Bruck, P. (2006). Microlearning: Emerging concepts, practices and technologies after e-Learning. In *Proceedings of Microlearning Conference 2005: learning & working in new media* (Issue January 2016).
- Iqbal, M. Z., Alaskar, M., Alahmadi, Y., Alhwiesh, B. A., & Mahrous, A. A. (2021). Perceptions of Residents on the Microlearning Environment in Postgraduate Clinical Training. *Education Research International*, 2021, 1–6. <https://doi.org/10.1155/2021/9882120>
- Jahnke, I., Lee, Y. M., Pham, M., He, H., & Austin, L. (2020). Unpacking the Inherent Design Principles of Mobile Microlearning. *Technology, Knowledge and Learning*, 25(3). <https://doi.org/10.1007/s10758-019-09413-w>
- Javorcik, T., & Polasek, R. (2019, November 21-22). *Comparing the Effectiveness of Microlearning and eLearning Courses in the Education of Future Teachers*. [Conference]. International Conference on Emerging Elearning Technologies and Applications (ICETA), Starý Smokovec, Slovakia. <https://doi.org/10.1109/iceta48886.2019.9040034>
- Javorcik, T., & Polasek, R. (2019, November 21-22). *Practical Application of MicroLearning in Education of Future Teachers* [Conference]. International Conference on Emerging Elearning Technologies and Applications (ICETA), Starý Smokovec, Slovakia. <https://doi.org/10.34190/eel.19.031>
- Kirkpatrick D. L, & Kirkpatrick J. D. (2006). *Evaluating training programs: The four levels* (3<sup>rd</sup> ed.). Berrett-Koehler Publication.

- Kirkpatrick, D.L. (1976). Evaluation of Training. In: Craig, R.L., Ed., *Training and Development Handbook: A Guide to Human Resource Development*, McGraw Hill, New York.
- Kirkpatrick, D. L. (1994). *Evaluating Training Programs: The Four Levels*. San Francisco: Berrett-Koehler Publishers.
- Koedinger, K.R., Anderson, J.R., Hadley, W.H., & Mark, M.A. (1997). Intelligent Tutoring Goes To School in the Big City. *International Journal of Artificial Intelligence in Education*, 8, 30-43.
- Lally, P., van Jaarsveld, C. H. M., Potts, H. W. W., & Wardle, J. (2009). How are habits formed: Modelling habit formation in the real world. *European Journal of Social Psychology*, 40(6), 998–1009. <https://doi.org/10.1002/ejsp.674>
- Lau, R. W. H., Yen, N. Y., Li, F., & Wah, B. (2013). Recent development in multimedia e-learning technologies. *World Wide Web*, 17(2), 189–198. <https://doi.org/10.1007/s11280-013-0206-8>
- Link Click Counter - free URL shortener and monitor. (n.d.). Retrieved from <https://www.linkclickcounter.com>
- London, M., & Diamante, T. (2018). The science of learning and development. In M. London & T. Diamante, *Learning interventions for consultants: Building the talent that drives business* (pp. 25–45). American Psychological Association. <https://doi.org/10.1037/0000094-003>
- Mak, B. L., & Sockel, H. (2001). A confirmatory factor analysis of IS employee motivation and retention. *Information & Management*, 38(5), 265–276. doi:10.1016/s0378-7206(00)00055-0
- Mann, S. (1996). What should training evaluations evaluate? *Journal of European Industrial Training*, 20(9), 14–20. <https://doi.org/10.1108/03090599610150264>
- Mayer, K. U., & Solga, H. (2008). *Skill formation: Interdisciplinary and cross-national perspectives*. Cambridge: Cambridge University Press.
- Mazzoni, G., & Cornoldi, C. (1993). Strategies in study time allocation: Why is study time sometimes not effective? *Journal of Experimental Psychology: General*, 122(1), 47–60. <https://doi.org/10.1037/0096-3445.122.1.47>
- Mitrović, A., & Holland, J. (2020). Effect of Non-mandatory Use of an Intelligent Tutoring System on Students' Learning. *Lecture Notes in Computer Science*, 386–397. [https://doi.org/10.1007/978-3-030-52237-7\\_31](https://doi.org/10.1007/978-3-030-52237-7_31)
- Naser, S. S. A. (2008). Developing an intelligent tutoring system for students learning to program in C++. *Information Technology Journal*, 7(7), 1055–1060. doi:10.3923/itj.2008.1055.1060

- Nkambou, R., Bourdeau, J., & Psyché, V. (2010). Building Intelligent Tutoring Systems: An Overview. *Studies in Computational Intelligence*, 361–375. [https://doi.org/10.1007/978-3-642-14363-2\\_18](https://doi.org/10.1007/978-3-642-14363-2_18)
- Olsen, J. K., Rummel, N., & Aleven, V. (2016). Investigating Effects of Embedding Collaboration in an Intelligent Tutoring System for Elementary School Students In Looi, C. K., Polman, J. L., Cress, U., and Reimann, P. (Eds.). *Transforming Learning, Empowering Learners: The International Conference of the Learning Sciences (ICLS) 2016, Volume 1*. Singapore: International Society of the Learning Sciences.
- Orwoll, B., Diane, S., Henry, D., Tsang, L., Chu, K., Meer, C., Hartman, K., & Roy-Burman, A. (2017). Gamification and microlearning for engagement with quality improvement (GAMEQI): A bundled digital intervention for the prevention of central line-associated bloodstream infection. *American Journal of Medical Quality*, 33(1), 21–29.
- Ozturan, M., & Kutlu, B. (2010). Employee satisfaction of corporate e-training programs. *Procedia - Social and Behavioral Sciences*, 2(2), 5561–5565. <https://doi.org/10.1016/j.sbspro.2010.03.907>
- Pane, J. F., Griffin, B. A., McCaffrey, D. F., & Karam, R. (2014). Effectiveness of Cognitive Tutor Algebra I at Scale. *Educational Evaluation and Policy Analysis*, 36(2), 127–144. <https://doi.org/10.3102/0162373713507480>
- Pokorny, R. A., Hall, E. P., Gallaway, M. A., & Dibble, E. (1996). Analyzing components of work samples to evaluate performance. *Military Psychology*, 8(3), 161–177. doi:10.1207/s15327876mp0803\_3
- Puah, S., Bin Mohmad Khalid, M. I. S., Looi, C. K., & Khor, E. T. (2021). Investigating working adults' intentions to participate in microlearning using the decomposed theory of planned behaviour. *British Journal of Educational Technology*. doi:10.1111/bjet.13170
- Ramesh, V. M., & Rao, N. J. (2012, July). *Tutoring and expert modules of intelligent tutoring systems* [Conference]. IEEE 4th International Conference on Technology for Education, Hyderabad, India. doi:10.1109/T4E.2012.52
- Reio, T. G., Rocco, T. S., Smith, D. H., & Chang, E. (2017). A Critique of Kirkpatrick's Evaluation Model. *New Horizons in Adult Education and Human Resource Development*, 29(2), 35–53. <https://doi.org/10.1002/nha3.20178>
- Rettger, E. (2017, March). *Microlearning with Mobile Devices: Effects of Distributed Presentation Learning and the Testing Effect on Mobile Devices*. Arizona State University. [https://keep.lib.asu.edu/\\_flysystem/fedora/c7/178994/Rettger\\_asu\\_0010E\\_16994.pdf](https://keep.lib.asu.edu/_flysystem/fedora/c7/178994/Rettger_asu_0010E_16994.pdf)

- Roll, I., Aleven, V., McLaren, B. M., & Koedinger, K. R. (2011). Metacognitive Practice Makes Perfect: Improving Students' Self-Assessment Skills with an Intelligent Tutoring System. *Lecture Notes in Computer Science*, 288–295. [https://doi.org/10.1007/978-3-642-21869-9\\_38](https://doi.org/10.1007/978-3-642-21869-9_38)
- Ronis, D. L., Yates, J. F., & Kirscht, J. P. (1988). Attitudes, decisions, and habits as determinants of repeated behavior. In A.R. Pratkanis, S. J. Breckler, & A. G. Greenwald (Eds.), *Attitude structure and function* (pp. 213–239). Hillsdale: Lawrence Erlbaum Associates.
- Ross, J. A., Scott, G., & Bruce, C. D. (2012). The Gender Confidence Gap in Fractions Knowledge: Gender Differences in Student Belief-Achievement Relationships. *School Science and Mathematics*, 112(5), 278–288. <https://doi.org/10.1111/j.1949-8594.2012.00144.x>
- Sabitov, R. A., Smirnova, G. S., Sirazetdinov, B. R., Elizarova, N. U., & Eponeshnikov, A. V. (2017). The concept of intelligent tutoring for enterprise staff as a component of integrated manufacturing control system development. *Advances in Systems Science and Applications*, 17(1), 1-8. Retrieved from [www.scopus.com](http://www.scopus.com)
- Sarma, B. H. S., & Ravindran, B. (2007). Intelligent Tutoring Systems using Reinforcement Learning to teach Autistic Students. *Home Informatics and Telematics: ICT for The Next Billion*, 65–78. [https://doi.org/10.1007/978-0-387-73697-6\\_5](https://doi.org/10.1007/978-0-387-73697-6_5)
- Sawarynski, K. E., & Baxa, D. M. (2019). Utilization of an online module bank for a research training curriculum: development, implementation, evolution, evaluation, and lessons learned. *Medical Education Online*, 24(1), 1611297. <https://doi.org/10.1080/10872981.2019.1611297>
- Schmitz, C. (n.d.). LimeSurvey. Retrieved March 2, 2022, from <https://www.limesurvey.org/solutions/universities>
- Shail, M. S. (2019). Using Micro-learning on Mobile Applications to Increase Knowledge Retention and Work Performance: A Review of Literature. *Cureus*. <https://doi.org/10.7759/cureus.5307>
- Sim, J. (2017). Using Kirkpatrick Four Level Evaluation model to assess a 12-week accelerated ultrasound intensive course. *Sonography*, 4(3), 110–119. <https://doi.org/10.1002/sono.12116>
- StataCorp. 2021. Stata Statistical Software: Release 17. College Station, TX: StataCorp LLC.
- Stevens, S. M., Forsythe, J. C., Abbott, R. G., & Gieseler, C. J. (2009). Experimental Assessment of Accuracy of Automated Knowledge Capture. *Foundations of Augmented Cognition. Neuroergonomics and Operational Neuroscience*, 212–216. [https://doi.org/10.1007/978-3-642-02812-0\\_25](https://doi.org/10.1007/978-3-642-02812-0_25)

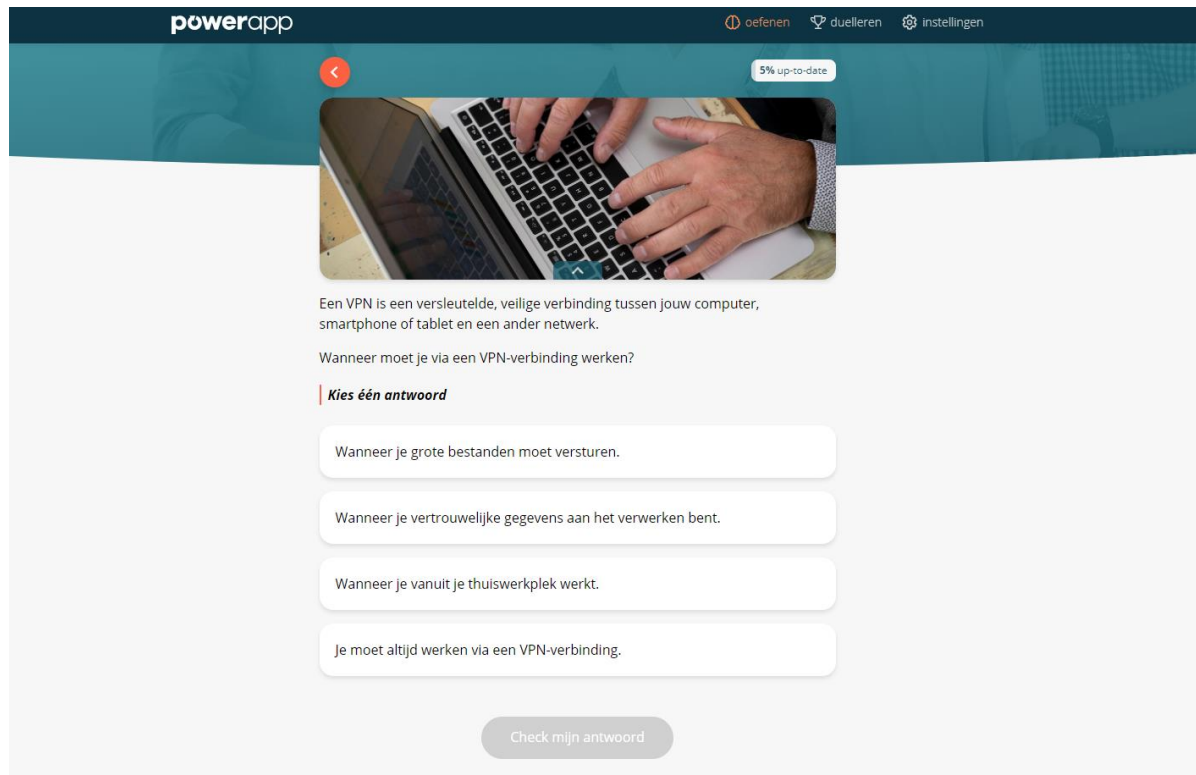
- Sykes, E. R., & Franek, F. (2003, June 30 – July 2). *A prototype for an intelligent tutoring system for students learning to program in java™* [Conference]. IASTED International Conference on Computers and Advanced Technology in Education, Rhodes, Greece, 78-83.
- Tolstikh, O., Pankova, V., & Krasnova, E. (2021). Microlearning in teaching English to students of engineering specialties. *E3S Web of Conferences*, 273, 12136. <https://doi.org/10.1051/e3sconf/202127312136>
- Tullis, J. G., & Benjamin, A. S. (2011). On the effectiveness of self-paced learning. *Journal of Memory and Language*, 64(2), 109–118. <https://doi.org/10.1016/j.jml.2010.11.002>
- Tumi, N. S., Hasan, A. N., & Khalid, J. (2022). Impact of compensation, job enrichment and enlargement, and training on employee motivation. *Business Perspectives and Research*, 10(1), 121-139. doi:10.1177/2278533721995353
- Ukamaka Betrand, C., Okwudili Anigbogu, S., & Adiele Okolie, S. (2021). Overview of Intelligent Tutoring System in Human Teamwork and Its Impact Generally on Learning. *International Journal of Intelligent Information Systems*, 10(1), 5. <https://doi.org/10.11648/j.ijis.20211001.12>
- United States Office of Personnel Management. (2008). *Best practice: Mentoring*. Retrieved from <http://www.opm.gov/hrd/lead/BestPractices-Mentoring.pdf>
- Van Ruysseveldt, J., & Van Dijke, M. (2011). When are workload and workplace learning opportunities related in a curvilinear manner? The moderating role of autonomy. *Journal of Vocational Behavior*, 79(2), 470–483. <https://doi.org/10.1016/j.jvb.2011.03.003>
- Vashistha, A., Okeke, F., Anderson, R., & Dell, N. (2018). 'You Can Always Do Better!'. *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. <https://doi.org/10.1145/3173574.3174126>
- Wikipedia contributors. (2022, January 18). HTTP 404. Retrieved from [https://en.wikipedia.org/wiki/HTTP\\_404](https://en.wikipedia.org/wiki/HTTP_404)
- Zandbergs, U., Judrups, J., Plane, E., & Uscins, R. (2021). Improvement of microlearning with help of learning analytics in enterprises. *20th International Scientific Conference Engineering for Rural Development Proceedings*. Published. <https://doi.org/10.22616/erdev.2021.20.tf338>

## Appendices

### Appendix A

**Figure A1**

*Screenshot of a question in the PowerApp*



Note. From Bright Alley (2020).

## Appendix B

**Table B1***Adaptation of the reaction/appreciation scale*

	Original item	Adapted item
1	The <b>training</b> took place at a suitable time for me.	The <b>learning session</b> took place at a suitable time for
2	The subject content in the <b>program</b> was relevant to my job.	The subject content in the <b>learning tool</b> was relevant to my <b>(future)</b> job.
3	The <b>training program</b> combined theory and practice.	The <b>learning tool</b> combined theory and practice.
4	The content of the <b>training program</b> included up-to-date theory and practical information.	The content of the <b>learning tool</b> included up-to-date theory and practical information.
5	The material was presented in a manner appropriate to the target group's training needs.	The material was presented in a manner appropriate to the target group's training needs.
6	The audio-visual aids were effective.	The audio-visual aids were effective.
7	The length of the <b>training program</b> was suitable and adequate.	The length of the <b>learning session(s)</b> was/were suitable and adequate.
8	The handouts provided will help me to meet all my training needs.	-
9	The <b>training program</b> was linked to my training needs and my current job tasks.	The <b>learning tool</b> was linked to my training needs and my current job tasks.
10	I feel that the <b>program</b> will help me do my job better in the future.	I feel that the <b>learning tool</b> will help me do my job better in the future.

Note. Original scale from Alsalamah and Callinan (2021)

**Table B2***Adaptation of the attitude scale*

	Original item	Adapted item
1	My knowledge and information developed as a result of the <b>training</b> .	My knowledge and information developed as a result of the <b>learning with the learning tool</b> .
2	Through the <b>training programs</b> , I learned about some laws, theory and practices and learned information I did not know before.	Through the <b>interaction with the learning tool</b> , I learned about some laws, theory and practices and learned information I did not know before.
3	<b>Training programs</b> provided me with practical skills in my field that I did not have before.	<b>The learning tool</b> provided me with practical skills in my field that I did not have before.
4	<b>Training programs</b> provided an opportunity for the exchange of new information, knowledge and experiences among <b>participants</b> .	<b>The learning tool</b> provided an opportunity for the exchange of new information, knowledge and experiences among <b>(future) colleagues</b> .
5	<b>Training programs</b> helped me to succeed in my work in a way that I would not have been able to before.	<b>The learning tool</b> helped me to succeed in my <b>(future)</b> work in a way that I would not have been able to
6	The <b>training programs</b> motivated me and made me interested in learning more.	The <b>learning tool</b> motivated me and made me interested in learning more.
7	The <b>training program</b> has helped to change my attitude towards the topic and training area.	The <b>learning tool</b> has helped to change my attitude towards the topic and training area.

Note. Original scale from Alsalamah and Callinan (2021)



**Table B3***Adaptation of the information security behavior scale*

<b>Device securement:</b>	
	I set my computer screen to automatically lock if I don't use it for a prolonged period of time.
	I use a password/passcode to unlock my laptop or tablet.
	I manually lock my computer screen when I step away from it.
	I use a PIN or passcode to unlock my mobile phone.
<b>Password generation:</b>	
	I do not change my passwords, unless I have to.
	I use different passwords for different accounts that I have.
	When I create a new online account, I try to use a password that goes beyond the site's minimum requirements.
	I do not include special characters in my password if it's not required.
<b>Proactive awareness:</b>	
	When someone sends me a link, I open it without first verifying where it goes.
	I know what website I'm visiting based on its look and feel, rather than by looking at the URL bar.
	I submit information to websites without first verifying that it will be sent securely (e.g., SSL, "https://", a lock icon).
	When browsing websites, I mouseover links to see where they go, before clicking them.
	If I discover a security problem, I continue what I was doing because I assume someone else will fix it.

Note. Original scale from Egelman and Peer (2015)

**Table B4***Adaptation of the behavior scale*

	Original item	Adapted item
1	The training program helped me to organise my role as head teacher more effectively.	-
2	The <b>training programs</b> inspired me to improve my achievement.	The <b>learning tool</b> inspired me to improve my achievement.
3	The <b>training programs</b> increased my ability to perform well in my job role.	The <b>learning tool</b> increased my ability to perform well in my <b>(future)</b> job role.
4	The <b>training programs</b> helped me to develop <b>leadership</b> behaviour.	The <b>learning tool</b> helped me to develop <b>information security</b> behaviour.
5	The <b>training programs</b> developed some aspects of my behaviour.	The <b>learning tool</b> developed some aspects of my behaviour.
6	The <b>training programs</b> helped me to prove myself in my work as <b>head teacher</b> .	The <b>learning tool</b> helped me to prove myself in my <b>(future)</b> work.
7	My <b>job</b> behaviour changed after completing the <b>program</b> .	My <b>information security</b> behaviour changed after completing the <b>learning</b> .

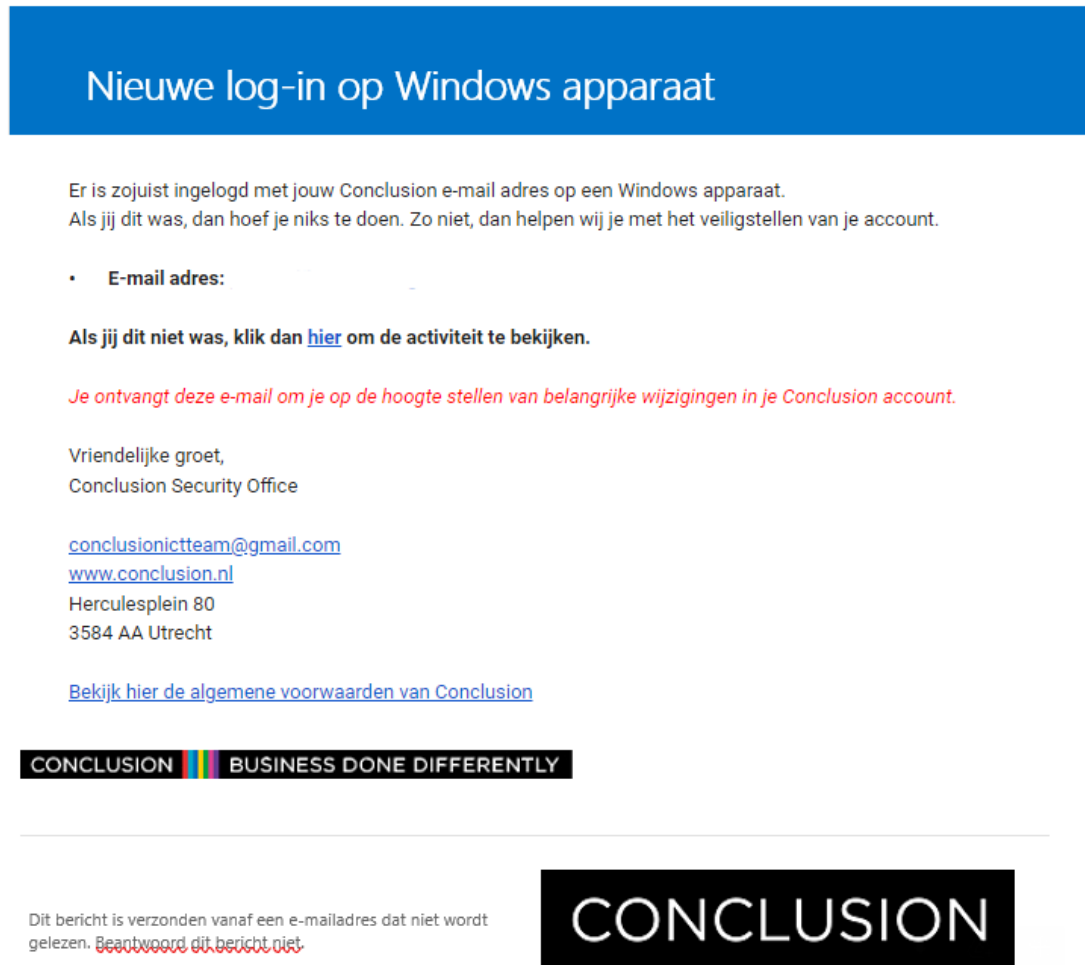
Note. Original scale from Alsalamah and Callinan (2021)



## Appendix C

**Table C1**

*Template of the pre-learning fake phishing e-mail to Conclusion B.V. participants, in Dutch.*



*Note.* The, Dutch, subject of the e-mail was: “Nieuwe log-in op Windows apparaat”.

**Table C2**

*Template of the pre-learning fake phishing e-mail to TU Eindhoven participants.*



Dear reader,

We noticed a new sign-in to your TU/e account on a Windows device. If this was you, you don't need to do anything. If not, we'll help you secure your account. Check the activity [here](#).

Kind regards,  
ICT Team TU/e

**Contact**

Email: [tueictteam@gmail.com](mailto:tueictteam@gmail.com)  
Website: [www.tue.nl](http://www.tue.nl)

Note. The subject of the e-mail was: "A new sign-in on Windows".

**Table C3**

*Template of the post-learning fake phishing e-mail to Conclusion B.V. participants, in Dutch.*

## Urgent: verplicht wachtwoordherstel

Uw beheerder heeft een verplichte wachtwoordwijziging ingesteld voor uw account, vanwege een verdachte hackpoging. Maak direct een nieuw wachtwoord aan om de veiligheid van uw account te verzekeren.

- E-mail adres:


Klik [hier](#) voor meer informatie en om uw wachtwoord te wijzigen.

*U ontvangt deze e-mail om u op de hoogte stellen van belangrijke wijzigingen in uw Conclusion account.*

Vriendelijke groet,  
Conclusion ICT Office

[conclusionictoffice@gmail.com](mailto:conclusionictoffice@gmail.com)  
[www.conclusion.nl](http://www.conclusion.nl)  
Herculesplein 80  
3584 AA Utrecht

[Bekijk hier de algemene voorwaarden van Conclusion](#)

CONCLUSION  BUSINESS DONE DIFFERENTLY

---

Dit bericht is verzonden vanaf een e-mailadres dat niet wordt gelezen. ~~Beantwoord dit bericht niet.~~

CONCLUSION

*Note. The, Dutch, subject of the e-mail was: "Urgent: verplicht wachtwoordherstel".*

**Table C4**

*Template of the post-learning fake phishing e-mail to TU Eindhoven participants.*



Dear reader,

Your administrator has initiated a mandatory password recovery for your account, due to a suspicious hacking attempt. Please change your password directly to ensure the safety of your account.

**E-mail address:**

You can find more information [here](#), as well as a TU/e guideline for changing your password.

Kind regards,  
TU/e ICT Office

**Contact**

Email: [tueictoffice@gmail.com](mailto:tueictoffice@gmail.com)

Website: [www.tue.nl](http://www.tue.nl)

*Note. The subject of the e-mail was: "Urgent: mandatory password recovery".*