IMPROVING KNOWLEDGE RETENTION OF DIGITAL WORK-BASED LEARNINGS USING SPACED LEARNING

Does the instructional method of spaced learning cause better learning in a work-based e-learning environment?
An investigation to enhance learning outcomes of digital work-based learning offerings

Dissertation

at the Faculty of Business Administration and Economics
of Heinrich Heine University Düsseldorf

by
Hanan Kondratjew

Strategic and International Management Series

No. 52

Dissertation for obtaining the degree of Doctor Rerum Politicarum submitted to the Faculty of Business Administration and Economics of Heinrich Heine University Düsseldorf.

Supervisor: Prof. Dr. Andreas Engelen

Co-supervisor: Prof. Dr. Steffen Strese

Chairman examination panel: Prof. Dr. Marius Wehner

The oral examination of the dissertation took place in Düsseldorf on March 29, 2023.

1st edition

D61

Printed in Germany on acid-free paper

Foreword

It is probably one of the most prominent and widespread truisms that continuous learning has developed into one of the main success factors for executives and employees across industry sectors. The ever-changing conditions, e.g., in terms of technologies and customer preferences, require that companies enable their executives and employees to learn continuously. And indeed, many statistics indicate that companies invest more than ever in executive and employee education programs. Traditionally, such programs are held “on the spot” in classroom-like settings. Yet, the digital transformation and the covid crisis have advanced possibilities and the acceptance of online training quite substantially. In contrast to classroom settings, such online training is much more flexible, e.g., in terms of how to split learning sessions. Therefore, decision-makers in personnel development and executive and employee education in general need to design such pieces of training and take various variables (e.g., timing, pace) into account. While the broader learning literature provides insights, it is relatively silent on how work-based learning online formats should be designed.

Hanan Kondratjew recognizes this gap in the learning and also management literature and presents a dissertation that sheds light on how management training formats should be designed based on the spaced learning approach. Spaced learning refers to learning interventions across time that are assumed to enhance learning retention. Specifically, Hanan Kondratjew asks whether the instructional method of spaced learning causes better learning in a work-based e-learning environment. To provide insights into this domain, based on sound theoretical arguments and literature analyses, she presents two experiments. Experiment 1 taught factual and conceptual knowledge on the exemplary topic of “platform business models”, whereas experiment 2 taught procedural knowledge on the exemplary topic of “time management”. Findings indicate that there are important nuances suggesting that a tailor-made approach is needed depending on the subject at hand. Overall, the reader learns that spaced learning also works in this work-based context with pieces of training and education on management-related topics.

Hanan Kondratjew presents a very good and convincing dissertation. She recognizes an important and very interesting research area and provides theoretically derived empirical insights that enable decision-makers in companies to develop fine-grained online training formats. I wish the dissertation the large readership it surely deserves.

Prof. Dr. Andreas Engelen

Acknowledgements

Since it is often very difficult for me to put into words what I would like to say, I will keep it very short and just say it out loud and with all the sincerity I can: Thank you.

First and foremost, Knut Alicke, Andreas Engelen and Nils Müller.

Also, Jan Bachmann, Juliana Gottschling, Anna Gründler, Deryk Hopley, Marion Kahrens, Dennis Küsters, Patrick Schwaferts, Matthias Stadler, Asbjørn Berg Sørensen, Pascal Lange, Anna Liashenko, Antje Keahiolalo, Taylor Kowalchuk, Fabio Serena and Jan Veira.

Finally, I dedicate this work to my two children. Alba, who will always live in my heart, and without whose whole short life I might never have had the opportunity to do a PhD. Edgar, who ultimately gave me the strength and support to finish this work and who has had to listen to the results of my studies on so many walks. We will see if the inter-session intervals on these walks were long enough for lifelong learning.

Summary

Labour market projections indicate that by 2030, about one third of the global work force needs to be up- and reskilled due to advancements in automation and digital transformations (Zahidi, 2020). To ensure long-lasting competitive advantage, formal work-based learning interventions have become unavoidable for businesses. Even though more than USD 350 billion have been invested globally into work-based learning in 2019, this investment was not efficient since the learning offering was neither satisfying, nor focusing on the right content to upskill employees (Beer et al., 2016; Gartner, 2018). Employees even cite poor training and development interventions as the main reason for changing employers (Bersin, 2018). It appears that designers and providers of work-based learning interventions as well as executives in charge have little clarity on how to design learning interventions which lead to long-term learning and knowledge transfer (Beier, 2021; Billett, 2014; Glaveski, 2019; Tuijnman & Boström, 2002; Vargas, 2017). Hence, learning budgets are largely mis-invested, as learning contents are quickly forgotten. Getting work-based learning 'right' is positively correlated with e.g., job satisfaction, organisational performance, innovation, and improved decision-making (e.g., Bersin, 2018; Ellis & Kuznia, 2014; Ellinger, 2004; Kontoghiorghes et al., 2005; Pfeffer & Veiga, 1999; Rose et al., 2009; Ryu & Moon, 2019; Spicer & Sadler-Smith, 2006). Although there is basic agreement for these positive aspects, actual implementation is lacking.

It is important to understand that learning depends on three memory functions: encoding, storage, and retrieval (Spielman et al., 2018). Even though all three functions are necessary to learn, especially the functions of storage and retrieval are of importance when it comes to work-based learning, since these enable learners to transfer and apply what they have learnt. The most reliable and meaningful phenomenon of human memory to do so is the spaced learning effect. Studied for more than 100 years, it refers to a powerful long-term memory advantage, resulting from deliberately scheduled repetitions of what has to be learnt (Carpenter et al., 2012; Cepeda

et al., 2006; Delaney et al., 2010; Dempster, 1989; Gerbier & Toppino, 2015; Vlach et al., 2019).

The extant literature however has paid little attention to its effects on real-world educational learning interventions. Past research has mostly been focused on verbatim and language learning, i.e., factual knowledge (Anderson, Krathwohl et al., 2001), and research within the field of lifelong learning and skills such as technological, social, emotional, and higher cognitive skills, i.e., procedural knowledge (Anderson, Krathwohl et al., 2001), is missing. The latter skills however are said to be of highest importance for the future of work (Bodem-Schrötgens et al., 2021; Bughin et al., 2018; Cepeda et al., 2006). Furthermore, and driven by the COVID-19 pandemic, the importance and use of digital learning offerings such as e-learning interventions (as defined by Clark and Mayer, 2016) has increased tremendously. Yet, no spaced learning research is known that has investigated if effects become evident in e-learning interventions.

To address the above-described research gap, the following overarching research question was formulated:

Does the instructional method of spaced learning cause better learning in a work-based e-learning environment?

In order to answer this research question, eight hypotheses were derived from the existing literature. These were evaluated in two separate field experiments, in which primary data in the form of knowledge tests and surveys were collected. Experiment 1 taught factual and conceptual knowledge on the exemplary topic of "platform business models", whereas experiment 2 taught procedural knowledge on the exemplary topic of "time management". Within both experiments, a between-subjects experimental design was used, differentiating seven groups. One followed a massed learning condition, studying all learning content in one day. This group took one knowledge test after a retention interval of two weeks, serving as control group for groups 2 to 5 and another test after a retention interval of four weeks, serving as control group for groups 6 and 7. Groups 2 to 7 followed equal spaced learning conditions, whereby learning sequences, inter-session and retention intervals varied. At the end of the experiments, learning outcomes and metacognitive beliefs were compared between all groups in the final tests. Furthermore, a survey was conducted with those participants who took part in both experiments.

The results from experiment 1 indicated that spaced learning leads to better knowledge retention than massed learning, especially at long retention intervals. These findings are in line with previous spaced learning research. In contrast, experiment 2 indicated none or very limited difference in learning outcomes for neither the spaced learning groups nor the massed learning group. When comparing the effects of the two field experiments, a strong indication arose that the spaced learning effect is not just modulated by learning schedule design but also by the type of knowledge taught. Thereby, all results were largely independent from other factors such as testing, immediate feedback and number of learning sessions, which were said to enhance the effect. In terms of learner's preference, in both experiments independent of the knowledge type taught, metacognition analyses showed that learners in general preferred and asked for the application of spaced, interactive, and guided work-based learning interventions, hence supplementing the partial positive effect on knowledge retention with a positive learning experience from spaced learning.

This research contributes to the literature by being the first known effort to find evidence of the spaced learning effect in real-world work-based learning environments that teach factual and conceptual knowledge. Further, it extends the previous literature by postulating that the occurrence of the spaced learning effect depends, among other aspects, on the knowledge type taught. On a practical note, this research offers meaningful and beneficial insights for any designer and provider of work-based e-learning interventions and for managers and executives seeking long-term knowledge retention of individuals taking part in these interventions. As, regardless of type of knowledge taught, learners overwhelmingly preferred spaced, interactive, and guided learning sessions, designers and providers of work-based learning interventions shall ensure learning interventions are designed this way, even though learning outcomes are only enhanced for factual and conceptual knowledge. Managers and executives in charge of learning and development are encouraged to rethink the way learning is conducted in their organisations and how they spend learning budgets, since it has been demonstrated that the approaches in use today are not always economical or desirable.

Table of contents

List of Figures

List of Tables

List of abbreviations, acronyms, and symbols

ACT-R	Adaptive control of thought-rational
ANOVA	Analysis of variance
df	Degrees of freedom
EUR	Euro
F	Test statistic for ANOVA
ISI	Inter-session interval
JOL	Judgment of learning
n/a	Not applicable
OECD	Organisation for Economic Co-Operation and Development
p	Probability of evidence for the credibility of the null hypothesis
PPE	Predictive Performance Equation
p_{tukey}	Probability result from Tukey post hoc test
RI	Retention interval
SAM	Search of Associative Memory
SE	Standard error of means
t	Calculated difference represented in units of standard error
UNESCO	United Nations Educational, Scientific and Cultural Organization
USD	US Dollar
η^2	Effect size η^2

1. Introduction

1.1 Context and purpose of the study

For businesses, working in ever faster and changing economies forces their employees to learn constantly (Zuber, 2014). The Organisation for Economic Co-Operation and Development (OECD; 2021) as well as the World Economic Forum (Zahidi, 2020) claimed that at the end of 2020, about 114 million jobs had disappeared globally compared to the preceding year. It is expected that by 2030, about a third of all workers globally (1 billion individuals) need to be reskilled because of digital transformations and automation. This development is said to make manual, physical, and basic cognitive skills redundant since these could be mastered by machines. Whereas skills that machines cannot master, like technological, social, emotional as well as higher cognitive skills are becoming increasingly important (Bodem-Schrötgens et al., 2021; Bughin et al., 2018). Amongst the top skill groups in rising demand over the next years are critical thinking, analysis, problem solving, self-management, stress-tolerance, resilience, and flexibility (Zahidi et al., 2020).

Skill shortages are regarded as one of the biggest threats to company growth as they, among other things, are claimed to hinder innovation, diminish quality, and obstruct market positions (Moritz et al., 2019). Taking the view of employees, skill shortages are associated with a higher chance of unemployment, lower incomes, and job dissatisfaction (Dondi et al., 2021). Having surveyed about 18,000 people in 15 countries, Dondi and team (2021) came up with a range of skills which would benefit every employee regardless of sector or occupation and which in return would also benefit companies' growth. Similar to Zahidi and teams' findings (2020), their study revealed 56 so-called foundational skills or distinct elements of talent (see Table 1 for an overview), which are said to help people succeed in the future labour market, since they assist them in working and operating in digital environments, being able to constantly adapt to new professions and work procedures and generating added value to what artificial intelligence, robotics or machines have to offer (Dondi et al., 2021). These skill categories align with Anderson, Krathwohl and teams' (2001) definition of *procedural knowledge*, while skills more at risk to being automated or digitalized would be considered as *factual* and *conceptual* (Adams, 2015).

Table 1 – Foundational skills needed to succeed in the future of work, adapted from Dondi et al., 2021, p. 3

Cognitive	Interpersonal	Self-leadership	Digital
▪ Critical thinking ▪ Planning and ways of working ▪ Communication ▪ Mental flexibility	▪ Mobilizing systems ▪ Developing relation-ships ▪ Teamwork effective-ness	▪ Self-awareness and self-management ▪ Entrepreneurship ▪ Goals achievement	▪ Digital fluency and citizenship ▪ Software use and development ▪ Understanding digital systems

To overcome potential skill shortages and to stay economically competitive, constant up- and reskilling, especially in the listed areas is required. *Work-based learning interventions* can be seen as a good means for this, given the huge amount of time individuals spend in the workplace (European Commission, 2018; Poquet & de Laat, 2021). It is widely agreed that work-based learning opportunities, i.e., "*the systematic development of the knowledge, skills and attitudes required by an employee to effectively perform a given task*" (Patrick, 2000, cited in Ryu & Moon, 2019, p. 482), are also key for high-performing businesses and organisations (Kontoghiorghes et al., 2005; Pfeffer & Veiga, 1999; Porth et al., 1999; Spicer & Sadler-Smith, 2006).

However, learning and instruction are not of value if learners do not retain what they have learnt or are unable to transfer or put it into actual use, i.e. no skill is created (Connors, 2021; Gallardo, 2021; Göldi, 2011; Ryo & Moon, 2019; Sala & Gobet, 2017). And so, organisations and executive managers have found it very difficult to create appropriate work-based learning offerings, and the approaches to learning and development largely relied on methods which are not appropriate for work-based learning (Ellinger, 2004; Tauber et al., 2019). As a result, work-based learning happened largely for the wrong reasons, at the wrong time, with competencies and skills being taught that have little or no relevance to the role that must be performed. In many organisations learning was regarded as simply having something on offer but not about learning the right skills needed to perform a certain role (Glaveski, 2019).

This discrepancy derives from two different standpoints: whilst executive managers focus on generating measurable financial impact (resulting from having employees with the right *skills* to perform a certain role), learning and development departments focus on aspects such as satisfaction and completion of courses – all related to the *learning* experience (Tauber et al., 2019). But learning and skilling cannot be equated. Even though learning is the foundation of real skilling, it cannot be benchmarked or measured and if not applied, it will be forgotten (Ebbinghaus, 1885/1962; Tauber et al., 2019). Skills on the other hand can be benchmarked, measured, and evaluated and investing in those is at the heart of today's ever changing business world (Tauber et al., 2019).

Work-based learning therefore requires transformative, adaptable learning strategies, which keep up with changing circumstances, leading to transferrable skills allowing for long-term application (Glaveski, 2019; Kane et al., 2018; Tauber et al., 2019). Over the last few years, changes have been observed in how learning and development departments are regarded within organisations and how they have stepped out of a secondary role in the company and have taken on one of the most important roles for the strategic orientation of companies to overcome skill shortages and sustain competitive advantage, which is evident in increased budgets and executive management support (see Figure 1; Chelovechkov & Spar, 2019).

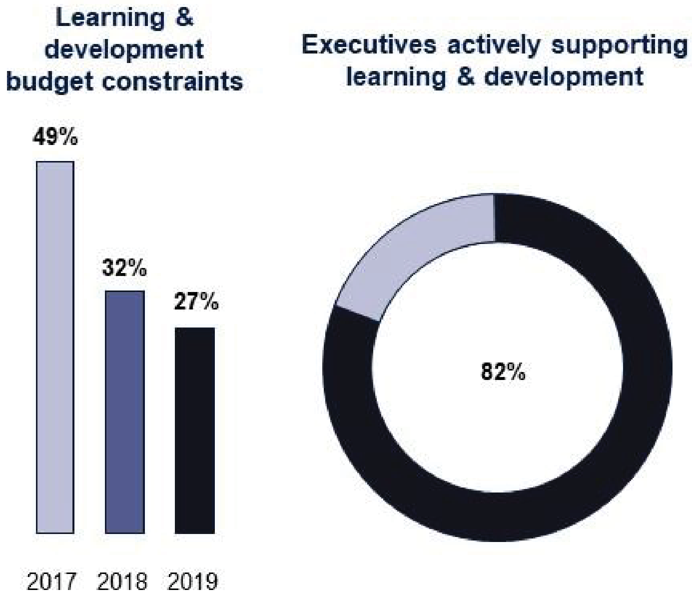


Figure 1 – Increased budgets and executive support for work-based learning interventions, adapted from Chelovechkov & Spar, 2019, p. 7

Globally, organisations and businesses invested about USD 359 billion on work-based learning interventions in 2019 (Glaveski, 2019) and business managers are expecting to re- and upskill about 70 percent of their workforce by 2025 (Zahidi et al., 2020). Beer and associates (2016) argued in their survey that this investment did not pay off as, among other things, the majority of the 1,500 managers asked at 50 different organisations were not satisfied with their organisations' learning and development offering. Also, 52 percent out of 5,997 employees receiving work-based learning interventions stated, that they need better upskilling, i.e., they do not possess the right skills to do their jobs (Gartner, 2018). Only one-quarter of respondents to a 2010 McKinsey survey were convinced that their work-based learning offerings significantly impacted business performance whilst most organisations do not even track the return on their investment in training (Gryger et al., 2010) and importantly, only 42 percent of workers is predicted to participate in offered re- and upskilling learning interventions (Zahidi et al., 2020). Employees are directly addressing their desire to receive guidance on what needs to be learnt to make their individual learning more relevant, e.g., 61 percent out of 772 executives, managers and employees surveyed stated that they would like to align their learning to actual skill gaps, and another 48 percent of respondents asked for assessments of knowledge and skills to find out in which areas improvements are needed (Tauber et al., 2019). Yet, those are largely missing (Tauber et al., 2019). Both external providers of work-based learning offerings as well as employers have little clarity on how to effectively design work-based learning interventions to maximise their effect on knowledge advancement, retention and eventually long-term application by incorporating biological and neuroscientific research findings (Beier, 2021; Billett, 2014; Glaveski, 2019; Tuijnman & Boström, 2002; Vargas, 2017). Disregarding this shortcoming leads to billions being spent on learning content that is quickly forgotten and the money is simply mis-invested (Glaveski, 2019). Additionally, ineffective learning opportunities at work were found to be the most important thing that would make employees look for a new job (see Figure 2; Bersin, 2018), whereas companies that get learning 'right' are 21 percent less likely to lose employees (Tauber et al., 2019).

Figure 2 – Reasons that would make employees look for a new job, adapted from Bersin, 2018

Summarising, learning is seen as the new currency of global job markets, where career development and even just the possibility of it are valued more than rising salaries (Joseph, 2019). 94 percent of employees would consider staying longer with a company if they would invest in their learning (Chelovechkov & Spar, 2019). The fact that work-based learning is positively correlated with job satisfaction, organisational performance, organisational innovation, improved decision-making, organisational commitment, employee productivity and a reduction in work-related stress has been confirmed by business research for a very long time (e.g., Bersin, 2018; Ellis & Kuznia, 2014; Ellinger, 2004; Kontoghiorghes et al., 2005; Pfeffer & Veiga, 1999; Porth et al., 1999; Rose et al., 2009; Ryu & Moon, 2019; Spicer & Sadler-Smith, 2006; Udemy, 2018). Especially, the returns of *lifelong work-based learning* are two-fold: for individuals it is said to enhance employability and salaries, for organisations it is promised to lead to higher productivity and innovation (European Commission, 2018; Hanushek & Wößmann, 2010; Rees, 2010; Tynjälä, 2008; Vargas, 2017). Therefore, organisations, managers and departments in charge as well as external providers of work-based learning offerings would do well to create high-quality learning opportunities, taking into account biological and neuroscientific research findings, that lead to an improvement of job-relevant skills for individual employees and at the same time create conditions that allow employees to integrate learning opportunities into their daily work (Glaveski, 2019; Ryu & Moon, 2019; Tauber et al., 2019).

All of the above applies similarly to digital, online, or e-learning offerings, i.e., the delivery of learning material and instructional design through digital devices (Clark & Mayer, 2016; Ellis & Kuznia, 2014). These offerings, as well as investments in these, have risen tremendously over the last years, allowing for delivering learning opportunities at scale (Figure 3; Chelovechkov & Spar, 2019; Seyda & Placke, 2020; Tauber et al. 2019).

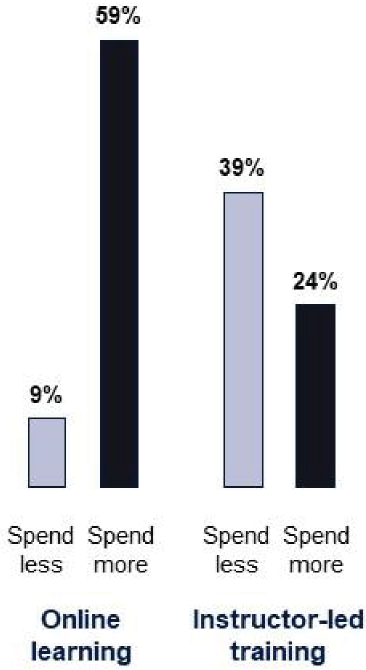


Figure 3 – Learning budgets shift to online learning offerings, adapted from Chelovechkov & Spar, 2019, p. 14

Even though e-learning offerings might never fully replace in-person learning, they a) lead to higher productivity and faster onboarding times (Udemy, 2018) and b) allow for the flexibility learners stemming from different generations ask for, ideally leading to cross-generational alliances during the learning process (Chelovechkov & Spar, 2019). More precisely, learners are asking for:

- Mobile, social, and collaborative learning environments
- Self-directed and independent learning during spare time at work
- Sufficient time to learn

(Chelovechkov & Spar, 2019; Tauber et al., 2019; Figure 4)

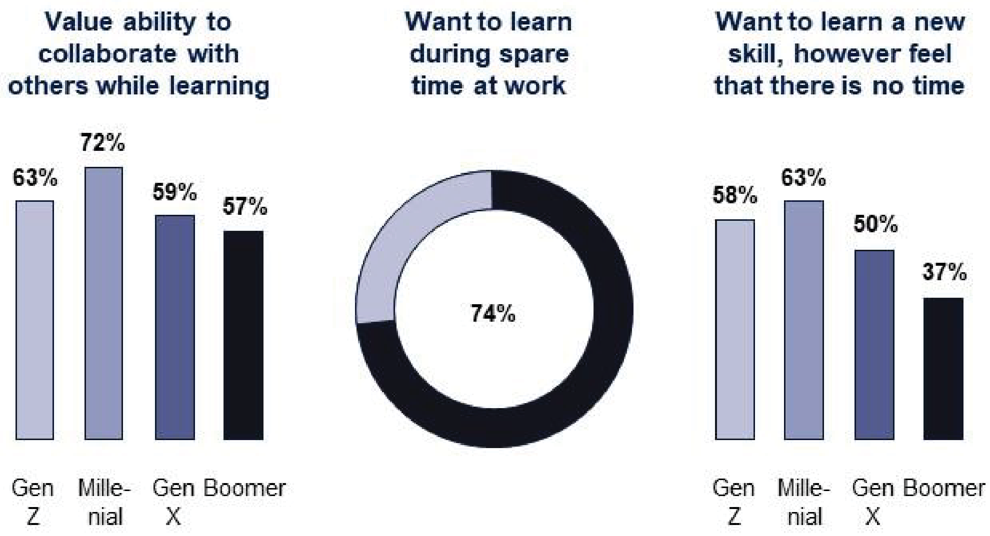


Figure 4 – Analysis what cross-generational learners are asking for, adapted from Chelovechkov & Spar, 2019, pp. 34

However, none of the benefits of learning will occur in the company if there is no sustained management support or if a company's culture is against change (Ellis & Kuznia, 2014). But it is precisely this change that distinguishes high-engagement and high-growth companies such as Kellogg, Allianz or Service Titan from others and generates competitive advantage (Chelovechkov & Spar, 2019; Udemy, 2018). Psychological and business-related research on learning has shown an importance to

- individually assess and benchmark skills needed to perform a (future) role, to stay ahead of interferences,
- immediately transfer what was learnt to real-life situations to allow skill building,
- provide immediate feedback and guide learners through the learning experience and
- repeat these steps along the process.

Despite all this knowledge, there is clear lack in implementation (Beier, 2021; Billett, 2014; Glaveski, 2019; Tuijnman & Boström, 2002; Vargas, 2017). Any learning is only going to happen when it is built around three basic memory functions, namely *encoding*, *storage* and (Spielman et al., 2018). Whilst encoding refers to getting new information into memory through different instructional methods, storage and retrieval are concerned with retaining, recalling, and recognising the information learnt and ideally, transferring it to new situations (Spielman et al., 2018). In the context of work-based learning environments, the functions storage and retrieval are of particular importance, as learning takes place every day at any time, also during work. However, for what is learnt to be of lasting relevance, it must either be repeated at fixed intervals over a longer period and thus consolidated or integrated into the work so that repetition creates an automation effect. After all, about 50 percent of information taught in a learning unit is forgotten if it is not repeated or applied within 24 hours and another 75 percent after six days (Blanchard, 2013; Ebbinghaus 1885/1962).

A well-known and researched instructional method to do exactly that and thus enhance long-term knowledge retention is *spaced learning* (e.g., Cepeda et al., 2006; Delaney et al., 2010; Vlach et al., 2019). First mentioned in research in 1885, it argues that distributing learning interventions across time slows overall learning but enhances knowledge retention (Cepeda et al., 2006; Delaney et al., 2010; Latimier et al., 2021; Walsh et al., 2018a/b). The benefits of spaced learning on knowledge retention have been analysed in depth, mainly derived from low-complexity, verbal learning research in a laboratory with relatively short time delays between learning interventions (Balota et al., 2011; Cepeda et al., 2006; Mozer et al., 2009). Interestingly, learners perceive spaced learning interventions as less helpful than traditional ones (so-called *massed* learning), which is why the topic of learner's self-perception (or *metacognition*) must be kept in mind when designing learning interventions (Vlach et al., 2019).

According to the current state of knowledge, there is no sustainable application of spaced learning in any educational context, neither schools nor work-based learning settings (Dempster, 1988; Kapler et al., 2015; Walsh et al., 2018b). Even worse is its application in a management context with only one study examining the effects of spaced learning to an experiential management learning (Kondratjew & Kahrens, 2018). What thus remains critical to be investigated is if the proven positive effect of spaced learning on learning retention also occurs when applied in a complex work-based learning intervention in a business environment.

This study aims to address this need. Concretely, the purpose of this study is to *improve the learning outcomes of digital work-based learning interventions with the help of spaced learning*, whilst considering the learners' self-perception. For this, two e-learning field experiments (Gerber, 2011) on the exemplary topics of "platform business models" (to test factual and conceptual knowledge) as well as "time management" (to test procedural knowledge) have been designed. Based on the results of this research, recommendations for future digital work-based learning interventions are derived.

1.2 Importance of the study

Formal work-based learning interventions are inevitable for both employees and employers if they want to ensure competitive advantages on both individual and organisational levels. Yet very little is done to ensure long-term application and transfer of what was learnt, which would, in turn, extend the economic advantages resulting from education. In this context, this research is led by the overarching research question:

Does the instructional method of spaced learning cause better learning in a work-based e-learning environment?

To answer the research question, the following research objectives are addressed:

a. Critically review the concept(s) of lifelong learning and the role of work-based learning
b. Critically review the theories of spaced learning and its benefits
c. Analyse how to optimally design multimedia e-learning interventions
d. Review different knowledge types and how memory is best measured

e. Experimentally explore the effects of the application of spaced learning activities on learner's learning, i.e., recognition and recall to work-based e-learning interventions

f. Draw up recommendations for future work-based e-learning interventions.

The outcomes are derived through the method of experimental research and aims to be beneficial for any designer and provider of digital work-based learning interventions who seek long-term knowledge retention of learners, which in turn shall be transferred into work-related skills. Moreover, this study might contribute to other academic research in the field of learning as it offers new insights about a specific phenomenon, namely if spaced learning can be applied to complex, non-laboratory, work-based e-learning interventions.

1.3 Structure of the study

This research is structured along five chapters. The first chapter, the *introduction*, establishes the context and purpose of the study. Thereby, outlining the significance of the study for research and practice. The guiding research question and objectives are introduced.

The second chapter, the *literature review*, provides a basic understanding of the neurosciences of *learning*, discusses the concept(s) of lifelong learning, the role of work-based learning and its meaning for the wider economy. The theory of *spaced learning* as a means to enhance long-term knowledge retention of educational interventions is reviewed while also touching upon *retrieval* practice and on the learners' declarative *metacognitive* knowledge. It further elaborates on basic *instructional design* concepts, thereby bridging learning theory and educational practice, introducing *human cognitive architecture*, *cognitive load theory*, *multimedia instruction* as well as fundamentals of the *evaluation of learning*. Hereafter, in chapter three, a summary of the *gap in current state of research* with regards to the topic of managerial, lifelong learning is discussed. Within this chapter, *eight hypotheses* are derived. The fourth chapter then outlines the *experimental objects of this research* (two e-learning interventions) and explains how they were designed from a *learning design* perspective. Afterwards, the chapter explains the *method and results and discusses the two experiments* conducted to answer the research question and hypotheses as well as compares the two experiments against each other. In the fifth and final chapter, a general conclusion is drawn from the *key findings* of the research at hand, *limitations* of the experiments are highlighted and *recommendations for future work-based e-learning interventions* as well as *future research* are made. The structure of this research is shown in Figure 5.

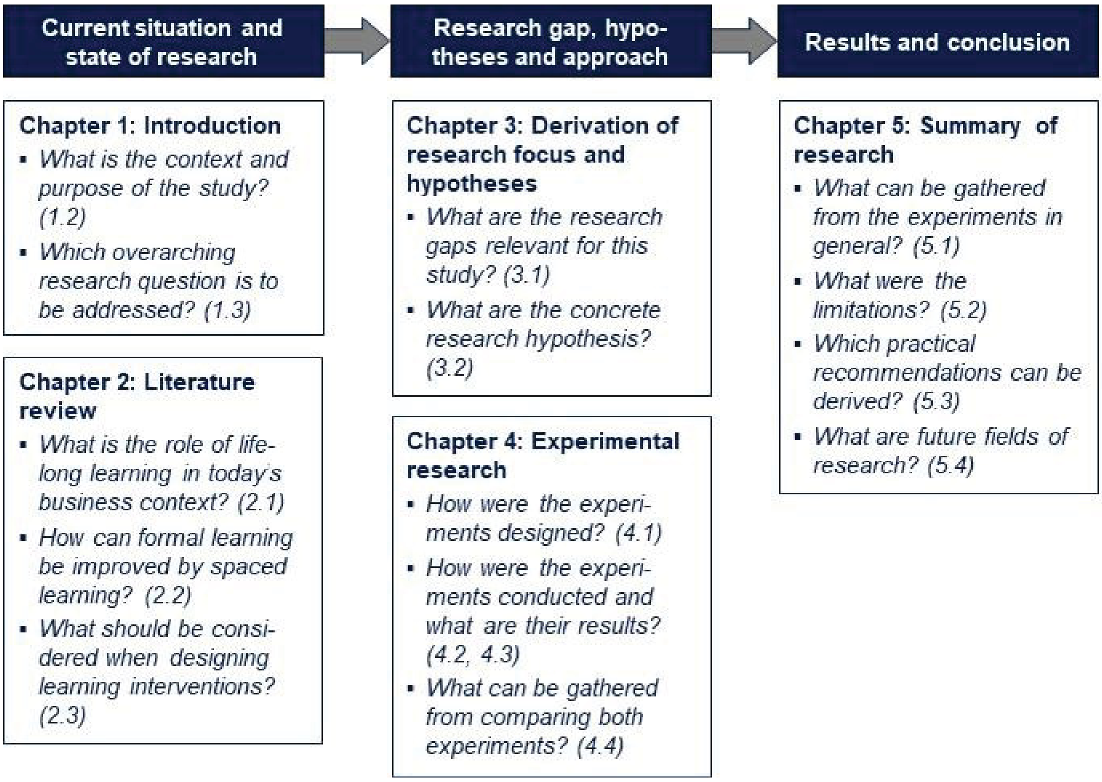


Figure 5 – Structure of the research; author's own compilation

2. Literature review

2.1 Learning – From psychology to economics

The first chapters of this review provide an *overview on learning, memory, and retrieval* from a cognitive psychological perspective (chapter 2.1.1) to allow for a better understanding of the subsequent chapters. Chapter 2.1.2 then elaborates on the *concept of lifelong learning* from both a philosophical and the much broader political perspective. In this context, the role of *work-based learning* (chapter 2.1.2.1) is described, highlighting the importance of it for individuals, organisations as well as the broader economy. By doing so, light is shed upon the *economics of education* (chapter 2.1.2.2). After the case of application of *work-based learning in Germany* in 2019 is presented in chapter 2.1.2.3, the findings are summarised (chapter 2.1.3).

2.1.1 Learning, memory, and retrieval

Several definitions of the terms *learning*, *memory,* and *retrieval* exist. Learning and memory psychologists as well as researchers have an ongoing debate about these (Ertmer & Newby, 2013; Khalil & Elkhider, 2016) and still, not everything is known about the nature of learning and memory – only ideas and models exist on how both processes work (Dirksen, 2016). Schunk (2012) argued that this is because different researchers put emphasis on different aspects central to learning. For example, Gazzaniga and associates (2014) defined learning as "*the process of acquiring [...] new information, and the outcome of learning is memory. That is, a memory is created when something is learned, and this learning may occur either by a single exposure or by repetition of information, experiences, or actions*" (p. 380). Based on this, one could assume that learning is equivalent to memory formation – which is true only one way: learning cannot happen without memory, yet memory still exists without long lasting learning success (Roth, 2011). To address this, Schunk (2012) proposed a more explicit definition of learning which tries to capture the term in full, and defined learning as "*[...] an* *enduring change in behaviour, or in the capacity to behave in a given fashion, which results from practice or other forms of experience*" (p. 3). This definition focuses on the following three aspects: first, learning is a change. Second, learning does not merely mean a change in what those learners know but also a change in what the learner does, i.e., how the learner behaves. Third, the change results from an experience.

Overall, one can argue that learning is only successful when it involves the three basic cognitive memory functions of *encoding*, *storage,* and *retrieval* (Spielman et al., 2018). In brief, *"[e]ncoding is the act of getting information into our memory system through automatic or effortful processing. Storage is retention of the information […]"* (Spielman et al., 2018, p. 79). The most common distinction of memories is made between the *chronological* and *functional* memory (Roth, 2011). Within chronological memories, memory psychologists generally distinguish between very short-lived memories such as sensory memory, short- to medium-lived memories such as short-term memory (at times identical to the so-called working memory) as well as long-living memories, the so-called *long-term-memory* (Table 2; Dirksen, 2016; Gazzaniga et al., 2014; Huppelsberg & Walter, 2009).

Table 2 – Types and characteristics of memory, adapted from Gazzaniga et al., 2014, p. 380

	Characteristic of memory			
Type of memory	**Time course**	**Capacity**	**Conscious awareness**	**Mechanism of loss**
Sensory	Milliseconds to seconds	High	No	Primarily decay
Short-term and working	Seconds to minutes	Limited (7±2 items)	Yes	Primarily decay
Long-term nondeclarative	Days to years	High	No	Primarily interference
Long-term declarative	Days to years	High	Yes	Primarily interference

Long-term memory can be divided again by functionality, depending on the kind of stimuli which needs to be stored (Gazzaniga et al., 2014; Roth, 2011). Squire (1987) and Schacter (1996) distinguish between *declarative/explicit* and *procedural/implicit* memory.

Declarative memory relates to all knowledge which we acquired consciously through repetitive exposure and can talk about (knowing-that/-what), whereas procedural memory refers to all learnt cognitive and motoric skills that we know we possess (knowing-how), gained mostly unconsciously through physically trying and performing (Roth, 2011). A third type, *emotional* memory, which was previously seen as a sub-memory of procedural memory, is now viewed as a self-standing memory which shows characteristics of both declarative and procedural memory (Roth, 2011). Figure 6 shows a brief overview of the three types of memory and their further division into sub-memories. The research at hand explores declarative and procedural memory, dealing with the effects of repeated opportunities to retrieve taught facts and concepts as well as knowledge about how to accomplish something from memory.

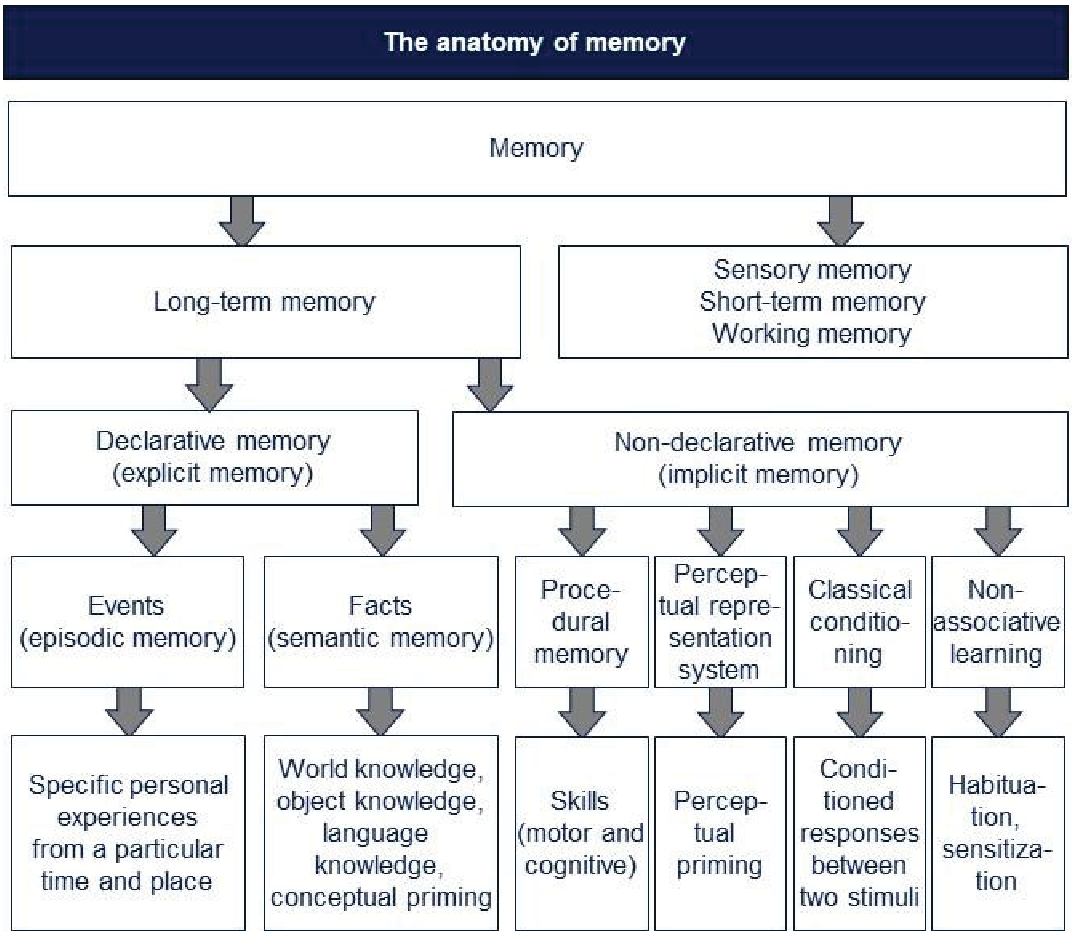


Figure 6 – Hypothesised overview of memory, adapted from Gazzaniga et al., 2014, p. 381

Retrieving information from memory, i.e., from storage, can happen through recall, recognition, and relearning (Spielman et al., 2018). Through repeated interventions, retention can be enhanced as retrieval is eased. Being able to retrieve knowledge or just parts of it from memory enhances efficiency of completing tasks related to that knowledge and allows individuals to handle increasingly complex problems. It also enables individuals to focus only on relevant information and disregard everything unrelated to a certain topic. This in turn eases further knowledge accumulation of new or related topics (Benjamin & Tullis, 2010). Further information on human cognitive architecture follows in chapter 3.

2.1.2 Lifelong learning and the role of the workplace

Even if some believe all life-relevant learning takes place in educational institutions provided mainly by governments and religious organisations or that this learning is superior to all other kinds of learning, this notion is considered too short-sighted (Billett, 2014). Learning, as explained in the preceding chapter, unconsciously happens all the time and everywhere (Alheit & Dausien, 2018). From a philosophical perspective, "*learning must be understood in its entirety and as the result of interaction between the individual and the individual's environment*" (Bjursell, 2020, p. 677). Learning is a continuous, lifelong process in which everyone participates, and which offers the opportunity to expand and improve abilities in a wide range of skills. It happens throughout our life beyond graduation from (high) school and/or university. One part of lifelong learning is (lifelong) education (Billett, 2018; OECD, 2021; Serrat, 2017). Three modes of learning are classified (OECD, 2021; Rubenson, 2019; Tuijnman & Boström, 2002):

- *"Formal learning refers to intentional and systematic learning in a (state-run) institution which is dedicated to education and provides certificates;*

- *non-formal learning generally refers to intentional and systematic learning outside of a state-run institution;*
- *informal learning generally refers to non-intentional and non-structured learning in a life context such as the workplace, the family, etc.*" (Rubenson, 2019, p. 299)

Although learning is a purely personal, individual process, the concept of *lifelong learning* (English & Mayo, 2021) has made it onto global political agendas through the work of various transnational organisations such as the OECD, the United Nations Educational, Scientific and Cultural Organization (UNESCO), and the European Union (EU) (Milana, 2012). From a political-economic perspective, lifelong learning is seen as a means to safeguard national or even transnational economic goals by ensuring individuals' lifelong employability in ever increasing working life, ever-faster changing working environments as well as technological advances (Billett, 2018; Commission of the European Communities, 2000; European Parliament 2008; Milana, 2012; OECD; 2021; The World Bank, 2003). As such, learning does not stop, but continues throughout life (Bjursell, 2020), whereby adults can individually develop themselves as a measure to address educational and knowledge deficits (Schmidt-Hertha, 2018).

Lifelong learning, from a political standpoint and as proposed by the OECD (2004) "*covers all purposeful learning activity, from the cradle to the grave, that aims to improve knowledge and competencies for all individuals who wish to participate in learning activities*" (p. 1). Originating in the Paul Lengrand report (1970), in which *lifelong education* was introduced as a response to several global crises, that emphasised social responsibility and viewed learning as a human right (English & Mayo, 2021; Milana, 2012). The concept has shifted over the last decades not just in terms of its terminology from *education* to *learning* but further in terms of *focus* and *meaning* (Billett, 2018; Bjursell, 2020; Elfert, 2015). This shift implied a change of accountability of learning outcomes: from a position of seeing education as an institutional fact and a public good with governments in charge to provide structures and educational institutions to counteract social inequalities, towards a capitalist position which stresses education or rather learning as a consumer product, i.e., a private, individual matter (Bjursell, 2020; Commission of the European Communities, 2000; English & Mayo, 2021; OECD, 2007; Tuijnman & Boström, 2002; UNESCO, 2015; Vargas 2017).

When seen as a consumer product, major concerns arise, for example Vargas (2017) argued that "*the conceptualization of education as a service or a commodity that is subject to transaction assumes that it yields individual benefits and represents an advantage for individuals to position themselves better in the social realm than others who do not acquire such a commodity, or who cannot afford it*" (p. 8). This leads to social inequalities and decay (English & Mayo, 2021): whilst some people have the (monetary) means to maintain and acquire new skills needed for newly evolved jobs, others do not and are falling behind (OECD, 2021).

Additionally, requirements for specific roles and tasks are changing at a high speed. To take one example, according to the OECD (2019), six out of ten adults are lacking information and communications technology skills which would be needed for newly emerging jobs, 32 percent of all jobs globally will change drastically due to automation, and a strong shift from jobs in the manufacturing market towards the service sector has taken place. Further still, the share of highly skilled and low skilled jobs has increased over the last twenty years whereas the share of middle-skilled jobs has decreased (OECD, 2019).

This need for lifelong learning to adapt to a changing work environment has further been accentuated since the outbreak of the global COVID-19 pandemic. According to the OECD (2021), roughly 114 million jobs disappeared globally at the end of 2020, compared to 2019. In addition, the World Economic Forum (Zahidi, 2020) expects that the previous estimate of the McKinsey Global Institute (Manyika et al., 2017) of the need to reskill 375 million people globally, i.e., 14 percent of the global workforce, due to digital transformations and automation, dramatically accelerated to one billion people (roughly one third of all workers), mainly caused by the global pandemic. New demands have been placed on organisations as well as their employees, including different skill and competence sets to achieve organisational targets (Billing et al., 2021; European Commission, 2020). Only a small group of individuals in higher skilled jobs or the ones in information and communication technology roles are not said to fear loss of employability – and the majority of individuals are asked to up- or reskill to close the skill gaps compared to the pre-pandemic situation; thus, inequality and exclusion are entrenched within employment statuses and incomes (Billing et al., 2021; James & Thériault, 2020; OECD, 2021; Strack et al., 2021). However, inequalities do not solely occur in work-related areas, but also in relation to "*access and participation to lifelong learning education, which in turn has had consequences for well-being and mental health*" (Watts, 2020, cited in James & Thériault, 2020, p. 129).

As many employees faced sudden unemployment, questions on learning and training were put on hold. In addition, neither were all households equipped with the technological resources needed nor did all adults have the digital skills to participate in online learning formats (James & Thériault, 2020). Also, from the side of employers it was recorded that organisationally provided learning opportunities came to a halt, with interruptions for almost 90 percent of global employees. And although trying to move learning towards digitally enabled trainings and thereby bringing education further along the path of the economy (English & Mayo, 2021), organisations faced challenges in delivering these due to infrastructure issues, insufficient user knowledge or inability to transfer on-site learnings into remote courses (International Labour Organization, 2021). Yet, most people as well as organisations have become more open to online learning and have started partnering with external parties to continue up- and reskilling (International Labour Organization, 2021; Strack et al., 2021).

Within the group of working people, about 68 percent of all individuals are open to switch job roles as a response to the disruptions in working and only four percent were unwilling to reskill under any circumstances. Especially individuals at the beginning or halfway through their career as well as those in roles that face the highest risk of being replaced due to technological advancements showed a high willingness to pursue new career paths (Strack et al., 2021). This said and recalling World Economic Forum's (Zahidi, 2020) assumption of the need to reskill about one third of all working adults, global lifelong learning, especially in a work-based context, is needed at scale (Billett, 2014; CEDEFOP, 2011). More importantly "*lifelong education [and hence, lifelong learning] must be viewed again as a public rather than a private good, beneficial to the economy and the wider society and funded by states rather than individuals accordingly*" (Waller et al., 2020, p. 245).

Hence, timely learner-centred work-based learning solutions are needed, which no longer impart professional knowledge in its full depth, but support individuals in learning competences on and during the job, promote learning processes in the broadest sense, and stimulate the ability to learn (CEDEFOP, 2011; European Commission, 2020; James & Thériault, 2020). These solutions must both be backed up by governments (e.g., by creating awareness of the importance of work-based learning; providing policies and infrastructures and putting funding schemes towards work-based learning in place) and by organisational leadership (e.g., by encouraging all employees regardless of qualification to participate in learning opportunities; making sure organisational innovation policies and training agendas are brought together to ensure both social justice and organisational well-being) (Billing et al., 2021; CEDEFOP, 2011; OECD, 2021). In an ever-changing environment, organisations with a high ability to learn and investment into work-based learning to adapt to new situations are said to have the ultimate competitive advantage as it is this learning which indicates an organisation's resistance and flexibility in uncertain times (Deakin Crick et al., 2013).

To strengthen this argument, the following sub-chapters introduce the concepts of *work-based learning* as well as the *economics of education*.

2.1.2.1 Work-based learning

Within the context of lifelong learning, workplaces represent quite a unique environment. Firstly, due to the large amount of time adults spend there and secondly, due to how adults learn, both in a structured manner and incidentally (European Commission, 2018; Poquet & de Laat, 2021). Workplaces do not just enable individuals to enhance their work-related skills and abilities but also their cross-cutting competences that make them more resistant to any changes occurring in their life or careers (Beier, 2021; European Commission, 2018). This happens mostly through work-based learning which continues to gain attention and importance since the early 1990s (Tynjälä, 2008).

As with all other terms related to learning, no universal definition of the term *work-based learning* exists (Ogunleye, 2013). Yet, it can be seen as "*the interlinked practices of performing job-related tasks, building capacity to perform those tasks (as in work and learning) and measuring outcomes of those efforts in terms of both the individual's capacity to perform and the impact on the organisation that sponsored the learning programme*" (Carliner et al., 2006, cited in Ogunleye, 2013, p. 182). Work-based learning thus allows individuals to secure employability through up- and reskilling and is an efficient way for employers to build a competitive advantage, to remain productive, innovative and modern as well as to improve employee morale, job satisfaction and reputation as the employer (Beier, 2021; Billing et al., 2021; Deakin Crick et al., 2013; European Commission, 2018). Initial prompts for employers to offer work-based learnings are derived out of work-related processes, updates in industry standards, market demands and innovation intentions (Brandi & Iannone, 2017).

Learning, as discussed in the previous part of this chapter, can occur formally, non-formally or informally. Whilst some may argue that learning only takes place through official educational programmes, most of the learning in individuals' adult life stems from their working experience and experiences at work, more precisely from informal learning situations (Brandi & Iannone, 2017; Milligan et al., 2014). Yet, not everything that enables adults to sustain the skillset of their current work position or develop skills for an alternative one can be learnt from work experiences – whilst informal learning mostly leads to implicit knowledge, formal learning leads to explicit knowledge (Tynjälä, 2008). Thus, formal learning opportunities are needed alongside informal ones – both even supplement each other – although intended learning outcomes of such provisions might not be achieved as individual learning depends on several factors including the individuals situation, their engagement during the learning experience, and their professional and educational background (Billett, 2014; Tynjälä, 2008).

Even though researchers argue that learning is a natural human lifelong process and in its essence places learners' needs and aspirations at the very core, tensions arise between employers and employees. Especially when it comes to trying to find a consensus on what needs to be learnt, how these learning interventions are designed and how learning outcomes are assessed, assuming they exist at all (Beier, 2021; Billett, 2014; Tuijnman & Boström, 2002; Vargas, 2017). Without addressing the appropriate content, learning interventions provided by employers or external learning providers can at times be viewed as being a superfluous business objective only (Vargas, 2017). Thereby, it is rather common that employees, i.e., the learners, are not consulted regarding content creation and third parties get to decide on and regulate learning frameworks without canvassing the learners' opinions (Billett, 2014; Tuijnman & Boström, 2002). However, Brandi and Iannone (2017) found out that a company's "*learning needs are largely identified by individuals [and thus claimed that] (e)nterprises rely on their employees to identify knowledge and skill gaps and also fulfil learning needs*" (p. 4).

As a result, is it arguable that most work-based learning interventions are designed without considering learners' actual needs. This then even discourages learning, especially learning that should enable learners, to develop work-related skills and abilities which allow transfer to current or other types of situations (Tynjälä, 2008). Further, work-based learning, as many lifelong learning arrangements, is rather seen as a product the learner consumes and not as a process in which the individual must pass all stages of learning as outlined in chapter 2.1.1 (English & Mayo, 2021). Thereby, training design and strategy do not exemplify the holistic perspective of lifelong learning (Beier, 2021; Dunlosky et al., 2013; Pashler et al., 2007; Tuijnman & Boström, 2002).

Nevertheless, employers are agreeing that investing into their employees' learning and skill building is most important to overcome the skill gaps and shortages caused by global economic changes, compared to hiring, contracting, redeploying and releasing (Billing et al., 2021). Work-based learning must be supported from top management, always bearing in mind that it is the individual who learns (Agrawal et al., 2020; Billing et al., 2021; Commission of the European Communities, 2000; OECD, 2001; Rees, 2010). Individual as well as institutional needs must be balanced when it comes to work-based learning, to ensure companies' investments into employee learning "pay off" (in terms of additional benefits these employees generate for their employers).

2.1.2.2 Economics of education

As per the prevailing paradigm of the so-called economics of education, the *human capital theory* (Becker, 1962), education and training are investments which improve individuals' productivity by enhancing the quality of employees, i.e., the knowledge, skills, competencies and attributes these individuals possess (Hanushek & Wößmann, 2020; OECD, 2007; Schönherr & Tiberius, 2014). Investing into one's employees' education and health is considered as important as investing into a company's physical assets, by which individuals are being commodified through skill acquisition (Klees, 2016). As each organisation's human capital differs, it represents an uninterchangeable competitive advantage with the potential to significantly improve the organisation's performance (Acedo et al., 2006; Crook et al., 2011).

For employers, but also the wider economy, return of this investment results in economic growth (European Commission, 2018; Rees, 2010; Tannenbaum, 1997) as an investment in employee skills via learning interventions, amongst other things according to Hanushek and Wößmann (2010), can

1. "*increase the human capital inherent in the labor force, which increases labor productivity and thus transitional growth toward a higher equilibrium level of output […],*
2. *increase the innovative capacity of the economy, and the new knowledge on new technologies, products, and processes promotes growth (as in theories of endogenous growth), […]*
3. *facilitate the diffusion and transmission of knowledge needed to understand and process new information and to successfully implement new technologies devised by others, which again promotes economic growth […]*" (p. 245).

From the employees' point of view, a positive correlation between the level of education and individual monetary income is affirmed, arguing that higher productivity resulting from better education and higher knowledge levels justify higher incomes (Kugler et al., 2017; Vargas, 2017). As such, private investment into education and hence accumulation of human capital is expected to eventually yield higher incomes and to offset all incurred costs (Carneiro et al., 2010).

Based on these assumptions one could believe that the more time people spend in (formal) education, the better their individual incomes as well as the economy. Even though strong evidence exists that the more educated people are, the higher their wages (Blanden & Machin, 2010), cross-country comparisons on the wage returns to education differ (Hanushek & Wößmann, 2010), thus education can only be seen as one factor influencing economic growth (Marquez-Ramos & Mourelle, 2019). Further, not all (formal) learning is good learning, and despite the fact that "*the majority of the macroeconomic literature on economic returns to education employs measures of the quantity of schooling*" (Hanushek & Wößmann, 2010, p. 245) it is also about the quality of educational interventions, i.e., the measured cognitive skill outcomes of these (Tannenbaum, 1997).

Inevitably, work-based educational interventions can only be an influencing factor to enhance competitive advantage when they enable individuals to transfer the skills and knowledge gained during the intervention of their job. However, it is assumed that about half of these skills and knowledge is already forgotten one day after the intervention unless reinforced during work (Blanchard, 2013). Thus, employers as well as training designers must ensure that transfer of what was learnt during the formal work-based learning intervention actively takes place in order to enlarge economic advantages from education (Blume et al., 2010; Kirkpatrick, 1967). For this Zahidi and associates (2020), state that 66 percent of businesses assume that proper up- and reskilling programmes would lead to a return of investment within one year.

2.1.2.3 Work-based learning in Germany

In their analysis, Brandi and Iannone (2017) concluded that whilst some countries have firmly established the concept of lifelong learning, others have not. The majority of businesses in countries such as Germany, Ireland, Slovenia and the United Kingdom are granting their employees time off for further education measures and personal development (Brandi & Iannone, 2017). Given that and to verify the above-mentioned assumption and theories, a review of the real-life effects of these is explained in the following, taking work-based learning activities in Germany between 2016 to 2019 as an example.

Overall, it was recorded that in 2019, 87.9 percent of German companies offered work-based learning and training to continuously educate their employees, with the topic of digitalisation as a main driver and digitalised businesses as main investors (Seyda & Placke, 2020). In doing so, businesses spent an average of EUR 1,293 per employee – split into direct costs of EUR 629, including expenses for external and internal trainers and lecturers, participant fees, catering and travel costs, costs for media and teaching materials as well as room and equipment costs; and indirect costs of EUR 608, including the paid working time used for the learning interventions (Seyda & Placke, 2020). This is equivalent to an overall economic investment volume of EUR 41.3 billion and increased by 23 percent compared to 2016 (Seyda & Placke, 2020) and further accounted for 9.4 percent of the German gross domestic product (Statistisches Bundesamt, 2021). 89.2 percent of these learning and training interventions took part during working hours and were split across four formats: seminars and courses, information events, in/non-formal learning during work, and self-directed learning. Already before the COVID-19 pandemic, the last two formats grew the most, as shown in Figure 7.

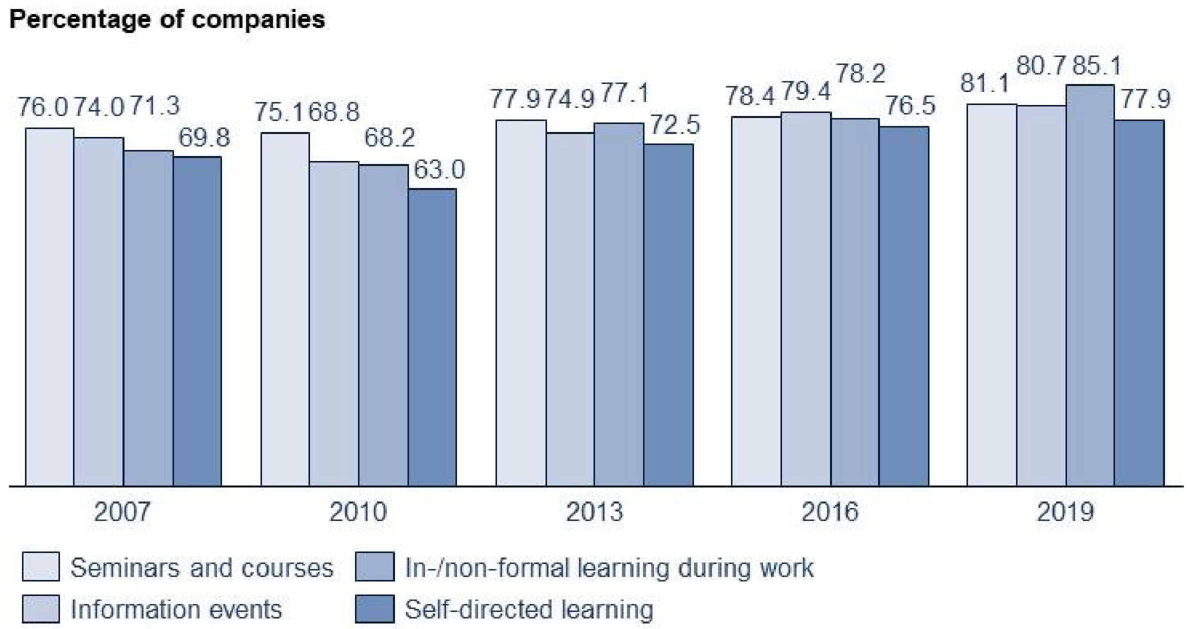


Figure 7 – Development of delivery formats of in-company continuing education in Germany, adapted from Seyda & Placke, 2020, p. 107

Even though digitalisation of the German learning and training market was accelerated during the pandemic with an increase from 35 percent pre-COVID to 54 percent post-COVID (Kirchherr et al., 2020), the largest share of work-based learning hours is still accounted for classical interventions such as seminars and courses (Seyda & Placke, 2020), which are rather designed to focus on learning to know and not learning to do (Rubenson, 2019). While about 84 percent of organisations state that work-based learning offerings are part of their board agenda, only half of these have a clear picture on which skills of employees are required in the future, and no proper learning strategies are in place (Kirchherr et al., 2020). Further, didactic formats are not adapted to the necessary qualifications and existing structures, no differentiation is made between online, offline and hybrid learning formats, and limited focus is placed on actual learning outcomes, i.e., knowledge retention. On top of that, neither learning contents nor learning successes are recorded and the increase in competence of employees is hardly assessed (Kirchherr et al., 2020). As a result, no systematic data evaluation of the learning outcomes, i.e., the return on learning – in form of knowledge retention and eventual application of the learnt material – of the employees can take place. Thus, employers cannot track if employees actually perform differently in their roles or if they could apply what was learnt in new roles (Blume et al., 2010; Kirchherr et al., 2020). Hence economics of work-based learning, while universally acknowledged, is so far not measured nor actively addressed.

2.1.3 Summary

The pressure for individuals to learn has increased considerably due to the escalating speed and amount of change taking place with regards to technical and organisational innovation cycles and society (Zuber, 2014). Every single individual can play an important role in this process by engaging in lifelong learning and thereby achieving both individual but also broader economic and social targets, such as employability (Billett, 2014). Most learning takes place in informal and non-formal settings (Milligan et al., 2014; Rubenson, 2019) even though not everything related to new forms of work and new roles due to up- or reskilling situations can be learnt through these settings (Billett, 2014). Thus, investments into formal work-based learning arrangements must be made (Tynjälä, 2008).

Work-based trainings are seen as important and inevitable for both individuals and organisations. Individuals' returns on learning are employability and higher wages while organisations' returns on learning are, among others, higher productivity and innovation (European Commission, 2018; Hanushek & Wößmann, 2010; Kugler et al., 2017; Rees, 2010; Tynjälä, 2008). For the case of Germany, it can be said that even though the investments made by organisations to enhance their human capital has increased by nearly a quarter from 2016 until 2019, little clarity existed on how learning interventions should be properly designed and assessed to maximise return on learning for both employees and employers.

Return on learning is seen as maximised permanent transfer of the learnt. Hence long-term knowledge retention needs to be aimed at enabling employees to grow personally as well as organisations to keep businesses competitive (Billett, 2018; Billing et al., 2021; Blume et al., 2010; Kirkpatrick, 1967; Vargas, 2017). To help organisations enhance their return on learning especially with the increase of digital learning offerings, it needs to be understood how learning can be designed more efficiently, with the research at hand putting focus on how knowledge retention of work-based learning can be enhanced.

2.2 Spaced learning – A strategy to enhance knowledge retention

To dive deeper into one very promising strategy for enhancing knowledge retention, the next chapters will review the existing state of research on *spaced learning* regarding their relevance for the outlined research question and objectives. The first chapter provides a basic understanding of the *spaced learning effect* (chapter 2.2.1), which is postulated to be an effective means to enhance long-term knowledge retention by temporarily distributing learning sessions over time. The second chapter critically summarises the findings and study setups of *previous spaced learning research* (chapter 2.2.2). Chapter 2.2.3 then investigates two of the most important moderating influences that affect the spaced learning effect, namely the comparison of two different *spacing schedules* (expanding and equal) and the *interaction between inter-session and retention intervals*. To explain these moderating influences as well as underpinning that the spaced learning effect exists, previous research proposes several different *theoretical accounts* (descriptive and computational), mainly deriving from neuroscientific research. These theories will be discussed in chapter 2.2.4. An expansion on spaced learning is the *retrieval practice*, which uses delayed tests to increase long-term knowledge retention (chapter 2.2.5). However, the spaced learning effect is detrimentally affected by the gap between learner's perceived and actual knowledge, which often leads to ruling out effective learning strategies such as spaced learning. In chapter 2.2.6, major research findings in the field of learners' *metacognition* are briefly reviewed.

2.2.1 Spaced learning effect

The *spaced learning effect* is among the most studied, most reliable and meaningful phenomena in the field of human memory and is at times referred to as distributed practice (Carpenter et al., 2012; Cepeda et al., 2006; Delaney et al., 2010; Dempster, 1989; Gerbier & Toppino, 2015; Vlach et al., 2019). The spaced learning effect refers to a powerful long-term memory advantage which occurs through the deliberate insertion of time intervals between repeated learning sessions compared to immediate repetition. Research has confirmed this finding for more than 100 years, especially in the field of language learning and verbatim recall. By engaging in the instructional method of spaced practice, overall learning is slowed, but retention is enhanced (for reviews see Cepeda et al., 2006; Delaney et al., 2010; Latimier et al., 2021; Walsh et al., 2018b).

Already in 1885, Ebbinghaus (1885/1962) conducted seminal experimental work on memorising a nonsense 12-syllable series using two learning schedules with himself as the subject. After 68 immediate consecutive repetitions on one day and another seven repetitions the day after, Ebbinghaus was able to recite the syllable series errorless. Yet, Ebbinghaus achieved the same result by distributing 38 repetitions over three days (Dempster, 1989). He concluded that learning and recalling depend on how often someone was exposed to the material (Schunk, 2012). Although several proposed reasons as to why spaced learning outcomes are superior to massed learning outcomes exist (Balota et al., 2011) and researchers from various disciplines, including cognitive psychology, applied psychology, neurosciences and pedagogy, continue evaluating the variables leading to the spaced learning effect, findings on it are of great importance to gain theoretical insight on human memory (Mulligan & Peterson, 2014).

All in all, research argues that stimuli which are relearnt and reviewed multiple times distributed across time are better remembered in the long-term than those that are *massed*, i.e., repeatedly studied without interruptions, which in turn is also of enormous relevance for educational practice (Bjork, 1979; Carpenter et al., 2012; Dempster, 1989; Dunlosky et al., 2013; Gerbier & Toppino, 2015; Greene, 2008; Kang, 2016; Kornell & Bjork, 2008; Sobel et al., 2011).

2.2.2 Previous research

Since Ebbinghaus' work, more than 1000 published researches have confirmed his findings accentuating the robustness of the instructional method "*spaced learning*" (Bird, 2010; Cepeda et al., 2008; Delaney et al., 2010; Dempster, 1989; Melton, 1970; Vlach et al., 2019). Summarising, the three main findings of spaced learning research are (Walsh et al., 2018b):

1. Initial acquisition is slowed
2. Retention is enhanced
3. Increasing the number of (temporally distributed) repetition sessions enhances retention to a point after which retention decreases again

The basic design of a spaced learning study is shown in Figure 8: Two succeeding learning sessions containing the same to-be-learnt information are separated through a manipulated time gap, the so-called *inter-session interval (ISI)*. The time between the multiple learning sessions can be described as either item-based, i.e., how many items are intervening, or time-based, i.e., how much time passes between the learning sessions (Latimier et al., 2021). The second learning session is followed by another manipulated time gap, the *retention interval (RI)*, which is followed by a test capturing learners' memory performance (Carpenter et al., 2012; Cepeda et al., 2006; Vlach et al., 2019).

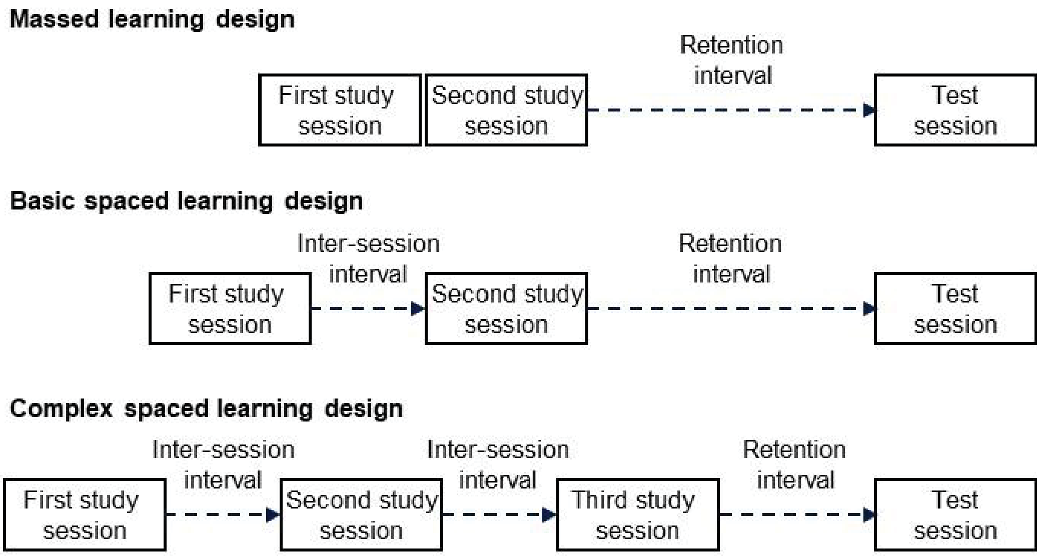


Figure 8 – Examples of basic and complex spaced learning research design, adapted from Wiseheart et al., 2019, p. 551, 555

A substantial part of the findings on spaced learning were derived from relatively brief laboratory studies, which allow for strict control of study-relevant variables (Balota et al., 2011; Bird, 2010; Carpenter et al., 2012; Cepeda et al., 2006). Being highly replicable across populations, domains, learning tasks and species, the spaced learning effect can be found across all human age groups from infancy and school children (e.g., Carpenter et al., 2009; Rea & Modigliani, 1987; Toppino, 1991; Toppino & DiGeorge, 1984; Vlach et al., 2008) to adulthood and older adulthood (e.g., Benjamin & Craik, 2001; Cepeda et al., 2006; Kornell et al., 2010; Seabrook et al., 2005; Yan et al., 2016), as well as for individuals with memory impairments (e.g., Balota et al., 2006; Fritz et al., 2007; Kalenberg, 2017; Wilson & Evans, 1996). The majority of studies within this human memory research focused on verbal or trivia factual learning material, such as word lists (Carpenter et al., 2012; Kapler et al., 2015) and associating names and faces (Landauer & Bjork, 1978). These studies asked participants to only retrieve these simple verbal facts from memory. Yet, fact learning alone does not equip any learner to fully apply the learnt material in a real-life setting as this would also require explaining, evaluating, analysing or even transferring the knowledge to new situations (Bloom, 1956; Foot-Seymour et al., 2019).

Studies exploring the effect of spaced learning on conceptually more complex higher-level skill learning such as mathematical and science concept learning (e.g., Kang & Pashler, 2012; Kornell & Bjork, 2008; Rohrer, 2009; Rohrer & Taylor, 2006/2007; Vlach et al., 2008), inductive category learning and making complex judgments (Kang & Pashler, 2012; Kornell & Bjork, 2008) are limited. Results are mixed as some studies showed that the spaced learning effect significantly dropped as the complexity of a task to-be-learnt increased (e.g., Bird, 2010; Donovan & Radosewich, 1999) whereas others have shown that the spaced learning effect can clearly be extended to at least moderately complex tasks as well as to higher-level thinking (Foot-Seymour et al., 2019; Kapler et al., 2015; Rohrer & Taylor, 2006). It was further proven that spacing enhances learning of perceptual and coordinated motor tasks (Baddeley & Longman, 1978; Carpenter et al., 2012; Cepeda et al., 2006; Dempster 1996). Also demonstrated for other species such as drosophila (e.g., Tully et al., 1994) and bees (e.g., Deisig et al., 2007), the view that the spaced learning effect resembles a "*fundamental principle of the memory system, shaped by evolution*" (Gerbier & Toppino, 2015, p. 50) was reinforced.

Over the last decade, research on the spaced learning effect in educational settings evolved (Wiseheart et al., 2019). Already in 1988, Dempster called for a wide-spread application of spaced learning in educational settings and classroom practice. Yet even in 2012, Carpenter and team were not able to confirm a systematic implementation of the spaced learning strategy in education programmes, although considerable positive enhancements of long-term knowledge retention in real-life educational settings were found (Carpenter et al., 2012; Kapler et al., 2015; Karpicke et al., 2016; Larsen, 2018; Mettler et al., 2016; Seabrook et al., 2005; Sobel et al., 2011). It was hypothesised that either research on spaced learning did not define a "*clear set of recommendations for how it can be used in everyday instruction*" (Carpenter et al., 2012, p. 375) or that the learning effects of translating the benefits of spaced learning from laboratories into complex educational settings might be too weak as students are not only required to retrieve information from memory but "*must manipulate and apply the remembered* *information to answer more complex, higher-level questions*" (Kapler et al., 2015, p. 38). The latter could be disproved from studies using real curricula (Wiseheart et al., 2019), thus, it appears as if the spacing effect remains "*a case study in the failure to apply the results of psychological research*" (Dempster, 1988, p. 627).

Based on a meta-analysis involving more than 400 spaced learning studies in the domain of verbal learning, Cepeda and team (2006, in Wiseheart et al., 2019) found that about 75 percent showed the positive spacing effect with a large effect size (d = 0.85). Another 15 percent of studies showed an advantage for the massing condition and 10 percent did not reveal any difference of massing versus spacing conditions. Keeping in mind the vast amount of existing literature on spaced learning, it is not surprising that they do not all report the same strong effects and some even showed mixed or no effects (Wiseheart et al., 2019).

An encompassing insight into existing spaced learning experimental research to date is shown in the following Table 3. In this, a distinction has been made on the following criteria: whether the research was laboratory-based, whether it was management related, timing of inter-session and retention intervals, sample size, and effect size. Not all studies led to the same results, partially due to different confounds in different experimental designs, which are also pointed out in the overview table. However, it is important to note that to the best of our knowledge, only one study applying spaced learning in a real-life, managerial relevant setting has been identified. Yet this study did not consider the ISI-RI relationship (more on this in the following chapter 2.2.3).

Table 3 – Literature table on spaced learning, adapted from Wiseheart et al., 2019, pp. 558

Authors	Year	Journal	Title	Brief study description	Laboratory based?/ Management related?	Subject/ skill	Time scale of ISI/RI	Nr. of study sessions	Sample size	Effect size	Notes
Baddeley, A. D. & Longman, D. J. A.	1978	Ergonomics	The Influence of Length and Frequency of Training Session on the Rate of Learning to Type	Postal workers learning keyboard skills on mail sorting machines	No/ no	Vocation/ motor skills	Hours/ months	60	72	Weak	-
Bahrick, H. P.	1979	Journal of Experimental Psychology: General	Maintenance of Knowledge: Questions About Memory We Forgot to Ask; Experiment 1	Undergraduates learnt 50 English-Spanish word pairs	Yes/ no	Language/ verbal learning	Days and months/ months	6	30	Weak	Study contains a potential confound
Bahrick, H. P. & Phelphs, E.	1987	Journal of Experimental Psychology: Learning, Memory, and Cognition	Retention of Spanish vocabulary over 8 years	Subjects learnt and relearnt 50 English-Spanish word pairs and were tested again after 8 years	No/ no	Language/ verbal learning	Seconds, days and months/ months and years	multiple	35	Weak	Study contains a potential confound
Bahrick, H. P., Bahrick, L. E., Bahrick, A. S. & Bahrick, P. E.	1993	American Psychological Society	Maintenance of Foreign Language Vocabulary and the Spacing Effect	Subjects' acquisition and retention of foreign language vocabulary were examined over several years	No/ no	Language/ verbal learning	Days and months/ years	13/26	4	Weak	Study contains a potential confound
Balota, D. A., Duchek, J. M., Sergent-Marshall, S. D. & Roediger, H. L.	2006	Psychology and Aging	Does expanded retrieval produce benefits over equal-interval spacing? Explorations of spacing effects in healthy aging and early stage Alzheimer's disease	Subjects learnt word pairs	Yes/ no	Medical treatment/ intellectual skills	Seconds/ minutes	5	103	Strong	-
Benjamin, A. S. & Craig, F. I. M.	2001	Memory & Cognition	Parallel effects of aging and time pressure on memory for source: Evidence from the spacing effect; Experiment 1	Students learnt word lists	Yes/ no	Language/ verbal learning	Seconds/ minutes	multiple	35	Weak	Study contains a potential confound

Table 3 continued – Literature table on spaced learning

Authors	Year Year	Journal	Title	Brief study description	Laboratory based?/ Management related?	Subject/ skill	Time scale of ISI/RI	Nr. of study sessions	Sample size	Effect size	Notes
Bird, S.	2010	Applied Psycho-linguistics	Effects of distributed practice on the acquisition of second language English syntax	Undergraduates corrected verb errors in texts	No/ no	Language/ intellectual skills	Days/ weeks and months	5	38	Strong	-
Braun, K. & Rubin, D. C.	1998	Memory	The Spacing Effect Depends on an Encoding Deficit, Retrieval, and Time in Working Memory: Evidence from Once-presented Words; Experiment 1	Undergraduates learnt word lists	Yes/ no	-	Seconds/ minutes	Multiple	48	Weak	Study contains a potential confound
Carpenter, S. K. & DeLosh, E. L.	2005	Applied Cognitive Psychology	Application of the Testing and Spacing Effects to Name Learning; Experiment 3	Undergraduates learnt face-name stimulus-response pairs	Yes/ no	Language/ verbal learning	Seconds/ minutes	3	69	Strong	-
Carpenter, S. K., Pashler, H. & Cepeda, N. J.	2009	Applied Cognitive Psychology	Using Tests to Enhance 8th Grade Students' Retention of U.S. History Facts	8th grade students learnt US history facts	No/ no	History/ verbal learning	Weeks/ months	2	75	Strong	-
Cepeda, N. J., Coburn, N., Rohrer, D., Wixted, J. T., Mozer, M. C. & Pashler, H.	2009	Experimental Psychology	Optimizing Distributed Practice Theoretical Analysis and Practical Implications; Experiment 2	Undergraduates learnt mot-well-known facts and names of unfamiliar visually presented objects	No/ no	Language/ intellectual skills	Minutes and months/ months	3	161	Mode-rate/ strong	-
Cepeda, N.J., Pashler, H., Vul, E., Wixted, J.T. & Rohrer, D.	2008	Psychological Science	Spacing Effects in Learning: A Temporal Ridgeline of Optimal Retention	Subjects learnt obscure but true trivia facts	Yes/ no	Language/ intellectual skills	Days and months/ days and months and year	3	1354	Strong	-
Delaney, P. F., Spirgel, A. S. & Toppino, T. C.	2012	Memory and Cognition	A deeper analysis of the spacing effect after "deep" encoding; Experiment 2	Undergraduates created "story links" between items from word lists	Yes/ no	Language/ intellectual skills	Seconds/ minutes	2	30	Strong	-

Table 3 continued – Literature table on spaced learning

Authors	Year	Journal	Title	Brief study description	Laboratory based?/ Management related?	Subject/ skill	Time scale of ISI/RI	Nr. of study sessions	Sample size	Effect size	Notes
Donovan, J. J. & Radosevich, D. J.	1999	Journal of Applied Psychology	A Meta-Analytic Review of the Distribution of the Practice Effect: Now You See It, Now You Don't	112 effect sizes were examined with regards to the relationship between conditions of massed and spaced learning practice performance	N/a	Literature review	N/a	N/a	63	Mode-rate	-
Foot-Seymour, V., Foot, J. & Wiseheart, M.	2019	Applied Cognitive Psychology	Judging credibility: Can spaced lessons help students think more critically online?	4th to 6th grade students learnt to judge the credibility of websites	No/ no	Critical thinking/ intellectual skills	Days/ weeks	3	388	Mode-rate	-
Grote, M. G.	1995	The Ohio Journal of Science	The Effect of Massed Versus Spaced Practice on Retention and Problem-Solving in High School Physics	High school students learnt physics concepts	Yes/ no	Science/ intellectual skills	Days/ weeks and months	Multiple	46	Strong	Study contains a potential confound
Kalenberg, K.	2017	Journal of Undergraduate Research at Minnesota State University	Spaced and Expanded Practice: An Investigation of Methods to Enhance Retention	3rd and 4th grade students learnt unknown math vocabulary words and definitions	Yes/ no	Language/ verbal learning	Days and weeks/ month	4	42	Strong	-
Kang, S. H. K. & Pashler, H.	2012	Applied Cognitive Psychology	Learning Painting Styles: Spacing is Advantageous when it Promotes Discriminative Contrast; Experiment 2	Undergraduates learnt to identify the artists of previously unseen paintings	Yes/ no	Arts/ intellectual skills	Seconds / minutes	10	90	Mode-rate	-
Kapler, I. V., Weston, T. & Wiseheart, M.	2015	Learning and Instruction	Spacing in a simulated undergraduate classroom: Long-term benefits for factual and higher-level learning	Undergraduate students learnt meteorology concepts	Yes/ no	Science/ intellectual skills	Days/ months	2	169	Strong	-

Table 3 continued – Literature table on spaced learning

Authors	Year	Journal	Title	Brief study description	Laboratory based?/ Management related?	Subject/ skill	Time scale of ISI/RI	Nr. of study sessions	Sample size	Effect size	Notes
Karpicke, J. D., Blunt, J. R. & Smith, M. A.	2016	Frontiers in Psychology	Retrieval-Based Learning: Positive Effects of Retrieval Practice in Elementary School Children	4th grade students learnt word lists	Yes/ no	Language/ verbal learning	Seconds and minutes/ minutes	2	88	Strong	-
Kondratjew, H. & Kahrens, M.	2018	Journal of Work-Applied Management	Leveraging experiential learning training through spaced learning	Students learnt hands-on lean warehouse logistics and processes	No/ yes	Vocation/ intellectual skills	Months/ months	2	14	Strong	Study contains a potential confound
Kornell, N. & Bjork, R. A.	2008	Psychological Science	Learning Concepts and Categories - Is Spacing the "Enemy of Induction"?; Experiments 1a ; 1b	Undergraduates studies multiple paintings by different artists	Yes/ No	Arts/ intellectual skills	Seconds/ seconds	Multiple	120/72	Strong	-
Kornell, N., Castel, A. D., Eich, T. S. & Bjork, R. A	2010	Psychology and Aging	Spacing as the Friend of Both Memory and Induction in Young and Older Adults	Subjects learnt to differentiate different artists based on their paintings	Yes/ No	Arts/ intellectual skills	Seconds/ seconds	Multiple	112	Strong	-
Mettler, E., Massey, C. M. & Kellmann, P. J.	2016	Journal of Experimental Psychology: General	A Comparison of Adaptive and Fixed Schedules of Practice; Experiment 2	Undergraduates learnt 24 country names and locations on a map of Africa	Yes/ No	Geography/ intellectual skills	Minutes/ weeks	2	48	Strong	-
Rea, C. P. & Modigliani, V.	1987	Memory & Cognition	The spacing effect in 4- to 9-year-old children; Experiment 2	Children of various age groups (pre-school, kindergarten, first-grade, and third-grade children) learnt words from word and picture lists	Yes/ No	Language / Intellectual skills	Seconds/ minutes	4	96	Strong	-
Reynolds, J. H. & Glaser, R.	1964	Journal of Educational Psychology	Effects of repetition and spaced review upon retention of a complex learning task; Experiment 1	Junior high school students learnt biology concepts and processes	Yes/ No	Science/ intellectual skills	Days/ days and weeks	Multiple	75	Strong	-

Table 3 continued – Literature table on spaced learning

Authors	Year	Journal	Title	Brief study description	Laboratory based?/ Management related?	Subject/ skill	Time scale of ISI/RI	Nr. of study sessions	Sample size	Effect size	Notes
Rohrer, D. & Taylor, K.	2006	Applied Cognitive Psychology	The Effects of Overlearning and Distributed Practice on the Retention of Mathematics Knowledge; Experiment 1	Undergraduates solved permutation problems	Yes/ no	Mathematics/ intellectual skills	1 week/ 1 and 4 weeks	3	116	Strong	-
Rohrer, D. & Taylor, K.	2007	Instructional Science	The shuffling of mathematics problems improves with learning; Experiment 1	Undergraduates solved permutation problems	Yes/ no	Mathematics/ intellectual skills	1 week/ 1 week	3	66	Strong	-
Seabrook, R., Brown, G. A. & Solity, J. E.	2005	Applied Cognitive Psychology	Distributed and Massed Practice: From Laboratory to Classroom; Experiment 3	Elementary school students learnt phonics	Yes/ No	Language/ verbal learning	Mins/ weeks	3	34	Strong	-
Smith, S. M. & Rothkopf, E. Z.	1984	Cognition and Instruction	Contextual Enrichment and Distribution of Practice in the Classroom	Undergraduates learnt statistics	Yes/ No	Mathematics/ intellectual skills	Day/ days	4	100	Strong	Study contains a potential confound
Sobel, H. S., Cepeda, N. J. & Kapler, I. V.	2011	Applied Cognitive Psychology	Spacing Effects in Real-World Classroom Vocabulary Learning	5th grade students learnt unfamiliar English words	No/ No	Language/ verbal learning	Weeks/ months	2	39	Mode-rate	-
Verkoeijen, P. P. J. L., Rikkers, R. M. J. P. & Özsoy, B.	2008	Applied Cognitive Psychology	Distributed Rereading can Hurt the Spacing Effect in Text Memory	Undergraduates read and comprehended expository narratives	Yes/ No	Language/ verbal learning	Seconds, days and weeks/ days	2	61	Mode-rate	-
Vlach, H. A., Sandhofer, C. M. & Kornell, N.	2008	Cognition	The spacing effect in children's memory and category induction	Three-year-old children learnt new categories and concepts using arts and craft supplies and objects from hardware stores	Yes/ No	Language/ intellectual skills	Seconds/ minutes	Multiple	36	Strong	-

Research in the field of spaced learning and the spaced learning effect continues to assess several moderating influences (Mulligan & Peterson, 2014), mainly the schedule and timing of repetitions, which will be discussed subsequently.

2.2.3 Moderating influences of the spaced learning effect

To practically achieve the best long-term knowledge retention, past and ongoing spaced learning research is concerned with two important questions: first, are there any specific spacing schedules to follow and second, what, if any, relationship exists between RI and ISI (e.g., Carpenter et al., 2012; Cepeda et al., 2008; Karpicke & Bauernschmidt, 2011; Lindsey et al., 2009). Pimsleur (1967) was one of the first to sketch out the impact spaced learning interventions have on a learner's "probability of correct response" of any word to-be-learnt: just after having learnt a single foreign-language word, probability of correct response is almost 100 percent, but as time decays and other information intervenes, the probability of recalling the word correctly declines rapidly and could eventually be fully forgotten not too far after it was initially learnt. Supposing that the word is repeated during the learning session, forgetting is stopped for a moment and, through the repetition of the word, slowed going forward. As spacing interventions continue, the memory trace of the word is strengthened, and forgetting is slowed even further. Pimsleur (1967) postulated that the time between spacing interventions should become longer after every revision as he claimed that once a word was repeated and its memory trace was boosted, the learner will remember the word for longer and longer.

Thus, he was the first one to state that any spacing schedule is expanding exponentially. However, since then, much more research arose on first, how frequently spaced learning sessions shall occur, i.e., which schedule they shall follow, and second, in which sequence these should be organised, i.e., what determines an optimal gap between learning sessions. In the following part of this chapter, the research at hand aims to address these questions, starting with comparing the most studied spacing schedules.

2.2.3.1 Comparing two spacing schedules – the expanded condition versus the equal condition

Previous experimental research on the spaced learning effect included at least two succeeding learning sessions, temporarily distributed over time. When the research design included more than two learning sessions, the ISI is varied whereas the RI is a fixed lag following the last learning session (Lindsey et al., 2009).

Karpicke and Bauernschmidt (2011) shaped the terms *absolute* and *relative spacing*. Absolute spacing refers to the total number of learning sessions taking place between all tests, whereas relative spacing links to how the learning sessions are relatively distributed to one another (just as the ISI). In line with previous findings (e.g., Karpicke & Roediger, 2010), both researchers argued that the absolute spacing variable has more importance on long-term knowledge retention enhancement compared to how learning sessions and tests are temporarily distributed. However, further research has shown that it is also the ‘how’ which influences successful knowledge retrieval and retention (Cepeda et al., 2008). Following this, the ISI can be designed following either an *equal*, *expanding* or *contracting schedule* (Cepeda et al., 2006; Karpicke & Bauernschmidt, 2011; Landauer & Bjork, 1978; see Figure 9).

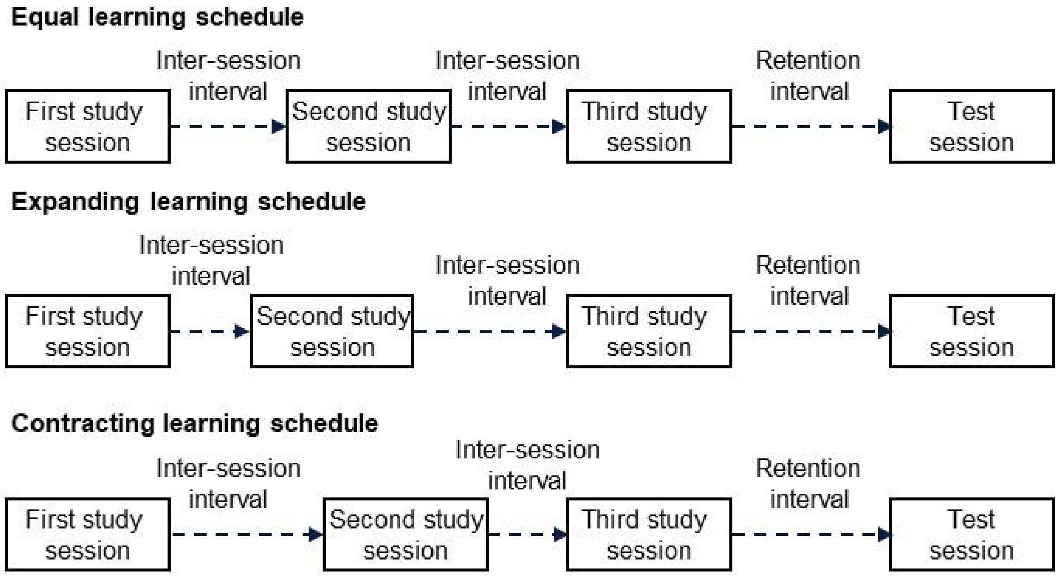


Figure 9 – Spaced learning research design with three different learning schedules, adapted from Wiseheart et al., 2019, p. 555

In an expanded spaced learning schedule, repetitive learning sessions are temporarily distributed at increasing intervals, whereas in an equal spaced learning schedule, repetitive learning sessions are consistently distributed over time (Logan & Balota, 2008). In contrast to the contracting spaced learning schedule, in which repetitive learning sessions are temporarily distributed at decreasing intervals, equal and expanding learning schedules have been frequent subjects of many spaced learning experiments (Latimier et al., 2021). Due to this and the finding that the equal and expanding spaced learning schedules appear to be more advantageous for longer RIs such as months or years and contracting schedules are rather beneficial for short RIs, (Cepeda et al., 2008; Küpper-Tetzel et al., 2014; Mozer et al., 2009; Wiseheart et al., 2019) this review focuses on the equal and expanding spacing schedules only.

In 1978, Landauer and Bjork were the first to experimentally compare the expanded and equal spacing schedule. Within two experiments, they were able to show the significant impact both equal and expanding spaced learning sessions had over massed learning and further noted a substantial but small effect of 7 percent of the expanded schedule over the equal one (Balota et al., 2006). Shaughnessy and Zechmeister (1992) followed in comparing both schedules and found an effect of 6 percent. At the time, it was hypothesised that "*expanding retrieval works because it introduces desirable difficulties during learning that improve later retention*" (Bjork, 1994/1999, in Karpicke & Roediger, 2007a, p. 705; see also Carpenter & DeLosh, 2005; Landauer & Bjork, 1978) as the increasingly longer ISIs in the expanding condition heighten difficulty of actual retrieval compared to the equally long ISI during the equal spacing condition. As long as retrieval success occurs during the expanding condition, memory is fostered (Balota et al., 2006).

Ever since, experiments with short RIs yielded mixed results with almost half of the experiments revealing an advantage of the expanding condition over the equal condition and the other half revealing no or the opposite advantage (e.g., Balota et al., 2006; Carpenter et al., 2012; Cull et al., 1996; Karpicke & Bauernschmidt, 2011; Karpicke & Roediger, 2007a, 2010; Logan & Balota, 2008; Spitzer, 1939; Toppino et al., 2018). Contrarily, experiments with long-term RIs showed strong effects of the equal spacing condition over the expanding spacing condition (Cull, 2000; Karpicke & Roediger, 2007a; Logan & Balota, 2008). These notions raised the assumption that it is the RI, i.e., when the learner would like to recall something, which determines the specific spacing schedule to use (Latimier et al., 2021). When comparing fictious learning curves of both conditions over the same time, one can then assume that even though there are differences of retrieval success in the learning acquisition phase, the learning curve of the expanding condition drops faster with expanding sessions and tests. Contrarily, the learning curve of the equal condition rises slower but steadily and thus, both curves are potentially at the same height at the final retrieval test (Carpenter & DeLosh, 2005; Logan & Balota, 2008). Yet, it could also be that the longer the RI becomes, the better retrieval gets in the equal spacing condition.

Overall, it is argued that the expanding spacing condition leads to initial improved performance during the learning phase which is said to be due to an easier retrieval process in the first place: having a first repetition relatively short after the initial learning evening resembles a massed learning session as the memory trace might still be active and the learner can easily recall the information. This benefit however is short-lived, decreases over time and might actually be completely lost in substantially delayed retrieval situations (Balota et al., 2011; Karpicke & Roediger, 2007a; Logan & Balota, 2008). This is because any expanding learning schedule has always one learning event less than an equally spaced schedule. Thus, overall variability of retrieval cues is lower than the equal condition and probability of retrieval temporarily distributed long-term is decreased (Balota et al., 2011). This finding is quite different to the assumption of Landauer and Bjork (1978), who argued that the expanded spacing schedule would be optimal for long-term memory retention.

Proponents of the equal spacing condition also argued that the desirable difficulty created within this schedule makes it more successful in long-term memory retention than the expanding schedule: delaying the first repetitive learning session involves greater effort in retrieving the information to-be-learnt from memory. This leads to more errors when tested, which is said to produce a great advantage for long-term memory, particularly when the learner receives feedback on errors thereby enhancing the learning experience (Karpicke & Roediger, 2007b; Roediger & Karpicke, 2006a). The learning acquisition phase during an equal spacing condition might be slower, contains more errors and therefore, harms accessibility of the to-be-learnt material. It, on the other hand, does allow for more encoding and sampling, as more neurologic processing is taking place and thus increases likelihood of later successful recall and better long-term memory is yielded (Balota et al., 2006; Cull et al., 1996; Logan & Balota, 2008).

There are two variables that can be attributed to the differences in the study results: First, some researchers did not compare the expanded with the equal spacing condition but rather confirmed superiority of spaced learning over massed learning (e.g., Rea & Modigliani, 1985; Spitzer, 1939). Second, most learning experiments in the field of spaced learning did not provide *corrective feedback* when errors occurred as research was instead concerned with the direct effects of spaced learning than with mediated ones (Karpicke & Bauernschmidt, 2011). However, providing feedback seems to have a remarkable impact on knowledge retention (Balota et al., 2006; Cepeda et al., 2006; Pashler et al., 2007).

Feedback on errors helps the learner see the gaps in their knowledge, overcome the metacognitive discrepancy of real and perceived knowledge as well as counteract forgetting (Black & William, 1998; Butler et al., 2007, 2008; Hattie & Timperley, 2007). Therefore, it does not matter if feedback is immediate or delayed (Corral et al., 2021; Pashler et al., 2007). Past research questioned whether one form of feedback is superior to the other and several different opinions were posited and contradictory results presented (e.g., Clariana et al., 2000; Corral et al., 2019; Kulik & Kulik, 1988; Mullet et al., 2014; Skinner, 1968). Butler and team (2013) however argued that elaborative (but at the same time not too dense) feedback which seeks explanations as to why a response is correct or incorrect helps the learner "*move from superficial factual knowledge to a more complex understanding of the concept*" (p. 291) in the learning process. In addition, and also supporting the research at hand, it was argued that when providing explanatory feedback temporarily distributed after a test has taken place, the delayed feedback can be considered spaced: the learner has two differing opportunities to engage with the learning content compared to immediate feedback for which no deep retrieval from memory needs to take place (Corral et al., 2021). Thus, learning might take place at a much higher pace when feedback is spaced out in time, relative to the immediate feedback and therefore might be ideal for long-term knowledge retention.

In accordance with the above, providing corrective feedback in any spacing condition showed a direct improvement during the acquisition phase as well as during retrieval for the final test (Balota et al., 2006; Cepeda et al., 2006; Cull et al., 1996). Yet, in a study by Cull (2000) which was replicated by Karpicke and Roediger (2007a) comparing the expanding spacing condition with the equal spacing condition, providing feedback within both, the researchers found a slightly higher advantage of the equal spacing condition over the expanding spacing condition during an immediate retrieval test and an even higher advantage during a retrieval test several days after the last learning session.

To summarise, when reviewing all research comparing the expanding spacing schedule to the equal spacing schedule, one can argue that there is no prevalent evidence that one schedule is per se superior to the other, but rather that the justification for one schedule over the others still depends on different aspects, "*such as the difficulty of the to-be-learned material, the type of review (rereading or retrieval practice), and the specific time frame*" (Storm, Bjork & Storm, 2010, in Kang, 2016, p. 14).

2.2.3.2 The inter-session interval and retention interval interaction

Driven by the fact that most students forget a lot of what has been taught in class and the finding that re-exposing them again to the already-forgotten material improves acquisition and accessibility and in turn prevents forgetting (Carpenter et al., 2009, 2012), a lot of complex research conducted in controlled experiments took place to find out the optimal *inter-session interval* (ISI) (the passing of time between two learning sessions) and *retention interval* (RI) (the time between the last learning session and final test) as guidance for universal classroom application (Bird, 2010; Mozer et al., 2009; Weinstein et al., 2018). Assuming that spaced learning has a positive impact on lifelong learning, it would be logical to say that longer ISI and RI enhance learning even better than shorter ones (Carpenter et al., 2012).

Yet, most spaced learning research had been conducted with brief single-ISI designs and short RIs ranging from seconds to minutes only (Cepeda et al., 2006, 2008, 2009). Also, in most of these experiments, memory was assessed via single recall tests, which "*only provide (...) information about whether the accessibility of an item in memory is above or below the threshold for successful retrieval*" (Nelson, 1985 in Walsh et al., 2018b, pp. 1326). In their meta-analysis on more than 400 spaced learning studies in verbal recall, Cepeda and team (2006) noted that only thirteen studies examined RIs as long as a day or even longer than a week, with most containing methodological flaws. Nevertheless, these studies showed a much better recall performance than those with a shorter ISI (Cepeda et al., 2006). Over short time scales very little is forgotten and hence, relevance of these experiments for educational practice aiming at long-term learning and application of the to-be-learnt material seems limited (Walsh et al., 2018b).

One of the first studies to examine more educationally relevant RIs was the one by Bahrick and team (1993) who examined learning acquisition and retention of Spanish vocabulary over a time period of nine years. Four subjects had to learn and re-learn 300 words until they knew them by heart, which were assigned to different spacing conditions (either 13 or 26 learning sessions took place at either 14 days, 28 days, or 56 days ISIs). Findings of this longitudinal study revealed that even though learning acquisition was slowed at highest ISI (56 days), significant higher retention was achieved. Cepeda and team (2008) claimed that this study "*appear(s) to suggest that over these long intervals, spacing effects may be monotonic, rather than showing an inverted-U shape, as found in shorter-term studies*" (p. 1096). On closer observation, Bahrick and team (1993) as many other researchers (e.g., Lapkin et al., 1998; Lightbown & Spada, 1994) did not control the ratio of ISI and RI (Bird, 2010). Additionally, having learners learn to a set performance criterion, i.e., having as many repetition in one learning sessions as needed to achieve perfect retrieval as done by the above mentioned study by Bahrick and team (1993) but also previously Bahrick and Phelphs (1987) and later Mozer and team (2009), no conclusions can be drawn with regards to the 'real' spaced learning effect for which the design of learning sessions should be fixed (Bird, 2010; Cepeda et al., 2008).

Notwithstanding the number of studies which proved long ISIs and RIs as robustly beneficial, several others exist which contradict this finding (Donovan & Radosevich, 1999; Robinson, 1921; Toppino & Gracen, 1985; Verkoeijen et al., 2008). Carpenter and team (2012) postulated that "*one potential danger of waiting too long before reviewing information is that students may forget much of what they have learned previously, and this forgetting may offset any benefits that would have occurred due to spacing*" (p. 373). Overall conclusion of previous research was that there is not a single schedule to follow to enhance memory retention, one rather must know until when it is expected to recall the information to-be-learnt.

Both intervals interact with each other with a power-law relationship as the length of the ISI depends heavily on the length of the RI: the longer the RI, i.e., the longer a learner would like to remember something, the longer the ISI between learning sessions (Balota et al., 2011; Cepeda et al., 2008; Glenberg, 1976; Rohrer & Pashler, 2007). This is also referred to as the "*temporal ridgeline of optimal retention*" (Cepeda et al., 2008, p. 1095). Consequently, one cannot determine an ISI if the RI is not clear. Hence, if a learner would like to remember information for years, ISIs of months seem highly applicable, whereas if the learner would like to remember information for weeks only, days would seem more appropriate (Balota et al., 2011; Cepeda et al., 2008, 2009; Dunlosky et al., 2013; Gerbier & Toppino, 2015; Kang, 2016; Pashler et al., 2007; Smolen et al., 2016).

As previous research has shown, the positive effect of long ISIs is not monotonically compared to the RI, i.e., merely lengthening the time between learning sessions does not automatically lead to long-term memory success (Bird, 2010). On the contrary, the spaced learning effect is a nonmonotonic one or as Cepeda and team (2009) framed it: "*[...] optimal [inter-session interval] increases as a function of [retention interval]. Second, the ratio of optimal [inter-session interval] to [retention interval] appears to decrease as a function of [the retention interval]*" (p. 243). This relationship plotted in a graph shows the so-called spacing function which always follows a similar curve (Figure 10): after the first learning session, high recall accuracy is achieved. Recall accuracy increases at first with ever-increasing ISIs but falls again gradually once a certain ISI is exceeded. On the other hand, having one massed learning session only would only lead to a rapid decrease of recall accuracy. Plotting this recall decay in a function would resemble the forgetting function (Mozer et al., 2009).

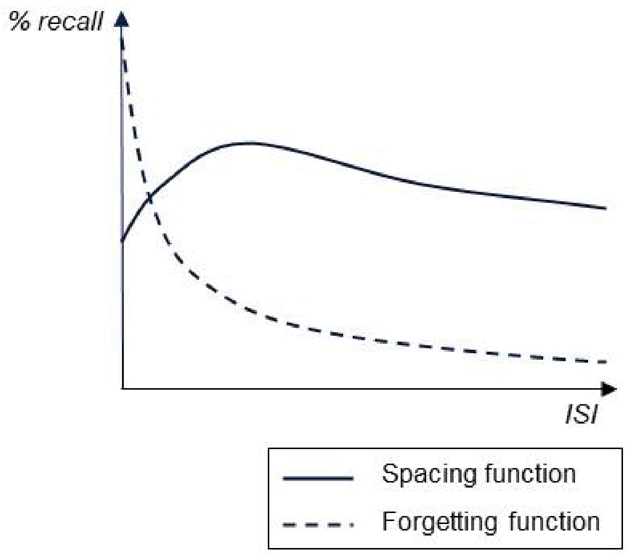


Figure 10 – Spacing function, adapted from Mozer et al., 2009, p. 1322

In a universally acknowledged paper, Cepeda and team (2008) examined memory recall performance of 1350 participants taking part in spaced learning over a duration of up to one year. Four different RIs with different ISI conditions were tested and it was found that the interpolated ISIs for the respective RIs were the following (Figure 11):

- RI: 7 days – ISI: 1 day (43 percent of RI)
- RI: 35 days – ISI: 11 days (23 percent of RI)
- RI: 70 days – ISI: 21 days (17 percent of RI)
- RI: 350 days – ISI: 27 days (8 percent of RI)

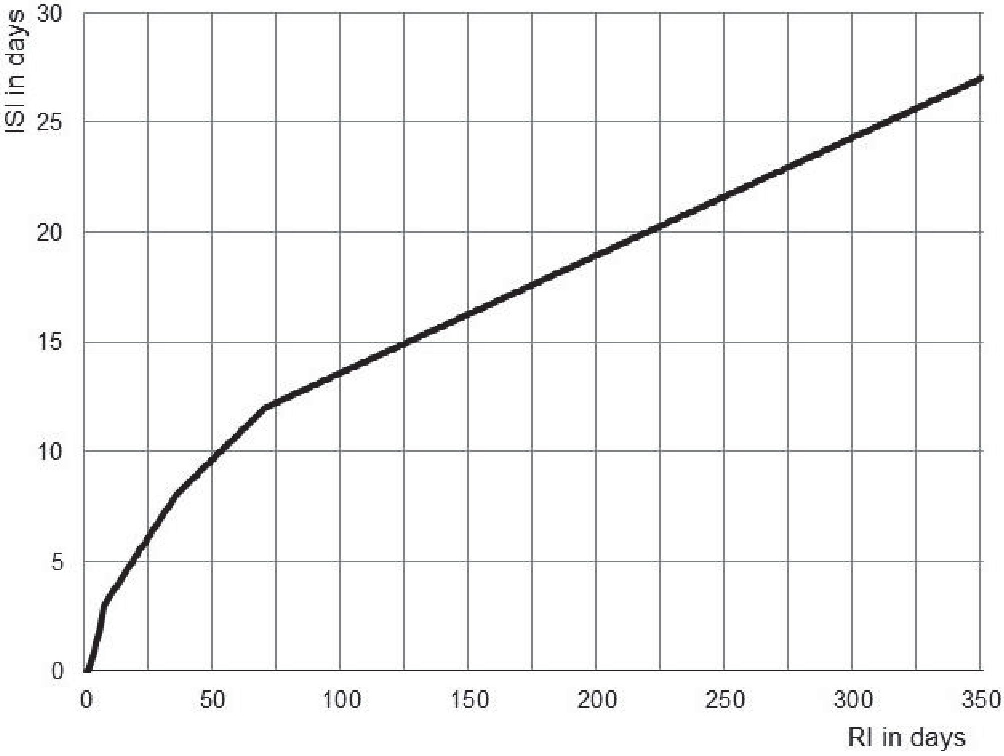


Figure 11 – Optimum gap for restudy, based on Cepeda et al., 2008

From their work, Cepeda et al. (2006, 2008, 2009) concluded that the optimal ISI must be approximately 15 percent of the RI. In line, Rohrer and Pashler (2007) suggested a ratio of 10 percent to 30 percent of ISI to RI and similarly, Pashler and team (2007) estimated the ratio to be 10 percent to 20 percent. All these findings account for the large as well as nonmonotonic spacing effects, showing that for every single RI, performance grew with increasing ISI and fell with ever-increasing ISI (Cepeda et al., 2008). Additionally, both laboratory and real-world experiment findings show that "*final memory performance is an inverted-U-shaped function of the degree of spacing between the study episodes […]. This inverted-U-shaped function has been observed when spacing is varied in the range of seconds and minutes […] although it may be seen more readily when study episodes and memory tasks are distributed over longer, more educationally relevant intervals of several days or more*" (Gerbier & Toppino, 2015, p. 53). Gerbier and Toppino (2015) further concluded that one can never space too much, it is rather the opposite: if too little spacing occurs, it can have detrimental effects for long-term memory formation.

2.2.3.3 Summary

No one ideal spacing schedule has been recommended in previous research. Contracting schedules seem to be beneficial for short RIs, hence for circumstances in which knowledge is only required for short intervals, whereas expanding and equal schedules are better to follow for longer RIs, i.e., when long-term knowledge retention is the aim (Cepeda, et al., 2008; Küpper-Tetzel et al., 2014; Mozer et al., 2009; Wiseheart et al., 2019). Again, no clear statement of preference for either expanding or equal schedule can be made. One rather must consider several different aspects such as the difficulty of the material to be learnt, how retrieval shall take place, and the given time frames (Storm et al., 2010 in Kang, 2016). Yet, when including explanatory feedback on errors, some studies have shown a slightly higher advantage of the equal spacing condition over the expanding one (Cull, 2000; Karpicke & Roediger, 2007a).

With regards to feedback in general, any type of feedback (e.g., immediate or delayed) on errors during a test is always beneficial for subsequent knowledge retention (Balota et al., 2006; Cepeda et al., 2006; Corral et al., 2021; Cull et al., 1996; Pashler et al., 2007). It should nevertheless be mentioned that delayed feedback might even be more beneficial compared to immediate feedback as it resembles an additional learning session which asks the learner to retrieve information from memory and hence memory traces are strengthened (Corral et al., 2021).

Based on the findings mentioned above, this research is going to experimentally assess different equal spaced learning schedules compared to a massed learning schedule. When setting up the experiments, it was ensured that the nonmonotonic relationship of ISI and RI, as postulated by Cepeda and team (2008) was considered (see chapter 4.1 for details).

2.2.4 Theoretical accounts

Several theoretical neuroscientific accounts have been proposed to explain the spaced learning effect as well as all above-mentioned related findings (Balota et al., 2011; Walsh et al, 2018b). These theories evolved mainly from vocabulary research and can be classified as either *descriptive* (i.e., trying to capture the underlying conceptual cognitive frameworks) or *computational* (i.e., trying to make quantitative predictions about the spaced learning effect) (Walsh et al., 2018b). Both sets will be explained next.

2.2.4.1 Descriptive theories of the spaced learning effect

Descriptive theories include the *encoding variability theory* (Glenberg, 1979; Melton, 1970), *study-phase retrieval theory* (Greene, 1989; Thios & D'Agostino, 1976), *retrieval difficulty theory* (Bjork & Bjork, 1992), *deficiency processing* (Hintzman, 1974) and its variants *consolidation theory* (Wickelgren, 1972), and *inattention theory* (Hintzman, 1974). The theories themselves can one by one be "*distinguished [...] with respect to the underlying memory processes they put forward to explain the [spaced learning effect]*" (Küpper-Tetzel & Erdfelder, 2012, p. 38). Research (e.g., Delaney et al., 2012; Dunlosky et al., 2013, Toppino & Gerbier, 2014) postulated that due to the effect size and extent of the spaced learning effect, it might be the case that the theories are not mutually exclusive and various accounts are to be considered.

From their meta-analysis, Cepeda and team (2006) concluded that *encoding variability theory*, *study-phase retrieval theory,* and *consolidation theory* should be used to explain the spaced learning effect as these are the only theories which reflect the "*finding that optimal ISI increases as retention interval increases*" (Cepeda et al., 2006, p. 370). Based on these findings, this research solely immerses into these three theories and disregards all other proposed. Table 4 shows their critical review.

Table 4 – Overview of descriptive theories of the spaced learning effect, author's own compilation

Encoding variability theory	
Definition and basic assumptions	▪ "*The amplitude of the fluctuation of the learning context between two retrieval episodes*" (Latimier et al., 2021, p. 978). ▪ All information stored in memory is associated to the context it was learned in, so-called retrieval cues (Glenberg, 1979).
Impact on spaced learning	▪ Many contextual fluctuations occur due to multiple learning sessions (Smolen et al., 2016). Contrary to massed learning, which is encoded as one context only (Balota et al., 2011; Kang, 2016). ▪ Greater variability of stored contextual elements and greater overlap of the same contextual information encoded during the initial learning session and the repetitions, leads to stronger memory traces and higher probability of later successful retrieval (e.g., Bird, 2010; Bjork & Allen, 1970; Cepeda et al., 2006; Foot-Seymour et al., 2019; Glenberg, 1979; Wiseheart et al., 2019).
Handling of ISI and RI interaction	▪ This theory "*[nicely handles] the intriguing spacing by retention interval interaction*" (Balota et al., 2011, p. 101), revealing that the ideal ISI increases as the RI increases, however with the ratio decreasing with higher RI. ▪ An inverted U-shaped memory function occurs, confirming the nonmonotonic effects of ISIs (Cepeda et al., 2006; Küpper-Tetzel & Erdfelder, 2012; Smith & Scarf, 2017).
Critical review and applicability	▪ "*Lacks specificity (e.g., how can fluctuations in context and cue/trace overlap be measured?) and does not account for the full range of distributed practice effects reported in the literature*" (Wiseheart et al., 2018, p. 554). ▪ The theory cannot stand for itself but a hybrid account combining mechanisms of this, and study-phase retrieval theory is more applicable (e.g., Benjamin & Tullis, 2010; Delaney et al., 2010; Mozer et al., 2009).

Table 4 continued – Overview of descriptive theories of the spaced learning effect

	Study-phase retrieval theory
Definition and basic assumptions	▪ "*Greater difficulty retrieving an earlier learning instance of an item leads to greater strengthening of the memory trace for this item during a subsequent learning event*" (Wiseheart et al., 2019, p. 553), which is viewed as retrieval cue of the initial learning event (Cepeda et al., 2006; Küpper-Tetzel & Erdfelder, 2012). ▪ The more time elapses between initial and succeeding learning events, the more effortful the retrieval of an item becomes (e.g., Bjork & Bjork, 1992; Braun & Rubin, 1998; Delaney et al., 2010; Hintzman & Block, 1973; Latimier et al., 2021; Thios & D'Agostino, 1976). Contrary to massed learning, where no memory reactivation happens.
Impact on spaced learning	▪ Spacing learning sessions allow the learner to forget the item to-be-learned. As a result, the difficulty of re-accessing the item increases, memory traces are strengthened and less forgetting occurs (Braun & Rubin, 1998; Delaney et al., 2010; Latimier et al., 2021; Smolen et al., 2016; Vlach et al., 2019; Wiseheart et al., 2019).
Handling of ISI and RI interaction	▪ ISIs between spaced study sessions should not become excessively large as the item to-be-learned from the initial learning session might be forgotten before the succeeding repetition, thus the initial memory trace cannot be reactivated, and the long-term memory formation mechanism is unavoidably hurt (Küpper-Tetzel & Erdfelder, 2012; Smolen et al., 2016; Wiseheart et al., 2019). This implies the nonmonotonic, inverted U-shape spacing function: if the ISI becomes too long, the benefits of study-phase retrieval are offset and forgetting occurs (Gerbier & Toppino, 2015).
Critical review and applicability	▪ The theory does not make it clear whether it assumes that the best ISI increases monotonically with the RI (Cepeda et al., 2006). It also lacks explanations of the underlying factors affecting the retrieval process and thus, might not be sufficient as a reason for the spaced learning effect to occur (Wiseheart et al., 2019). ▪ Similar to above, the theory also cannot stand for itself, however a hybrid theory combining mechanisms of this and encoding variability theory is more applicable (e.g., Benjamin & Tullis, 2010; Delaney et al., 2010; Mozer et al., 2009).

Table 4 continued – Overview of descriptive theories of the spaced learning effect

Consolidation theory	
Definition and basic assumptions	▪ The conceptual model for consolidation proposed by Landauer, (1969) and extended by Wickelgren (1972) is based on two assumptions: ▪ First, if a second learning session is happening too close in time to the initial learning session this will not noticeably strengthen the consolidation of the memory trace from the first learning session. ▪ Second, the likelihood that a second learning session yields a successful retrieval of the memory trace of the initial learning session decreases with time.
Impact on spaced learning	▪ Temporarily distributing learning sessions allows memory traces to get more durable and long-term memory is eventually created. Contrary to massed learning, where processing is substantially poorer as no retrieval takes place, hence no durable memory is formed (e.g., Bird, 2010; Delaney et al., 2012; Gerbier & Toppino, 2015; Hintzman & Block, 1973; Smolen et al., 2016).
Handling of ISI and RI interaction	▪ It confirms the nonmonotonic interaction of ISI and RI as well as the inverted U-shaped spacing function: when the ISI becomes too long and initial memory traces cannot be retrieved, all benefits of spaced learning are offset (Cepeda et al., 2006).
Critical review and applicability	▪ It does not specify whether the ISI increases monotonically with the RI (Cepeda et al., 2006).

2.2.4.2 Computational theories of the spaced learning effect

Next to the theoretical accounts, several computational models to describe and explain the spaced learning effect have been proposed (Benjamin & Tullis, 2010; Lindsey et al., 2009; Mozer et al., 2009; Pavlik & Anderson, 2005; Raaijmakers, 2003; Walsh et al., 2018b). All computational theories "*are instantiated as a set of mathematical equations, implemented in running computational software. These models make testable predictions, which render them falsifiable*" (Walsh et al., 2018b, p. 1329). Mozer and team (2009) argued that these mathematical models are relatively complex as "*the brain contains multiple, interacting* *memory systems whose decay and interference characteristics depend on the specific content being stored and its relationship to other content. Consequently, these computational theories are fairly flexible and can provide reasonable post-hoc fits to spacing effect data*" (p. 1322). Yet, these models are of great usage to optimise spaced study schedules to maximise knowledge retention for situations in which the exact timing of when knowledge needs to be recalled is unknown. This is especially the case in real-life learning environments in which only a rough idea exists until when items should be available for recall (Lindsey et al., 2009). Among the most elaborated ones in use (Mozer et al., 2009; Walsh et al., 2018b) are:

- ACT-R (adaptive control of thought-rational) declarative memory equations (Pavlik & Anderson, 2005)
- PPE (Predictive Performance Equation; Jastrzembski & Gluck, 2009; Walsh et al., 2018a), which extends the general performance equation (Anderson & Schunn, 2000)
- SAM (Search of Associative Memory) model for spacing and repetition effects (Raaijmakers, 2003)

Both ACT-R and PPE postulate that spaced learning reduces retrieval losses, however, both computational models are based on different assumptions. Walsh and team (2018b) explained these differences as follows: *"[ACT-R] maintains separate decay rates and elapsed times for each instance of an item, and activation is summed across all instances at retrieval. Subsequent repetitions do not interact with stored instances, and their contribution to activation is additive. In contrast, PPE computes the average decay rate and elapsed time for all instances. The latter quantity, elapsed time, is weighted toward more recent experiences. The positive effects of practice and the negative effects of decay are multiplied to predict performance. By reducing the average decay rate or elapsed time, repetitions can have a super-additive effect on activation.*" (p. 1131). Being conceptually different to ACT-R equations and PPE, the SAM model revealed that spaced learning effects are not only dependent on the storage of multiple contextual elements within the same memory trace, but on correct retrieval or recognition during the repetitive learning sessions (Greene, 1989; Maddox, 2016; Raaijmakers, 2003). Yet, it postulates the same assumptions on relearning as ACT-R does (Walsh et al., 2018b).

Even though SAM, ACT-R, and PPE make use of different descriptive accounts as underlying principles (SAM – encoding variability and study-phase retrieval theory; ACT-R – study-phase retrieval and consolidation theory; PPE – study phase retrieval theory) (Cepeda et al., 2006; Walsh et al., 2018a/b), they all verify that:

1. Massed studying accelerates acquisition
2. Spaced studying increases retention
3. Knowledge retention develops nonmonotonically with the increase of the ISI (Walsh et al., 2018b)

2.2.4.3 Summary

Enhanced knowledge retrieval and retention can arise from any of the above-mentioned theoretical accounts, all of which "*have been derived primarily from studies involving memory for direct repetitions of the same stimuli*" (Corral et al., 2021, p. 796). Further, it could be shown that none of them can account exclusively for the broad range of spaced learning findings (Dunlosky et al., 2013; Küpper-Tetzel & Erdfelder, 2012; Wiseheart et al., 2019). Rather, they support and back-up one another, either as hybrid models or as sequenced brain processes. Nevertheless, all of the theoretical descriptive accounts presented in this review have been replicated through computational models, which were all able to underpin the findings of the spaced learning literature (Cepeda et al., 2008; Küpper-Tetzel & Erdfelder, 2012; Walsh et al., 2018b; Wiseheart et al., 2019).

Overall, one could derive the following remarks: once information is studied, a short-term memory trace is built in the long-term memory store. Assuming the item is reviewed again immediately (e.g., time frame of seconds or not at all), these repetitions will not lead to new encoding in memory. If the short-term memory trace of the item to be learnt has left short-term memory and was also well recalled from the long-term memory store, new contextual elements are stored and encoded along this specific memory trace. Either of these additional elements can be accounted as a retrieval cue. Further, the successful re-encoding of the memory trace in the long-term memory store represents the study-phase retrieval theory (Küpper-Tetzel & Erdfelder, 2012). All in all, "*one can, therefore, conceive of the total [learning] process as a cycle of acquisition, loss, and reacquisition of information, with diminishing amounts of information lost during the intervals between exposures until the information becomes part of permanent knowledge retrievable without further re- learning*" (Bahrick, 1979, pp. 297).

2.2.5 Retrieval practice

The so-called *retrieval practice* supplements the positive learning effects of spaced learning with the *testing effect*, which has been shown across a range of test formats, test materials, as well as learners' age groups, outcome measures, and time intervals (Cogliano et al., 2019; Dunlosky et al., 2013; Morehead et al., 2015; Roediger & Karpicke, 2006a/b). Testing itself is as of yet said to not only assess a learner's knowledge to be able to judge whether a particular instructional activity, such as spaced learning, yielded in envisioned learning outcomes (Wiliam, 2011), but to also enhance it in multiple ways such as (Dunlosky et al., 2013; Dunlosky & Rawson, 2015; Kang, 2016; Kapler et al., 2015; Roediger & Karpicke, 2006a):

- Improving the memory traces of the information tested
- Allowing transfer of the learnt information to new circumstances
- Slowing forgetting and boosting subsequent learning
- Enabling the learner to engage in much deeper encoding processes

Irrelevant of schedule (e.g., equal or expanding), spaced schedules of retrieval have yielded much better knowledge retention after intervals lasting months and years (e.g., Bahrick et al., 1993; Cepeda et al., 2009). In contrast, immediate tests only benefited massed learning practice as information are recalled from working memory only (Roediger & Karpicke, 2006a, b). Further, and in accordance with major findings of the spaced learning research, retrieval practice research yielded an inverted U-shaped function which shows that retrieval declines as a function of spacing, i.e., with time passing by the likelihood of correct recall declines (Gerbier & Toppino, 2015). Also, all before-mentioned theoretical accounts underlying the spaced learning effect have been proven correct for the testing effect (Balota et al., 2007; Gerbier & Toppino, 2015; Modigliani 1976).

Any form of retrieval practice, such as testing, is said to diminish the information forgotten as learners engage in much deeper and more comprehensive encoding processes which lead to far more retrieval cues for the information to be learnt and is applicable well beyond verbal learning (Pashler et al., 2007; Rodriguez et al., 2021; Roediger & Karpicke, 2006a). Therefore, it is interestingly not of relevance whether a learner receives feedback on test performance (Butler et al., 2008; Karpicke & Roediger, 2007a). Coupled with the fact that tests are still seen as a means of assessment and are neither favoured by teachers nor by learners, wide-spread application of retrieval practice is missing (Dunlosky et al., 2013; Roediger & Karpicke, 2006b).

2.2.6 Metacognition

Research on students' declarative metacognitive knowledge, i.e., the "*explicit factual knowledge about the variables that affect memory performance*" (Vlach et al., 2019, p. 117) revealed that learners inaccurately rated massed learning higher than spaced learning as well as retrieval practice, although the same learners had higher retrieval performances in any spacing condition (Bjork, 1999; Dunlosky et al., 2013; Morehead et al., 2015; Simon & Bjork, 2001; Vlach et al., 2019; Zechmeister & Shaughnessy, 1980). Learners need to be confident and must believe that a chosen learning strategy leads to success and errors in rating one's confidence through so-called *judgments of learning* (JOLs) (e.g., rating one's confidence of having correctly answered a test question on a scale from 0 to 100 percent) can have a heavy impact on metacognitive knowledge (Cogliano et al., 2019; Dunlosky & Metcalfe, 2009; Roediger & Karpicke, 2006a).

Erroneous assessment of one's encoding fluency (how easy the learning process was experienced) and retrieval fluency (how fast or easy any learnt information can be retrieved from memory) can lead to falsified assessments of one's own knowledge due to believing one has learnt something better than one actually has (overconfidence) or believing one has not learnt as much as one in fact has (under-confidence) (Finn & Tauber, 2015; Moore & Healy, 2008). When asked, many students reported that fast learning, i.e., quick encoding fluency, is equal to good learning, yet performed much worse on final tests than those students who reported the opposite. These learners can be classified as being overconfident caused by overestimating their learning (Moore & Healy, 2008). With regards to retrieval fluency, learners might become overconfident in their knowledge. They may assume they know something if it seems familiar to something they have once learnt and thus stop studying. When tested, however, it becomes apparent that they do not know as much as they thought they knew (Ariel & Karpicke, 2018; Cogliano et al., 2019; Kornell & Bjork, 2007, Zechmeister & Shaughnessy, 1980). Furthermore, overconfidence in learning can arise when believing that one has learnt or performed better than somebody else, i.e., overestimating their own abilities compared to those of others (Moore & Healy, 2008) or take too lightly how fast something that has been learnt can be forgotten (Putnam et al., 2016).

Falsified assessments of one's own knowledge can lead to ineffective study strategy selection such as massed instead of spaced studying and re-reading instead of self-testing or retrieval practicing (Ariel & Karpicke, 2018; Dunlosky & Rawson, 2015; Karpicke & Roediger, 2007a/b; Karpicke et al., 2009; Kornell & Bjork, 2007; Tullis et al., 2013). Surprisingly, a survey by Morehead, Rhodes and DeLozier (2015) revealed that students are technically aware that spaced practice is more beneficial to long-term knowledge retention than massed practice, yet students engage in massed studying. A proposed reason for this was that encoding during massed studying and re-reading appears more fluent and less difficult and thus, learners' confidence in their learning progress is higher (Bjork, 1999; Kornell & Bjork, 2008; Simon & Bjork, 2001; Son, 2010).

Theoretically, this perception is true as massed studying as well as re-reading "*gives the illusion that memory storage is strong and will last for a long time*" (Wiseheart et al., 2019, p. 554). Driven by this illusion, students inaccurately evaluate their JOL's (Tullis et al., 2013). However, research has shown en masse: learning acquisition processes which are experienced as more effortful (such as spaced learning and retrieval practice) and create desirable difficulties, promote significant better long-term memory retention, which is less prone to forgetting (Bjork, 1999). Yet again, learners' confidence in what was learnt is very low (Finn & Tauber, 2015; Tullis et al., 2013).

Overall, "*learners are not always cognizant of what learning conditions are best, and they can even develop misconceptions as to what processes do and do not benefit their own learning*" (Vlach et al., 2019, p. 116). To overcome this lack of consciousness, learners as well as learning instructors should be educated and explain the effectiveness of a chosen learning strategy which in turn enhances confidence in learning strategies that have been dismissed due to believed ineffectiveness (Ariel & Karpicke, 2018). It has been shown that through the application of retrieval practice, learner's metacognitive knowledge on how they actually did on the test was enhanced and thus, the accuracy of their JOLs increased and was even stronger when feedback was given (Butler et al., 2007; Cogliano et al., 2019; Dunlosky & Rawson, 2015). By better understanding the benefits spaced learning or retrieval practice has for long-term knowledge creation, learners automatically engage in more thorough encoding and processing strategies during the learning experience and a learner is better equipped to distinguish between items they know and those they do not know. In the end, this will lead to profound long-term memory (DeWinstanley & Bjork, 2004 in Tullis et al., 2013).

2.3. Instructional design – Bridging theory and practice

2.3.1 Instructional design and e-learning

The preceding part of this chapter outlined the current state of research and great importance of thoughtfully designed work-based learning interventions to foster cognitive activities for long-term application and transfer of the learnt material. For this, spaced learning has proven to be one of the most efficient instructional methods.

Taking again Schunk's (2012) definition of learning and the three elements it contains (see chapter 2.1.1), instructional design processes and models can be used to create formal learning experiences which lead to change in the learner's behaviour (Clark & Mayer, 2016). Instructional design "*refers to the systematic and reflective process of translating principles of learning and instruction into plans for instructional materials, activities, information resources, and evaluation*" (Smith & Ragan, 2005, p. 4). Yet, the literature review has also shown that a specific understanding of *instructional design* is lacking. Several approaches to instructional design exist as well as various processes and models to apply it – all aimed at helping people learn (e.g., Beck et al., 2001; Dabbagh & Bannan-Ritland, 2004; Diamond, 1989; Gagné et al., 2005; Gustafson & Branch, 1997, 2002; Merrill, 2002; Newby et al., 1996; Twelker et al., 1972; van Merriënboer, 1997). These instructional design models and approaches are most often based on common learning theories which provide clues as to how, why and what kind of learning is to be expected (Clark & Mayer, 2016; Ertmer & Newby, 2013;

Khalil & Elkhider, 2016; Taras, 2005). *Behaviourism*, *cognitivism* and *constructivism* are the relevant learning theories to date, among which the cognitive learning theories are the most influential ones (Ertmer & Newby, 2013; Khalil & Elkhider, 2016; Smith & Ragan, 2005).

Another finding from the literature review was that driven by digitisation efforts and the recent COVID-19 pandemic, digital work-based learning interventions are gaining traction and importance for work-based learning (for a review see chapter 2.1.2). Hence, to evaluate the research at hand, all experiments undertaken focused on digital learning, herein called *e-learnings*. *E-learning* is defined as "*a combination of content and instructional methods delivered by media elements such as words and graphics on a digital device intended to build job-transferable knowledge and skills linked to individual learning goals or organizational performance*" (Clark & Mayer, 2016, p. 457). E-learning interventions have become a very common means of delivering learning content in academia and the business environment as it allows learners to flexibly learn whenever and wherever they want and need. Additionally, 'unlimited' options for providing learning content are available (Halawi et al., 2009; Yam & Rossini, 2013).

All in all, this research seeks to understand if the *instructional method* of *spaced learning* results in better retrieval of newly acquired knowledge from long-term memory *within an e-learning environment*. As any formal learning intervention is said to only be beneficial when it is congruent with *human cognitive architecture* and the principles of instructional design derived from research-based practice (Caspersen & Bennedsen, 2007; Shavelson & Towne, 2002; Sweller, 2016; Sweller et al., 2019), the experimental e-learnings of this research were designed following the principles of cognitivism, more precisely *cognitive load theory* (Mayer, 2019; Mayer & Moreno, 2002; Richards & Taylor, 2015).

Thus, the following chapter will summarise the present state of knowledge on *human cognitive architecture* (chapter 2.3.1.1) which provides the foundation for cognitive learning theories. It is followed by an overview of the distinction between *biologically primary and biologically secondary knowledge* (chapter 2.3.1.2) which altered the understanding of the influence of human cognitive processes on instructional design. *Cognitive load theory and its implications on multimedia instruction* will be introduced (chapter 2.3.1.3) as the basis for the design of the e-learning experiments carried out for this research. Chapter 2.3.1.4 then summarises the findings on instructional design and e-learning.

2.3.1.1 Human cognitive architecture

The *human cognitive architecture*, which underlies all human learning, is based on the following mechanisms: a limited *working memory*, an unlimited *long-term memory* as well as the relationship between the two (Caspersen & Bennedsen, 2007; Paas & Sweller, 2012; Sweller et al., 1998, 2019; Sweller & Paas, 2017). Working memory is said to resemble an individual's consciousness, i.e., attentional focus, as only the information available in working memory can be actively perceived (Shipstead et al., 2014; Sweller et al., 1998). Information entering working memory is processed in one of the working memory's slave systems: the *phonological loop* – which stores and manages verbal information; the *visuospatial sketchpad* – which stores and manages visual information; or the *episodic buffer* or *attention control* – which helps organise incoming information in chronological order (Baddeley 1986, 2000; Kirschner et al., 2011; Sorden, 2005). These systems are independent from one another and cannot compensate for a limited capacity in the other systems (Brünken et al., 2003). This is of particular importance and should always be considered when designing learning interventions, especially those teaching novel, unfamiliar information (Sweller et al., 2019). The limitations of working memory, also referred to as "*the defining aspects of human cognitive architecture*" (Sweller et al., 1998, p. 262), amount to three to five elements of novel information that can be handled at once (Baddeley, 1986; Miller, 1956) for a storage duration of maximum 20 seconds if not recalled (Peterson & Peterson, 1959). Anything exceeding this capacity overloads the working memory which in turn is not able to process the overload (Caspersen & Bennedsen, 2007). Limitations of working memory do not apply to information that has already been learnt and that is stored and recalled from long-term memory (Paas et al., 2010).

Long-term memory is capable of unconsciously storing and organising an unlimited, unimaginable number of facets of information, varying in complexity and size (Caspersen & Bennedsen, 2007; Paas & Sweller, 2012; Sweller et al., 1998; Tricot & Sweller, 2014). Therefore, it is regarded to be the centrepiece of human cognition (Paas et al., 2010). Information is stored in long-term memory in the form of hierarchically structured mental knowledge constructs called *schemas*¸ which allow to deal with current problems using past problem-solving experiences (Paas et al., 2010; Sweller et al., 1990, 1998; see also Schema Theory by Chi et al., 1982). These schemas "*are used to store and organize knowledge by incorporating or chunking multiple elements of information into a single element with a specific function*" (Paas & Sweller, 2012, p. 28), i.e., working memory is able to process several elements of information as if these were one element (Caspersen & Bennedsen, 2007; Paas & Sweller, 2012).

While learning, working memory processes and encodes incoming information elements, alongside extracting and manipulating this information before creating or modifying it as a schema in long-term memory (Caspersen & Bennedsen, 2007; Sweller et al., 1998). This process is called *schema construction* (Sweller et al., 1998). Once a lower level schema is created in long-term memory, it can be reactivated from working memory and modified to build higher level schemas, which combine different interacting elements of lower level schemas, into a single higher level schema (Caspersen & Bennedsen, 2007; Paas et al., 2004; Sweller et al., 1998). Through extensive practice, *schema automation* occurs. Repeated practice leads to the development of increasingly complex schemas. As a result, working memory starts to unconsciously process a problem, a procedure or a task and thus, more capacity is available for other new tasks and activities which require conscious processing (Sweller et al., 1998).

Consequently, working memory capacity is freed as it no longer considers individual, lower level schemas but instead complex, higher level ones and in turn, more sophisticated processing takes place (Sweller et al., 1998). A basic model of human cognitive architecture is shown in Figure 12.

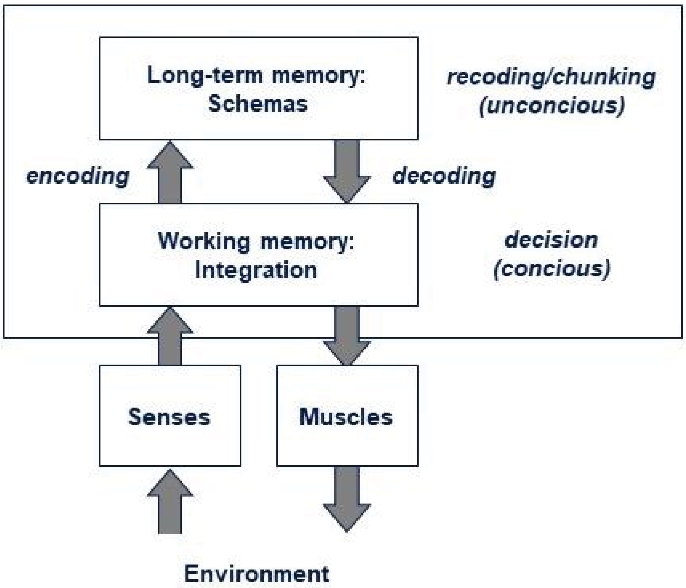


Figure 12 – The human cognitive architecture, adapted from Caspersen & Bennedsen, 2007, p. 113

It is argued that the level of domain knowledge held in an individual's long-term memory determines their performance (Paas et al., 2010). It was found that skilled learners, i.e., experts have a seemingly unlimited working memory capacity compared to novices when dealing with familiar information already stored in long-term memory (Chi et al., 1982; de Groot, 1965; Sweller & Paas, 2017). Schema construction and automation make experts in a specific domain feel as if what they have learnt is clear and easy, yet, they are forgetting about the complexity and difficulty of acquiring these schemas (Tricot & Sweller, 2014). Also, they tend to forget that what seems easy to retrieve for one individual due to schema automation is difficult for another individual with less expertise (Sweller et al., 1998).

Instructional design should therefore support both schema construction and schema automation to make maximum use of working memory limitations (Paas et al., 2004). *Cognitive* *load theory* is a commonly acknowledged instructional theory that has detected a number of strategies to facilitate cognitive learning, bearing in mind limited working memory capacity (Paas & Ayres, 2014; Paas & Sweller, 2012; Sweller, 1988, 2010; Sweller et al., 1998, 2011; van Merriënboer & Sweller, 2005). The theory as well as its recommended evidence-based implications for *multimedia e-leaning interventions* will be explained after a brief excursion into Geary's (2002, 2008, 2011) distinction of *biologically primary* versus *biologically secondary knowledge*, as it has severe implications on considering human cognitive architecture when it comes to instructional design (Tricot & Sweller, 2014).

2.3.1.2 Excursion: Biologically primary versus biologically secondary knowledge

In his work on evolutionary educational psychology, Geary (e.g., 2002, 2008, 2011) divided knowledge into two categories, namely a) *biologically primary knowledge* and b) *biologically secondary knowledge*. Biologically primary knowledge resembles all skills and knowledge gained over numerous decades and generations (Geary & Berch, 2016). It is acquired effortlessly, intuitively, without leveraging working memory capacity and most importantly, acquisition happens outside of educational settings (Sweller, 2011, 2016; Sweller et al., 2019; Sweller & Paas, 2017; Tricot & Sweller, 2014). Examples of biologically primary knowledge are learning to listen and to speak our native language, face recognition, basic social relations, problem solving of unfamiliar problems, transfer of existing knowledge to new circumstances – all of which represent generic-cognitive skills concerned mostly with learning, thinking, and problem solving (Sweller, 2011, 2016; Sweller et al., 2019; Tricot & Sweller, 2014). Without these, human survival is said to not be possible (Sweller, 2016).

Biologically secondary knowledge on the other hand "*is knowledge that has become culturally important and needs to be acquired in order to function appropriately in a society*" (Sweller, 2011, p. 40). Examples are reading, writing, mathematics, and even more comprehensive, all knowledge taught with explicit instruction in educational contexts such as schools, universities, workplaces, etc. (Paas & Sweller, 2012; Sweller, 2011, 2016; Tricot & Sweller, 2014). These taught skills are said to be domain-specific rather than generic-cognitive and their acquisition requires a lot of external motivation, consciousness, and is further associated with difficulty and effort (Paas & Sweller, 2012; Sweller, 2016). This also implies that in contrast to a generic-cognitive skill/knowledge, which "*is a mental process that can be applied to a wide variety of unrelated areas [...], a domain-specific skill is a procedure that only can be applied to a specific range of areas*" (Sweller, 2016, p. 361).

Working memory capacity limitations as outlined in the preceding chapter do not apply for biologically primary knowledge but only to novel biologically secondary knowledge since humans have not evolved to learn how to process the latter automatically (Paas & Sweller, 2012; Sweller, 2016). It follows that, if biologically primary knowledge does not require significant working memory capacity and teaching it is rather unsuccessful, instructional design and procedures taking into account working memory capacity and duration limits can only be applied for explicit tuition of biologically secondary knowledge (Sweller, 2016). Despite this, "*acquisition of biologically secondary knowledge is heavily dependent on the prior acquisition of primary knowledge*" (Tricot & Sweller, 2014, p. 266). Learning how to apply specific generic-cognitive skills and biologically primary knowledge in certain domains facilitates the acquisition of these domain-specific skills and biologically secondary knowledge (Sweller et al., 2019; Youssef-Shalala et al., 2014). Consequently, skilled performance as per Paas and Sweller (2012) is also dependent on how well an individual has learnt to make use of their generic-cognitive skills and thus, is able to relieve working memory capacities (Sweller & Paas, 2017). Thus, while biologically primary knowledge and cognitive generic skills are unteachable, they can support acquisition of domain-specific skills and biologically secondary knowledge (Paas & Sweller, 2012). Herein, instructional interventions shall always consist of "*a combination of primary and secondary skills with the secondary skill being the only part that is learned*" (Sweller et al., 2019, p. 272).

Cognitive load theory (Chandler & Sweller, 1991; Sweller, 1988) builds onto this distinction of the two above-mentioned knowledge categories, using biologically primary means to learn biologically secondary knowledge (Tricot & Sweller, 2014). The underlying information processing system consists of five biologically primary principles mirroring information processing aspects relying on Darwin's (see Sweller & Sweller, 2006) evolutionary theory (Sweller et al., 2011; Sweller, 2016; Tricot & Sweller, 2014). These are the information store, the borrowing and reorganizing, the randomness as genesis, the narrow limits of change and the environmental organizing and linking principle (Sweller, 2016).

Based on these insights on human cognitive architecture as well as Geary's distinction of biologically primary and secondary knowledge, it was concluded that instructional design's purpose is to ease the acquisition of novel domain-specific skills and biologically secondary knowledge whilst managing working memory limitations (Sweller, 2016; Tricot & Sweller, 2014). Several instructional effects, i.e., cognitive load effects, were derived and recommended for instructional usage (Sweller et al., 2019). These will be introduced in the next chapter.

2.3.1.3 Cognitive load theory and multimedia instruction in e-learning

In 1988, the first all-encompassing description of the *cognitive load theory* was presented (Sweller, 1988; Sweller et al., 2019). Since then, cognitive load theory has become one of the most acknowledged theories on learning, educational psychology, and instructional design for learning complex cognitive tasks (Ayres & Paas, 2012; Paas et al., 2004). In essence, cognitive load theory is concerned with eradicating working memory limitations when learning complex, novel, and biologically secondary knowledge which would easily overwhelm learners so as to enhance the extent of knowledge held in long-term memory (Paas et al., 2010; Sweller, 2011, 2016). As mentioned in the previous part of this chapter, cognitive load is all load on "*working memory during problem solving, thinking, and reasoning (including perception, memory, language, etc.)*" (Caspersen & Bennedsen, 2007, p. 113). Cognitive load can "*be caused by the intrinsic nature of the task or by the manner in which the information within the task is presented to them*" (Kirschner et al., 2011).

The first type of load is called *intrinsic cognitive load* and depends on two things: first, the so-called *element interactivity* and second, the expertise level of the learner (Schnotz & Kürschner, 2007). In that sense, intrinsic load cannot be changed or manipulated unless the learner's expertise level changes (Caspersen & Bennedsen, 2007; Sweller, 2011). The level of element interactivity is central to the cognitive load theory. Element interactivity refers to "*the amount of interacting elements that have to be processed simultaneously*" (van Merriënboer & Sluijsmans, 2009, p. 56). Element interactivity is low when elements can be learnt in isolation such as the vocabulary of a new language (Sweller et al., 1998). Each word can be learnt non-interactively in isolation, successively, and does not impose a high load on working memory (Paas et al., 2010). Element interactivity is high when several elements must be processed simultaneously in working memory, such as when learning grammar, and an excessive load is put on working memory (Sweller, 2011; Sweller et al., 1998).

The concept of element interactivity is also applicable to the second type of cognitive load, namely *extraneous cognitive load*. Extraneous cognitive load is a result of inappropriate presentation of the learning material which does not support but rather hinders learning (Caspersen & Bennedsen, 2007; Sorden, 2005). Extraneous cognitive load increases element interactivity and blocks schema construction and automation (Paas et al., 2004; Sweller et al., 2019).

In contrast, *germane cognitive load*, the third type of cognitive load, encourages learning processes by supporting schema construction and schema automation as well as reducing element interactivity (Caspersen & Bennedsen, 2007; Paas et al., 2004). It is assumed that germane cognitive load is linked to intrinsic cognitive load as it reallocates working memory capacity "*from extraneous activities to activities directly relevant to learning by dealing with information intrinsic to the learning task*" (Sweller et al., 2019, p. 264). Thus, the more germane cognitive load focuses on extraneous cognitive load, the less learning occurs as less working memory resources can be allocated to intrinsic cognitive load and vice versa.

All three types of cognitive load add up but just as with working memory's three slave systems, available working memory capacity of a learner should never be exceeded (Kirschner et al., 2011; Sorden, 2005). This relationship presented as a function could resemble (Caspersen & Bennedsen, 2007):

Cognitive load = Extraneous load + Germane load + Intrinsic load

Yet, this function varies for different learners on various knowledge levels (Brünken et al., 2003). Generally speaking, one can argue that individual performance decreases at both ends of the continuum: either extreme underload or extreme overload (Issa et al., 2011; Paas et al., 2004; Figure 13).

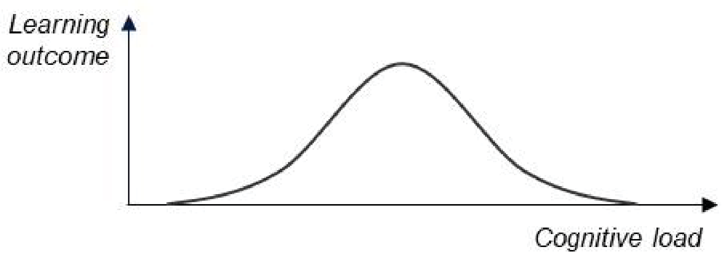


Figure 13 – Learning outcome as a function of cognitive load, adapted from Caspersen & Bennedsen, 2007, p. 113

When starting to learn something new, learners should not be overwhelmed through high element interactivity, otherwise successful performance will be impossible due to working memory capacity constraints (Ayres & Paas, 2012; Schnotz & Kürschner, 2007). However, as expertise develops, element interactivity can be increased as learners have cognitive schemas which allow them to manage higher intrinsic load, in other words, expertise mirrors understanding (Sweller et al., 1998; van Merriënboer & Sluijsmans, 2009). Nevertheless, misalignment of external learning task requirements and internal capabilities of the learners, i.e., their level of expertise, can lead to little or no learning even for experts: if a learning task is too easy for an individual and performance is highly automated, little if any working memory capacity is needed (Schnotz & Kürschner, 2007). Even though performance might be successful, learning does not occur.

It follows that "*the goal of the design of instruction is to optimize cognitive load for a particular learner*" (Brünken et al., 2003, p. 54). In doing so, it is essential to know the characteristics of the learners, i.e., understanding their maximum level of intrinsic cognitive load, and based on that design learning in a way that extraneous cognitive load (e.g., inadequate instructional methods, environmental distractions) is minimised and germane cognitive load (e.g., adequate instructional methods) is maximised, all aiming at easing the learners mental effort whilst engaged in the learning process (Andersen & Makransky, 2020; Brünken et al., 2003; Sweller, 2011; Sweller et al., 2019). This is especially important as what might be germane cognitive load for a novice could easily be extraneous cognitive load for an expert (Paas et al., 2004). Several cognitive load effects have been proposed to tie in with this (Sweller et al., 2019). Most of these effects are due to a decrease in element interactivity and thus a decrease in extraneous load, only few alter intrinsic cognitive load (Sweller et al., 2019). Table 5 summarises all of these and gives explanations and examples.

Table 5 – Overview of cognitive load effects, adapted from Sweller et al., 2019, pp. 266

Timeline of major cognitive load effects before and after 1998		
Effect	**First naming of the effect**	**Description**
Goal-free effect	Sweller and Levine (1982) in *Journal of Experimental Psychology: Learning, Memory and Cognition, 8*, 463-474.	Replace conventional tasks with goal-free tasks that provide learners with a non-specific goal.
Worked example effect	Sweller and Cooper (1985) in *Cognition and Instruction, 2*, 59-89.	Replace conventional tasks with worked examples that provide learners with a solution they must carefully study.
Completion problem effect	van Merriënboer and Krammer (1987) in *Instructional Science, 16*, 251-285.	Replace conventional tasks with completion tasks that provide learners with a partial solution they must complete.
Split-attention effect	Tarmizi and Sweller (1988) in *Journal of Educational Psychology, 80*, 424-436.	Replace multiple sources of information, distributed either in space (spatial split attention) or time (temporal split attention), with one integrated source of information.
Redundancy effect	Chandler and Sweller (1991) in *Cognition and Instruction, 8*, 293-332.	Replace multiple sources of information that are self-contained (i.e., they can be understood on their own) with one source of information.
Compound element interactivity effect	Sweller (1994) in *Learning and Instruction, 4*, 295-312.	Cognitive load effects that are found for high element interactivity materials are typically not found for low element interactivity materials. Actually, cognitive load theory is only relevant for complex learning.
Variability effect	Paas and van Merriënboer (1994a) in *Journal of Educational Psychology, 86*, 122-133.	Replace a series of tasks with similar surface features with a series of tasks that differ from one another on all dimensions on which tasks differ in the real world.
Modality effect	Mousavi et al. (1995) in *Journal of Educational Psychology, 87*, 319-334.	Replace a written explanatory text and another source of visual information (unimodal) with a spoken explanatory text and the visual source of information (multimodal).
Publications of the 1998 article cognitive architecture and instructional design		
Self-explanation effect	Tarmizi and Sweller (1988) in *Journal of Educational Psychology, 80*, 424-436.	Replace separate worked examples or completion tasks with enriched ones containing prompts, asking learners to self-explain the given information.

Table 5 continued – Overview of cognitive load effects, adapted from Sweller et al., 2019, pp. 266

Timeline of major cognitive load effects before and after 1998		
Effect	**First naming of the effect**	**Description**
Imagination effect	Cooper et al. (2001) in *Journal of Experimental Psychology: Applied, 7*, 68-82.	Replace conventional study of a procedure or concept to learn with imagination, where the learner is asked to imagine or mentally practice the concept or procedure.
Isolated elements effect	Pollock et al. (2002).in *Learning and Instruction, 12*, 61-86.	Replace a presentation of information/tasks with all interacting elements at once by initially presenting elements of information sequentially in an isolated form rather than in a fully interactive form.
Compound expertise reversal effect	Kalyuga et al. (2003) in *Educational Psychologist, 38(1)*, 23-31.	Cognitive load effects that are found for low expertise learners (e.g., worked example effect, goal free effect) are typically not found or even reversed for high expertise learners.
Compound guidance-fading effect	Renkl and Atkinson (2003) in *Educational Psychologist, 38*, 15-22.	Cognitive load effects that are relevant in the beginning of a longer educational program (e.g., guided problem-solving, worked examples) are no longer relevant in later stages of the program, after the learners acquired sufficient expertise.
Collective working memory effect	Kirschner et al. (2009) in *Educational Psychology Review, 21*, 31-42.	Replace individual learning tasks with collaborative tasks so that more cognitive resources become available.
Compound transient information effect	Leahy and Sweller (2011) in *Applied Cognitive Psychology, 25*, 943-951.	Cognitive load effects that are found for transient information (e.g., self-pacing effect, segmentation effect, modality effect) are typically not found for non-transient or less transient information.
Human movement effect	Paas and Sweller (2012) in *Educational Psychology Review, 4*, 27-45.	Replace static or unrealistic visualisations with visualisations showing human movements.
Compound self-management effect	Roodenrys et al. (2012) in Applied Cognitive Psychology, 26, 878-886.	Cognitive load effects that are found for ill-designed instructional materials (e.g., split-attention) are not found when learners are explicitly taught how to reduce the associated extraneous load.

Cognitive load theory is of particular importance for designing *multimedia learning environments* such as e-learning interventions, in which distractions and unimportant instructional units can easily be incorporated (Andersen & Makransky, 2020; Brünken et al., 2003; Clark & Mayer, 2016; Mayer & Moreno, 2002; Sorden, 2005). Multimedia learning environments or multimedia instruction are defined as a lecture with pictures and words (Mayer & Moreno, 2002). Nevertheless, as cognitive load theory has proven "*the way in which text and pictures (or diagrams) are combined can have both positive and negative effects*" (Ayres &

Paas, 2012, p. 828) on the learning process. Similarly, in 2002, Mayer and Moreno presented the *cognitive theory of multimedia learning*. This theory depicts the same three instructional design goals as cognitive load theory does, namely a) reducing extraneous load, b) promoting germane load and c) facilitating intrinsic load (Andersen & Makransky, 2020).

The cognitive theory of multimedia learning draws upon working memory's slave systems and argues that visual and verbal information are processed in a dual channel system (in the phonological loop and the visuospatial sketchpad) and that the processing capacities of these systems are limited (Mayer & Moreno, 2002; Sorden, 2005). It is argued that displaying a combination of static or dynamic verbal or graphic information in distinct modalities leads to well-designed multimedia learning interventions, as they make best use of working memory's different slave systems (Kirschner et al., 2011).

The cognitive theory of multimedia learning further assumes that meaningful learning occurs when learners are cognitively active, meaning they actively select important words and visuals, mentally organise these into particular representations and integrate these into already existing representations in long-term memory (Brünken et al., 2003; Issa et al., 2011; Mayer & Moreno, 2002; Sorden, 2005). Figure 14 illustrates this process.

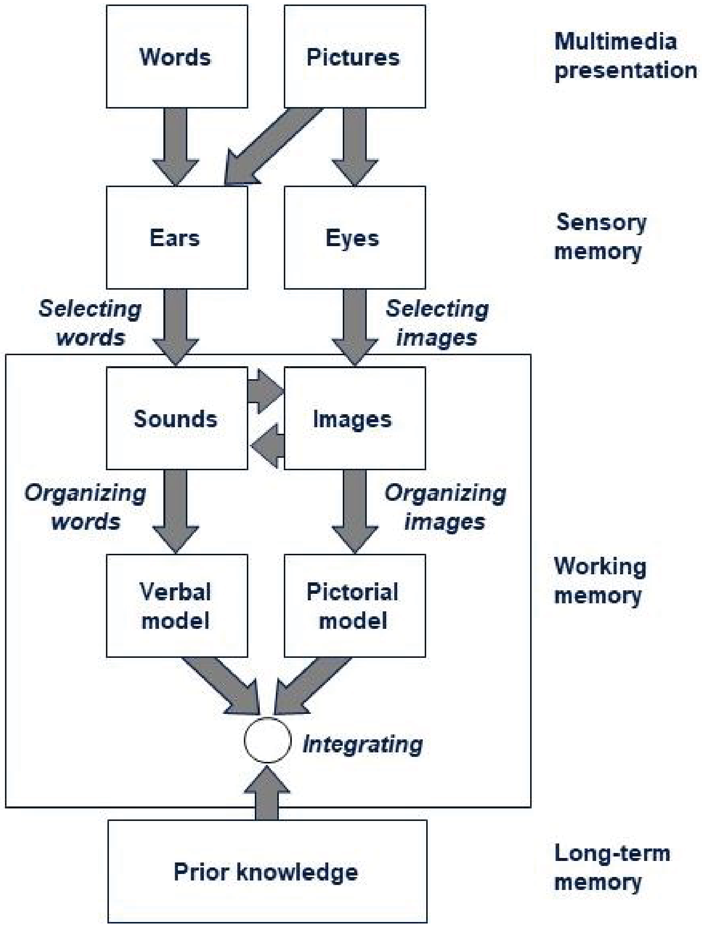


Figure 14 – Overview of dual-channel theory for multimedia learning, adapted from Issa et al., 2011, p. 820

Pitfalls of multimedia e-learning interventions are the seamless, unlimited possibilities of presenting content which can easily overwhelm working memory capacities and also, many interventions are designed to rather entertain learners than educate them and thereby clearly miss learning targets (Sorden, 2005). Regularly failing to learn and perform well on complex tasks, due to e.g., incorrect presentation of the material resulting in cognitive overload, might further lead to learners' decrease in motivation (Paas et al., 2004). Therefore, evidence-based measures, principles, and practice must be applied in instructional design to ease schema construction and automation (Brünken et al., 2003; Clark & Mayer, 2016; Issa et al., 2011; Sorden, 2005).

In addition, Mayer and Moreno (2002, 2003) have identified several evidence-based aids to guide multimedia-based e-learning processes aiming at promoting meaningful learning. These are mainly focused on two things: a) reduction of cognitive load (Mayer et al., 2004; Mayer & Moreno, 2003; Paas et al., 2003) and b) increasing interest of the learner (Mayer et al., 2004; Mayer & Moreno, 2003).

On a), several recommendations were introduced such as "*multimedia aids, in which students understand more deeply when they receive words and pictures rather than words alone; contiguity aids, in which students understand more deeply when words and pictures are presented simultaneously rather than successively; coherence aids, in which students understand more deeply when unneeded words and sounds are eliminated; modality aids, in which students understand more deeply when words are presented as narration rather than on-screen text; and redundancy aids, in which students understand more deeply when words are presented solely as narration rather than as narration and on-screen text*" (Mayer & Moreno, 2002, p. 117). In a nutshell, instructional design for multimedia e-learnings shall focus on a visually interesting and insightful layout with learning activities being centred around the concepts to be taught instead of trying to stimulate learners too much (Issa et al., 2011; Sorden, 2005). Referring to the previously introduced cognitive load theory, all aids and effects of the cognitive theory of multimedia learning can be classified as either focusing on the reduction of extraneous load, the increase of germane load or the facilitation of intrinsic load (Brünken et al., 2003).

With regards to b), i.e., increasing the learner's interest, it was found that personalisation within the learning experience e.g., a conversational communication style such as use of first and second person or comments addressed towards the learners during multimedia presentations, increases learners interest to cognitively engage in the learning process and as a result, yielded in higher transfer test scores compared to formal and unaddressed communication (Mayer et al., 2004).

2.3.1.4 Summary

From the above, it can be summarised that when designing any, and in particular multimedia e-, learning interventions, instructional designers must adhere to human cognitive architecture and more specifically to the limited working memory capacities (Mayer & Moreno, 2003; Sorden, 2005; Sweller et al., 2019). To achieve meaningful learning (Mayer & Moreno, 1998), learning activities should be designed in a way that reduces cognitive load and fosters learners' interest, both aiming at freeing capacity in working memory and then, using this free capacity for further processing (Issa et al., 2013; Mayer et al., 2004). Thereby, it is of importance to base all instruction on evidence-based procedures which "*help learners attend to relevant information, organise it into a coherent mental representation, and integrate it with prior knowledge*" (Issa et al., 2013, p. 389).

Cognitive load theory (Sweller et al., 1998) as well as the cognitive theory of multimedia learning (Mayer & Moreno, 2002) take all of this into consideration. They aim to decrease extraneous load, increase germane load, and facilitate intrinsic load. Hence, they should foster schema construction and automation to increase the amount of knowledge held in long-term memory (Paas et al., 2010; Sweller, 2011, 2016). Optimal learning design includes an analysis of the learners to better manage their cognitive load, as what seems easy for domain experts might put a huge load onto domain novices (Paas et al., 2004).

For this, several cognitive load effects have been proposed (for a review see Sweller et al., 2019). With regards to the design of multimedia e-learning interventions, it is especially important to understand how to most effectively use learning technologies and not base decisions on how exciting a tool is but rather how it best suits the cognitive learning process of the particular learner (Sorden, 2005). In line with their cognitive theory of multimedia learning, Mayer and Moreno (2002) introduced five multimedia aids (multimedia, contiguity, coherence, modality, and redundancy aids) to control extraneous, germane and intrinsic cognitive load, which were followed in designing the experimental e-learnings of this research (see chapter 4.1). Finally, personalisation during the learning intervention enhances learner's motivation and interest to cognitively engage in the learning experience and therefore, more meaningful learning takes place in the long-term (Mayer et al., 2004).

2.3.2 Evaluating learning – Learner's sophistication versus proficiency

Even though applying instructional design methods to learning course design is indispensable, it is not complete without clearly formulated learning objectives (Dirksen, 2016). Learning objectives lead to expected learning outcomes of what learners know and which skills and abilities learners possess after completion of any learning intervention (Harris & Clayton, 2019; Stanny, 2016). Every single part of the course design should be based on the learning objectives set – from the presentation of the learning content to the learning activities offered. If learners are to apply what they have learnt in their daily work, a real-world scenario could be part of the learning experience. If learners are meant to learn technical vocabulary, multiple-choice learning could be more suitable (Göldi, 2011). Clearly set learning objectives also help to examine the expertise levels of the learners, i.e., what is the overarching expectation of what the learning intervention achieves for the learners. Based on this, the assessment to be applied is decided (Bloom et al., 1956).

Bloom's learning taxonomy (Bloom et al., 1956) is the most known and applied psychological learning taxonomy to date with most influence on learning course design, learning outcome description as well as learning assessment creation (chapter 2.3.2.1; Göldi, 2011; Stanny, 2016). Even if the originally proposed learning taxonomy classified learning along three areas, namely cognitive (knowledge and intellectual skills), affective (interests, attitudes, and understanding of values), and psychomotor (physical, manual, motor, sensory or technical skills), it only applies to the cognitive learning area (Halawi et al., 2009). The taxonomy was reassessed in 2001 by Anderson, Krathwohl and team, underpinning the importance it has especially for the cognitive domain and adding latest research on cognitive learning processes (Adams, 2015; Radmehr & Drake, 2019). The following chapter will review both learning taxonomies and elaborate on what these mean for creating learning objectives as well as how to assess them (chapter 2.3.2.2).

2.3.2.1 Bloom's learning taxonomy

Bloom's learning taxonomy was originally intended to serve as a

- "*common language about learning goals to facilitate communication across persons, subject matter, and grade levels;*
- *basis for determining for a particular course or curriculum the specific meaning of broad educational goals, such as those found in the currently prevalent national, state, and local standards;*
- *means for determining the congruence of educational objectives, activities, and assessments in a unit, course, or curriculum; and*
- *panorama of the range of educational possibilities against which the limited breadth and depth of any particular educational course or curriculum could be contrasted*" (Krathwohl, 2002, p. 212).

The taxonomy was ground-breaking as it, for the first time, systematically organised cognitive skills for meaningful learning to take place (Adams, 2015; Halawi et al., 2009; Murphy, 2007). Learning and thinking were classified into six categories of cognitive skills, hierarchically organised, gathering lower-level cognitive skills (i.e., those requiring little cognitive processing such as knowing and understanding) and higher-level cognitive skills (i.e., those requiring deeper cognitive processing such as analysing, differentiating or evaluating) (Adams, 2015; Stanny, 2016). The six categories were knowledge, comprehension, application, analysis, synthesis, and evaluation (Bloom et al., 1956; Krathwohl, 2002; Murphy, 2007):

- "*knowledge […] focuses on memorization, recognition, and recall of information; […]*
- *comprehension […] focuses on organization of ideas, interpretation of information, and translation […]*
- *application […] focuses on problem solving, use of particulars, and principles […]*
- *analysis […] focuses on finding the underlying organization, and the division of a whole into components […]*
- *synthesis […] focuses on a combination of ideas to form something new, creating something unique whether verbal or physical […]*
- *evaluation […] focuses on making judgments about issues, resolving disparities or disagreements*" (Halawi et al., 2009, p. 375).

All of these categories, with the exception of application, were further broken down in subcategories (Krathwohl, 2002). Taken together, it was argued that the acquisition of all skills along the taxonomy is required for learners to develop critical thinking abilities (Murphy, 2007). All categories and subcategories are shown in Figure 15.

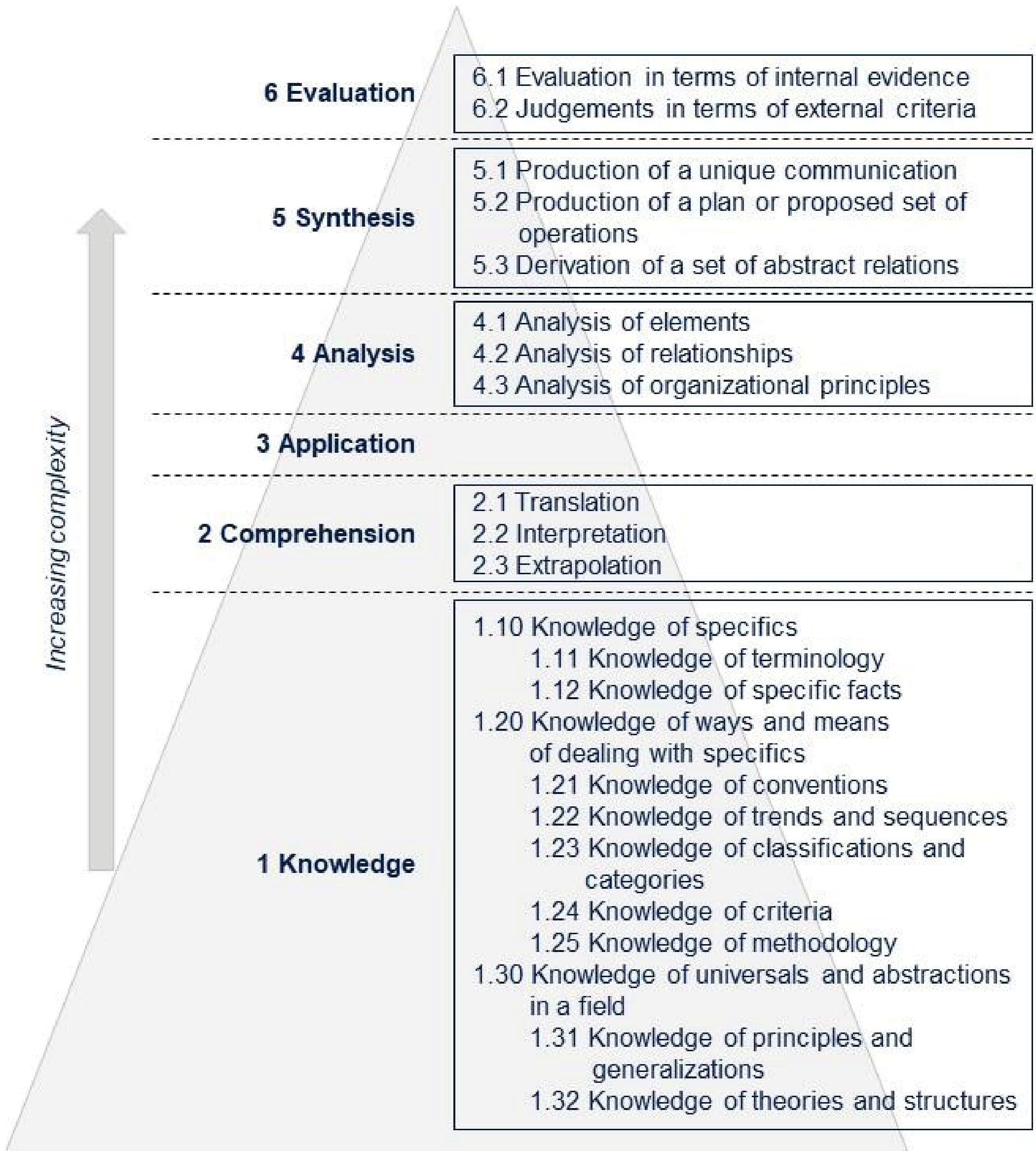


Figure 15 – Bloom's original taxonomy, adapted from Adams, 2015, p. 153 & Krathwohl, 2002, p. 213

Bloom et al. (1956) used several verbs describing each category, illustrating a continuing growth of cognitive abilities of the learner: whilst lower-level verbs define learning of terms, facts, knowledge, basic principles, and methods, higher-level verbs define learning of complex thinking skills, such as applying knowledge to real-world problems or examining opposing interpretations (Murphy, 2007; Stanny, 2016). Bloom et al. (1956) did not use the differentiation of lower-level and higher-level cognitive skills (Adams, 2015), yet it was done so afterwards, as Bloom et al. (1956) regarded their taxonomy as a gradual process, meaning that the next higher-level in the taxonomy requires more sophisticated cognitive skills than the previous (Conklin, 2005). As Agarwal put it: "*To achieve a higher objective or category, one must first master cognitive processes at a lower category. In other words, before comprehension, application, or analysis can take place, a student must first acquire knowledge.*" (Agarwal, 2019, p. 190). For each of these categories, different assessment tools are used, as shown in Table 6.

Table 6 – Learning and thinking categories with respective forms of assessment, adapted from Adams, 2015, p. 152

Category	Assessment
Knowledge	▪ Multiple-choice/short-answer questions for retrieval or recognition of information
Comprehension	▪ Paraphrasing information in one's own words ▪ Classifying items in groups ▪ Comparing/contrasting information with similar objects ▪ Explaining a principle to other people
Application	▪ Using knowledge, skills, methods in unfamiliar situations
Analysis	▪ Distinguishing between fact/opinion ▪ Breaking down information into single units
Synthesis	▪ Creation of a new product in a particular situation
Evaluation	▪ Reflection upon a learning session by using feedback and assessment results of learners to conclude value of a learning intervention

Even if the taxonomy was ground-breaking, several critiques arose both conceptually and empirically (Conklin, 2005). Weaknesses in the differentiation and hierarchisation of the six main categories were pointed out (Paul, 1995, in Stanny, 2016; Sitte, 2001); and, the evolving order of categories was questioned (Anderson, Krathwohl, et al., 2001). Further, it was argued that a lack of a common language typically leads to a lack of mutual understanding, meaning that the verb lists proposed by Bloom and associates "*function like a thesaurus or a rhyming* *dictionary. Authors use these lists to discover alternative words when they search for variety for their writing. But an author who uses a thesaurus has no guarantee that she will produce crisp prose or inspired poetry. Similarly, authors who use a list of Bloom's taxonomy verbs to write measurable [learning objectives] have no guarantee that the verbs they select for their learning outcomes will provide all the detail needed to clearly describe increasing levels of expertise*" (Stanny, 2016, p. 10).

Though being useful, Bloom's taxonomy does not provide a holistic picture of learning (Stanny, 2016). In 2001, Anderson, Krathwohl and team proposed a *revised taxonomy* incorporating latest findings of research in cognitive sciences, latest learning principles and theories as well as data of a meta-analysis conducted on Bloom's taxonomy (Adams, 2015; Conklin, 2005; Radmehr & Drake, 2019). The revised taxonomy is a two-dimensional framework, called *Taxonomy Table*, in which the knowledge dimension resembles the subcategories of the initial taxonomy (vertical axis), and the cognitive processes dimension resembles the former six categories of Bloom's taxonomy (horizontal axis) (Anderson, Krathwohl, et al., 2001). Thereby, verbs are being used to portray the aimed cognitive process and nouns are used to explain the knowledge a learner is supposed to use or acquire (Anderson, Krathwohl et al., 2001).

The knowledge dimension addresses four types of knowledge that could be taught in any learning intervention, namely *factual*, *conceptual*, *procedural*, and *metacognitive* knowledge (Adams, 2015). This differentiation, however, is not set, and no clear dividing lines exist. Rather, the different types of knowledge can overlap or merge and subtypes for each of the knowledge categories exist (Anderson, Krathwohl, et al., 2001; Radmehr & Drake, 2019). This structure is shown in Table 7.

Table 7 – Bloom's revised taxonomy, adapted from Krathwohl, 2002, p. 214

Structure of the knowledge dimension of the revised taxonomy

Factual knowledge

- **Definition**: The basic elements that students must know to be acquainted with a discipline or solve problems in it.
- **Types**:
  - Knowledge of terminology
  - Knowledge of specific details and elements

Conceptual knowledge

- **Definition**: The interrelationships among the basic elements within a larger structure that enable them to function together.
- **Types**:
  - Knowledge of classifications and categories
  - Knowledge of principles and generalisations
  - Knowledge of theories, models and structures

Procedural knowledge

- **Definition**: How to do something; methods of inquiry and criteria for using skills, algorithms, techniques and methods.
- **Types**:
  - Knowledge of subject-specific skills and algorithms
  - Knowledge of subject-specific techniques and methods
  - Knowledge of criteria for determining when to use appropriate procedures

Metacognitive knowledge

- **Definition**: Knowledge of cognition in general as well as awareness and knowledge of one's own cognition.
- **Types**:
  - Strategic knowledge
  - Knowledge about cognitive tasks, including appropriate contextual and conditional knowledge
  - Self-knowledge

The cognitive process dimensions has, as the original taxonomy had, six hierarchical dimensions; however, dimensions were either renamed or changed: the former knowledge dimension was renamed 'remember', the former comprehension dimension was renamed 'understand', 'synthesis' was renamed 'create' and placed highest on the hierarchy; all other dimensions were kept but changed to the respective verb forms: apply, analyse and evaluate (Anderson, Krathwohl et al., 2001) All cognitive process dimensions and subdimensions still build on each other, increasing in complexity from lower-level to higher-level cognitive processing. Yet, Anderson, Krathwohl and team designed it in a flexible way, allowing the overlap of dimensions and therefore, allowing for broader applicability than the original one

(Krathwohl, 2002). The structure of the cognitive process dimensions as well as inherent sub-dimensions can be seen in Figure 16.

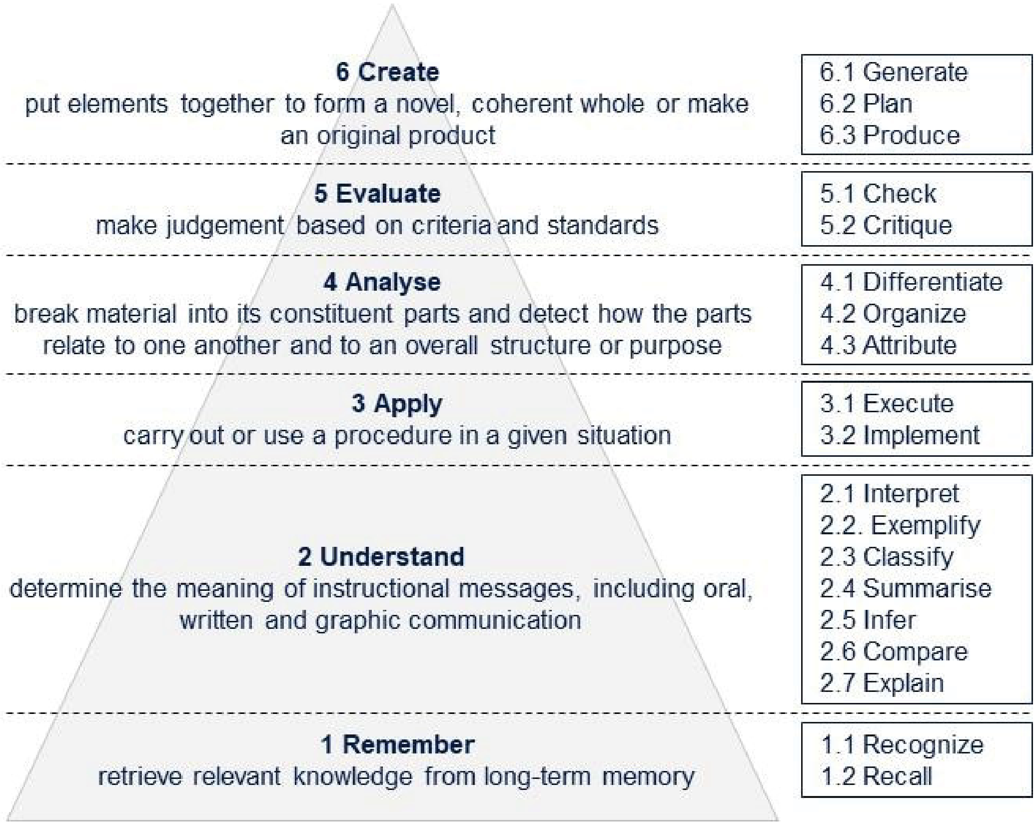


Figure 16 – The cognitive process dimensions of Bloom's revised taxonomy, adapted from Agarwal, 2019, p. 190 & Krathwohl, 2002, p. 215

Putting both dimensions together results in the two-dimensional framework: both the cognitive process dimension and the knowledge dimension can be used independently or linked for a comprehensive analysis of instruction, understanding, and assessment (Radmehr & Drake, 2019).

Thus to summarise, *Bloom's taxonomy for learning, teaching, and assessing* has been used for a long time by education and learning professionals and designers for traditional offline teaching but in recent years also for online teaching and learning interventions (Halawi et al., 2009). Providing a practical structure for knowledge transfer, the taxonomy as well its revision, is used to classify learning objectives, enabling a smooth transition from the simple to the complex and serve as an assessment tool to evaluate the performance of the learners (Adams, 2015; Anderson, Krathwohl, et al., 2001; Halawi et al., 2009; Radmehr & Drake, 2019).

2.3.2.2 Assessing learning

Embedded learning objectives formulated on the basis of a profound learning taxonomy, such as the revised Bloom's learning taxonomy (see previous chapter), as well as carefully planned and effectively delivered instruction, do not relate to the different levels of understandings various learners acquire when taking part in the same learning intervention (Wiliam, 2011). Therefore, *learning assessments* are needed, as "*it is only through assessment that we can find out whether a particular sequence of instructional activities has resulted in the intended learning outcomes*" (Wiliam, 2011, p. 3). Yet, assessments do not only help teaching and instruction professionals to gain an understanding of how well learners acquired what was supposed to be taught, they also serve as guidance for learners on what they regard as important, impacts the value they put on these tasks, and further affects future transfer of these (Schellekens et al., 2021).

In terms of measuring memory performance, *retention* and *transfer tests* are used to assess learning with the former "*measur(ing) how much of the presented material is remembered, [and the latter] measur(ing) how well the learner can apply what was learned to new situations*" (Issa et al., 2011, p. 821). Both assessments of memory performance, are measured sometime after a learning intervention has taken place (Schmidt et al., 2019). Even though some people regard both types of tests as equal, others argue that it is transfer tests which should be the overall aim of instruction as these provide a better understanding of learners' level of knowledge than retention tests do (Mayer et al., 2004; Schmidt et al., 2019; Sorden, 2005). Transfer can either be near or far: near transfer relates to applying what was learnt to a related domain and far transfer relates to applying what was learnt to an unrelated domain (Sala &

Gobet, 2017). If performance at test sometime after a learning intervention is as good as right after the learning intervention has taken place, one can argue that nothing was forgotten and vice versa, if performance is worse, that something was forgotten (Schmidt et al, 2019).

As discussed in chapter 2.1.1 successful learning is only achieved when the three basic cognitive memory functions of encoding, storage, and retrieval have been successful (Spielman et al., 2018). Performance of all three memory functions can be assessed retrospectively: encoding performance can be measured through *recognition tests*, storage performance is usually measured through *cued recall tests,* and retrieval performance is usually measured through *free recall tests* (Wise, 2020). *Recognition tests* require the least effort as learners are just asked to recognise something they have previously learnt (Dirksen, 2016; Wise, 2020). These tests are mainly held as multiple-choice tests, containing both answers that the learners haven been exposed to and not exposed to and are not meant to enable learners to apply the learnt material in real-life (Dirksen, 2016; Göldi, 2011; Wise, 2020).

The opposite is the case with *recall tests*, which enhance knowledge retention and are intended to enable real-life application of the learnt information (Dirksen, 2016; Göldi, 2011). *Cued recall tests* require even more cognitive effort of learners, as they are asked to remember the sequence of steps to be taken (Dirksen, 2016). *Free recall tests* which assess retrieval performance require the most effort, as learners are not given any clue to answer a test question but are asked to freely recall everything they know about a specific topic (Wise, 2020). To not distort performance results, it is recommended to first have a free recall test, then the cued recall test, and recognition tests last as the first type of test does not provide any clue at all and the last one provides stimulus clues and therefore a reversed order might promote recall which does not exist (Wise, 2020). Memory performance can be reported either as a percentage or number of correctly recognised/recalled items, however determination of correct items in cued and free recall tests is more difficult than for recognition tests, which can easily be evaluated (Dirksen, 2016; Wise, 2020). Recall tests require an individual to subjectively evaluate a response to determine the learner's proficiency, which is why it is advised to define what is regarded to be acceptable and profound beforehand (Connors, 2021; Dirksen, 2016; Wise, 2020).

Learning assessments can further be formative or summative, two terms formulated and introduced 1967 by Scriven. *Formative assessments* are also called *assessments for learning* (Black & Wiliam, 2009) as they are designed to improve overall student learning and are carried out regularly during the learning process (Connors, 2021; Schellekens et al., 2021). *Summative assessments*, on the other hand, are referred to as *assessments of learning*, as they take part after a learning process has come to an end and judge the learner's performance and whether learning has taken place (Connors, 2021; Schellekens et al., 2021). It is argued that both forms of assessment should be used together and build on each other (Dixson & Worrell, 2016; Houston & Thompson, 2017): during the learning intervention, formal or informal formative assessments should be used to enhance students' performance, providing both instructors and learners feedback on learning progress (Black & William, 1998; Broadbent et al., 2021). This information can then be used by instructors to modify instructional activities (Yam & Rossini, 2013) and by learners to become self-regulated through developing more accurate meta-cognition on their level of knowledge, a better understanding of learning objectives (McDaniel et al., 2011) as well as the provision of the opportunity to learn from mistakes (Goldfinch & Hughes, 2007). After a learning intervention has taken place, summative assessments should be used to holistically assess how much learners have actually learnt or retained from the intervention (Dixson & Worrell, 2016; Houston & Thompson, 2017; Perera-Diltz & Moe, 2014). Thereby, it has been shown that feedback provided during the formative assessments is central as it allows for learners to perform better during summative assessments, regardless if learning takes place online, offline or combined (Broadbent et al., 2021). Characteristics of both forms of assessment can be found in Table 8.

Table 8 – Formative and summative assessments, adapted from Dixson & Worrell, 2016, p. 154

Characteristic of formative and summative assessments		
Characteristic	**Formative assessment**	**Summative assessment**
Purpose	To improve teaching and learning To diagnose student difficulties	Evaluation of learning outcomes Placement, promotion decisions
Formality	Usually informal	Usually formal
Timing of administration	Ongoing, before and during instruction	Cumulative, after instruction
Developers	Classroom teachers to test publishers	Classroom teachers to test publishers
Level of stakes	Low stakes	High stakes
Psychometric rigor	Low to high	Moderate to high
Types of questions asked	What is working What needs to be improved How can it be improved	Does student understand the material Is the student prepared for next level of activity
Examples	Observations Homework Question and answer sessions Self-evaluations Reflections on performance Curriculum-based measures	Projects Performance assessments Portfolios Papers In-class examinations State and national tests

Both forms of assessment share several similarities such as "*assessors look for evidence of achievement, […] (j)udgements are made about the match between evidence and criteria, […] (j)udgements invokes communication, [and] (j)udgements are economic processes.*" (Knight, 2002, p. 277). Yet, one decisive difference of both is who the information regarding learning status is directed to (Knight, 2002): formative assessments inform learners about their learning progress and hence, enhances successive learning (Houston & Thompson, 2017; Perera-Diltz & Moe, 2014). Summative assessments, on the other hand, inform external parties about a learner's status of learning in the form of an evaluation or certification (Dixson & Worrell, 2016; Houston & Thompson, 2017). Hence, formative assessments are forward looking whereas summative assessments are backward looking (Taras, 2005).

However, formative assessments can also be used summatively, e.g., when external parties are informed about a learner's progress, and summative assessments can also be used formatively, e.g., when the results of a summative assessment are used to improve subsequent learning interventions (Dixson & Worrell, 2016; Perera-Diltz & Moe, 2014; Taras, 2005). Both forms of assessments are "*interdependent, as formative assessment feeds into summative and enhances the quality of information on which final judgements are made and communicated*" (Houston & Thompson, 2017, p. 5). Therefore, it is argued that assessments and subsequent feedback should not only take place at the end of the learning process but rather continuously throughout the learning process (Burke, 2009, in Schellekens et al., 2021).

2.3.2.3 Summary

This chapter outlined that not only a cognitive understanding of learning processes is required to design effective instruction, but also knowledge of learning goal description, learning objective setting, and assessment creation. The whole process of authentic instructional design is not a straightforward process but rather entails profound knowledge of educational and educational psychological research and theory (Connors, 2021; Gallardo, 2021).

Bloom's taxonomy for learning, teaching, and assessing (Bloom et al., 1956) and more importantly its revision (Anderson, Krathwohl et al., 2001) is the standard work for instructional design. Anderson, Krathwohl and teams' (2001) two-dimensional framework with its cognitive process dimension and knowledge dimension is comprehensively used to classify learning objectives, help learners transition from simple to complex as well as help learning professionals design assessments and support learner evaluation (Adams, 2015; Halawi et al., 2009; Radmehr & Drake, 2019). Therefore, no distinction is made between offline and online learning, the taxonomy finds application everywhere (Halawi et al., 2009).

To assess memory performance in terms of encoding, storage, and retrieval, three different tests can be used: recognition tests, cued-recall tests, and free recall tests (Wise, 2020). These tests can either be used to measure retention or in more applied settings to measure transfer (Issa et al., 2011). With regards to when assessments take place, formative and summative assessments are distinguished: with the first taking place along the learning process and the later taking place after the learning process has come to an end (Houston & Thompson, 2017).

Assessment design is an integral part of any instruction, as assessments are the only evidence of whether learning has taken place during the learning intervention (Wiliam, 2011). However, those responsible for writing assessments must be knowledgeable of the aim of the assessment: i.e., how sophisticated and proficient learners should be after the learning intervention (Connors, 2021; Dirksen, 2016; Gallardo, 2021). As sophistication and proficiency are subjective, instructional design and assessment writing require clear guidelines, starting with setting learning goals, developing learning objectives, agreeing on the appropriate assessment tools, and upfront definition of what is regarded correct and appropriate (Connors, 2021; Dirksen, 2016; Krathwohl, et al., 2001; Halawi et al., 2009; Radmehr & Drake, 2019; Wise, 2020).

3. Derivation of research focus and hypotheses

The following chapter 3.1 summarises the deficits in the spaced learning literature in the context of management and work-based learning, thereby justifying the overarching research question of whether the instructional method of spaced learning causes better learning in a work-based e-learning environment as introduced in chapter 1.2 and highlighting its importance. Afterwards, the directing research hypotheses that guide the experimental research at hand are presented in chapter 3.2.

3.1 Research gaps in the literature and overarching research question

As outlined in chapter 1, employee work-based learning and skill-building is becoming more and more important and success-critical in today's business environment. For this to be effectful, knowledge retention is critical (esp. when considering the challenges of lifelong learning). Since 1885, research argues that distributing learning sessions temporarily over time (spaced learning) results in better knowledge retention compared to one-off programmes (massed learning) (Cepeda et al., 2006). Ideally, applying spaced learning in educational situations would prevent general forgetting of information and accelerate subsequent relearning as well as skill acquisition (Blanchard, 2013; Glaveski, 2019; Ryu & Moon, 2019; Tauber et al., 2019; Walsh et al., 2018b). Yet, the number of experiments which address the context of this study with the requirement to improve on employee work-based and lifelong learning through spaced learning is very low – a fact that has been already stated by Dempster in 1988 and is still relevant today (Kapler et al., 2015).

An encompassing literature review (see table 3 in chapter 2.2.2) contains a section of 33 exemplary studies conducted within the experimental research of spaced learning. Even though 27 out of these 33 studies yielded moderate to strong effects (Balota, et al., 2006; Bird, 2010; Carpenter & DeLosh, 2005; Carpenter et al., 2009; Cepeda et al., 2008, 2009; Delaney et al., 2012; Donovan & Radosevich, 1999; Foot-Seymour et al., 2019; Grote, 1995; Kalenberg, 2017; Kang & Pashler, 2012; Kapler et al., 2015; Karpicke et al., 2016; Kondratjew & Kahrens, 2018; Kornell & Bjork, 2008; Kornell et al., 2010; Mettler et al., 2016; Rea & Modigliani, 1987; Reynolds & Glaser, 1964; Rohrer & Taylor, 2006, 2007; Seabrook et al., 2005; Smith & Rothkopf, 1984; Sobel et al., 2011; Verkoeijen et al., 2008; Vlach et al., 2008), only one of those provided insights on effects for managerial relevant learning (Kondratjew & Kahrens, 2018). Yet, this study contained experimental confounds as seven others did, for example by not considering the RI-ISI relationship (as discussed in chapter 2.2.3.2).

Furthermore, the extant literature in chapter 2 shows that the majority of spaced learning research has been conducted in the laboratory rather than in educational or life-relevant learning environments with comparably short ISIs and RIs (e.g., Balota et al., 2006; Benjamin & Craig, 2001; Braun & Rubin, 1998; Carpenter & DeLosh, 2005; Delaney et al., 2012; Kang & Pashler, 2012; Karpicke et al., 2016; Kornell & Bjork, 2008; Kornell et al., 2010; Mettler et al., 2016; Rea & Modigliani, 1987; Seabrook et al., 2005; Smith & Rothkopf, 1984; Verkoeijen et al., 2008; Vlach et al., 2008), a fact that has been asked for decades (Dempster, 1988). Within these studies, participants were mostly asked to learn facts (e.g., Mettler et al., 2016; Seabrook et al., 2005) or word lists/pairs (e.g., Balota et al., 2006; Benjamin & Craig, 2001; Braun & Rubin, 1998; Karpicke et al., 2016; Rea & Modigliani, 1987), recognising paintings (e.g., Kang & Pashler, 2012; Kornell & Bjork, 2008) or basic statistics (e.g., Smith & Rothkopf, 1984). Few studies considered more educational relevant ISIs and RIs on the scale of days, weeks, and months taught participants more intellectual skills such as more complex concepts such as physics concepts (Grote, 1995), meteorology concepts (Kapler et al., 2015) or solving permutation problems (Rohrer & Taylor, 2006, 2007). Thus, in summary, it can be stated that more non-laboratory studies with longer RIs on the scale of weeks, months, and years are needed from which knowledge on how to best balance ISIs and RIs can be derived (Balota et al., 2011; Cepeda et al., 2006; Mozer et al., 2009)

In addition, the critical review of literature identified three further areas still to be investigated: Firstly, more applied studies are needed which examine different spacing schedules as well as the influence of retrieval practice and feedback on learners in real-life learning environments with more than three re-learning sessions to understand the educational relevance of the spaced learning effect (e.g., Balota et al., 2011; Dunlosky et al., 2013; Kapler et al., 2015). Secondly, even with complex learning material, learners must not be "*trained to a criterion of perfect performance on all items during the second and subsequent learning sessions*" (Cepeda et al., 2006, p. 356) but re-learning sessions should be fixed, containing feedback to not derive misleading conclusions on how spacing impacts learning (Cepeda et al., 2006). Thirdly, some researchers claim that more studies examining the spaced learning effect outside the verbal learning domain are needed to showcase how to best bring laboratory-drawn results into the classroom (Dempster, 1988; Dunlosky et al., 2013; Gerbier & Toppino, 2015; Wiseheart et al., 2019), which does not even consider the additional demand on bringing these effects to use in work-based, business environments.

In addition, real-life implementation of spaced learning is lacking largely, even if it would be very relevant for the application in the field of digital work-based learnings and its impact on organisational development and employee retention and satisfaction (Bersin, 2018; Tauber et al., 2019). Regarding instructional design, not much clarity exits on how these work-based learning interventions must be designed to allow for long-term application of the learnt material (Beier, 2021; Billett, 2014; Blume et al., 2010; Kirchherr et al., 2020; Tuijnman & Boström, 2002; Vargas, 2017). Hence, guidance is needed on strategies to enhance long-term knowledge retention in the context of managerial work-based learning offerings.

It appears very interesting that there has not yet been a big push to integrate spaced learning into real-life educational contexts and bring it to actual application, even though the effects are universally confirmed. To contribute to filling this research gap, the research at hand aims at applying effects of spaced learning to complex digital work-based learning programmes to discover if and how far knowledge retention is enhanced through this addition or if spaced learning is solely applicable to low-level knowledge learning.

To address this, the overarching research question for the research at hand was formulated as follows:

Does the instructional method of spaced learning cause better learning in a work-based e-learning environment?

To answer this question with consideration of the above discussed shortcomings of the existing state of research, this experimental research will use (a) real-life, non-laboratory environments, (b) an RI of at least two weeks (to both consider optimal ISI and to pragmatically address needs for long-term/lifelong learning), (c) different moderating influences (i.e., repetition schedules, repetition frequencies, direct testing, and feedback), and (d) factual, conceptual as well as procedural knowledge with complex learning materials. It furthermore aims to (e) formulate clear guidelines to allow easier application of the spaced learning effect into modern day work-based trainings (i.e., e-learning), while (f) employing theories on human cognitive architecture, cognitive load as well as theories on multimedia instruction. Figure 17 shows this research focus and the research requirements.

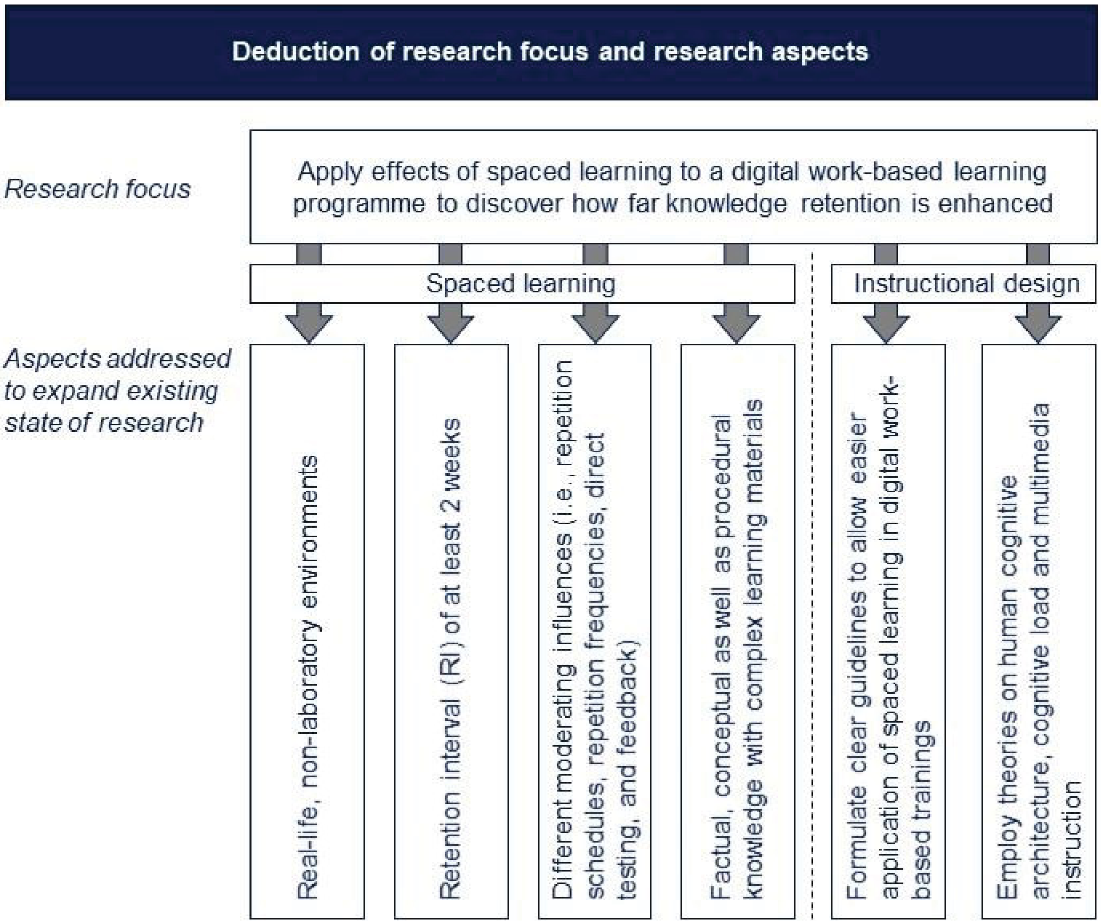


Figure 17 – Derived research focus and the research aspects, author's own compilation

By answering the research question, the research at hand aims at:

a) Firstly, contributing to the overall spaced learning and work-based learning literature. Educational as well as psychological research recognises the importance of spaced learning towards long-term knowledge retention, yet since this research is rather meagre in analysing or even considering effects of spaced learning to complex, higher-learning environments in which the long-term learning of new skills determines the success or failure of individuals, but also collectives such as companies, addressing the research question provides a valuable extension of the spaced learning and work-based learning literature.

b) Secondly, contributing to the calls for wide-spread application of spaced learning in real-life educational settings (Carpenter et al., 2012; Dempster, 1988) by addressing this research question in an empirical field experiment under lifelike learning conditions. By empirically testing if the effect can also be confirmed for complex, higher-level knowledge learning, a more holistic perspective to the spaced learning research is provided.

c) Thirdly, providing practical guidance for managers and learning professionals in charge to develop and design work-based learning programmes aiming at optimising knowledge retention, yielding into the build of applicable, transferrable skills, by potentially offering a sustainable means to enhance long-term knowledge retention for work-based learning. Although work-based learning opportunities are seen inevitable for both individuals and organisations (e.g., European Commission, 2018; Hanushek & Wößmann, 2010; Kugler et al., 2017; Rees, 2010; Tynjälä, 2008) and as a result are offered wide-spread globally (Glaveski, 2019), these are often offered for the wrong reasons, not considering biological brain processes underlying the human learning process, leading to dissatisfaction of participants which even results in employees leaving organisations to find better opportunities to grow elsewhere (Beier, 2021; Bersin, 2018; Billett, 2014; Chelovechkov & Spar, 2019; Glaveski, 2019; Tuijnman & Boström, 2002; Vargas, 2017). By addressing the mentioned research question, this research aims to help managers, external and internal professionals of work-based learning to design and develop learning programmes that do not waste investments because participants forget what they learnt in a short time.

3.2 Specific research hypotheses and their relevance

As will be described in chapter 4.1, this research follows an experimental research design, aiming at providing insights on the cause-effect relationship of spaced learning to knowledge retention of managerial e-learnings. Therefore, theoretically generated hypotheses are required to be proven or disproven to eventually formulate statements providing evidence about causalities and correctness of interpretation of the set research objectives per the above (Bortz & Döring, 2009, in Kubbe, 2018). As outlined in chapter 3.1, the overarching aim of the research at hand is to test if the positive effects of spaced learning can also be confirmed for complex managerial-relevant learning aiming at fostering lifelong learning in organisations. Thus, the hypotheses which will be introduced in the following derive mainly from the available spaced learning literature.

As discussed throughout chapter 2 and summarised in chapter 3.1, the robustness of the spaced learning effect has been shown to be highly replicable within the domains of verbal and trivia factual learning materials, associating names and faces across populations, age groups, and even for individuals with memory impairments (e.g., Balota et al., 2006; Carpenter et al., 2009, 2012; Fritz et al., 2007; Kalenberg, 2017; Kapler et al., 2015; Rea & Modigliani, 1987; Toppino, 1991; Toppino & DiGeorge, 1984; Vlach et al., 2008; Wilson & Evans, 1996). The majority of empirical studies in the field of spaced learning took place in relatively short laboratory studies (Balota et al., 2011; Bird, 2010; Carpenter et al., 2012; Cepeda et al., 2006), focusing on pure fact learning which enables learners to cognitively engage in proper skill building needed for lifelong application of the learnt (Bloom, 1956; Foot-Seymour et al., 2019). Since the spaced learning effect has been universally acknowledged, researchers started to question whether there is any specific way in which spaced learning interventions should be scheduled and if any relationship exists between how the single learning sessions are scheduled relative to one another and for how long knowledge should be retained (e.g., Carpenter et al., 2012; Cepeda et al., 2008; Karpicke & Bauernschmidt, 2011; Lindsey et al., 2009). All in all, three spacing schedules were considered: a contracting one, an expanding one, and an equal one. Yet, experiments comparing the different spacing schedules revealed mixed results throughout the spaced learning research (e.g., Balota et al., 2006; Carpenter et al., 2012; Cull et al., 1996; Karpicke & Bauernschmidt, 2011; Karpicke & Roediger, 2007a, 2010; Logan &

Balota, 2008; Spitzer, 1939; Toppino et al., 2018). Still, it has been found that whilst contracting schedules should be applied in situations in which knowledge must only be retained for short periods of time, i.e., short RIs, and expanding and equal schedules should be applied in situations where long-term knowledge retention is the aim, i.e., longer RIs (Cepeda, et al., 2008; Küpper-Tetzel et al., 2014; Mozer et al., 2009; Wiseheart et al., 2019). Deciding whether and when to take the expanding schedule or the equal schedule depends on a lot of factors, such as the complexity of the learning materials or what form of knowledge is to be taught (Storm et al., 2010 in Kang, 2016). But more importantly, it was found that it is most likely the RI which determines the ideal spacing schedule to follow (Latimier et al., 2021). Even though fictious learning curves comparing outcomes of expanding and equal spaced learning setups as well as carried out experiments yielded quite similar results of knowledge retention at final test, it is assumed that equal spacing schedules lead to more sustainable, long-lasting learning success (Carpenter & DeLosh, 2005; Cull, 2000; Karpicke & Roediger, 2007a; Logan & Balota, 2008).

As it is regarded as highly important to understand if the spaced learning effect is also applicable for complex, high-level knowledge, taught within digital learning environments, in educational and consequently work-based learning interventions to allow individuals as well as organisations to stay competitive in ever-faster changing markets, the following overarching research hypothesis H1 was set:

> ***Hypothesis H1:*** *Learning outcomes derived from e-learning interventions following an equal spaced learning design will be significantly higher than those derived following a massed design.*

When considering work-based learning interventions and the importance they represent for both individuals as well as organisations to ensure competitive advantage in local but also global job and economic markets, emphasis should be put on how to design work-based learning interventions to ensure long-term application and transfer of what was learnt. By doing so, the economic advantages resulting from education can be extended. It is widely agreed that return on learning is the maximised permanent transfer of the learnt into skills which can be automatically used (Connors, 2021; Gallardo, 2021; Göldi, 2011; Ryo & Moon, 2019; Sala & Gobet, 2017). It follows, that the overarching aim of work-based learning interventions is long-term knowledge retention to form the basis on which permanent transferable skills can be built (Billett, 2018; Billing et al., 2021; Blume et al., 2010; Kirkpatrick, 1967; Vargas, 2017).

Reverting to previously conducted spaced learning research, it has been outlined in depth that the best part of this research has taught and assessed factual knowledge rather than conceptual or procedural knowledge, which would rather be applicable for work-based learning environments. Further, the majority of experiments were set-up with relatively brief ISIs and RIs, which does not reflect the needs and requirements of work-based learning, in which long-term transfer and skill building is key to secure individual as well as organisational competitive advantage **(**Balota et al., 2011; Cepeda et al., 2006; Mozer et al., 2009**).**

Since the research at hand would like to add to the spaced learning as well as work-based learning literature and serve as guidance on how to design effective and efficient work-based learning interventions, whilst using complex factual, conceptual as well as procedural learning topics, an equal spaced learning setup, using RIs of two and four weeks were chosen to check by way of example if spaced learning can be applied to those knowledge dimensions. In doing so, the ISI-RI relationship as outlined in chapter 2.2.3.2 has been considered. Given that factual knowledge is on the lowest hierarchical level of Bloom's knowledge pyramid (see chapter 2.3.2), it is assumed that the spaced learning effect already occurs with relatively few learning sessions, separated by short ISIs, and assessed after short RIs (which has been shown in multiple studies throughout the last decades – for reference see chapter 2). Therefore, it is also assumed that more complex teaching materials (as used for this research), which require more demanding cognitive processing (as described in chapter 2.3.1), demand a higher number of learning repetitions at longer intervals for the spaced learning effect to become apparent. As a result, the following hypothesis H2 has been formulated, containing two sub-hypotheses:

> ***Hypothesis H2:*** *Learning outcomes derived from e-learning interventions instructing complex knowledge following an equal spaced learning design will be more pronounced with a) longer RIs and b) with increasing frequency of repetitions (with total time of learning being constant)*.

Additional streams of distributed practice research revealed that combining the spaced learning effect with the testing effect, i.e., the conscious testing of knowledge with the aim of knowledge retention, further enhances the proven spaced learning effect (Cogliano et al., 2019; Dunlosky et al., 2013; Morehead et al., 2015; Roediger & Karpicke, 2006a/b). By intentionally using testing as a learning strategy, forgetting of the subject or content learnt is minimised as individuals' encoding processes run much deeper and more comprehensive, therefore, retention is enhanced (Pashler et al., 2007; Rodriguez et al., 2021; Roediger & Karpicke, 2006a). Thereby it is said that retrieval practice leads to better knowledge retention, irrelevant of which spacing schedule is followed or whether corrective feedback on test performance is provided (Bahrick et al., 1993; Butler et al., 2008; Cepeda et al., 2009; Karpicke & Roediger, 2007a).

Although the research on retrieval practice has not established that corrective feedback would be necessary for additional learning success, spaced learning research providing corrective feedback revealed a remarkable impact on knowledge retention (Balota et al., 2006; Cepeda et al., 2006; Pashler et al., 2007). Interestingly, when it comes to corrective feedback, timing of when provided does not matter – results of studies on comparisons of immediate and delayed feedback showed that both help learners to understand their own knowledge gaps as well as conquer metacognitive misbeliefs of actual and perceived knowledge thereby diminishing forgetting (Black & William, 1998; Butler et al., 2007, 2008; Corral et al., 2021; Hattie & Timperley, 2007; Pashler et al., 2007). Additionally, it was found that when comparing both the expanding and the equal spacing condition including immediate testing and providing corrective feedback afterwards within both, a slightly higher advantage was found for the equal spacing condition (Karpicke & Roediger, 2007a).

Since Cepeda and associates (2006) called for additional spaced learning research using complex knowledge learning materials as research topic and concretely asking this research to use fixed schedules, providing feedback to the learners for them to create accurate awareness of their learning status, and business-related research on learning also highlighting the importance of providing immediate feedback on the learning process and guiding learners through the learning experience (e.g., Chelovechkov & Spar, 2019; Udemy, 2018), it would be of interest to understand if direct testing followed by immediate corrective feedback on test performance would enhance knowledge retention, bearing in mind the complexity of the learning materials as well as the e-learning environment. This also seems to be of relevance as it is precisely work-based learners who are asking to be tested and receive feedback in order to better understand where their knowledge and skill gaps are so that they can be addressed and continuously improve through learning (Tauber et al., 2019).

As a result, clearer understanding about the cause-effect relationship on how direct testing and feedback impacts subsequent knowledge retention of complex learning materials in an e-learning environment can serve as guidance on how to improve existing work-based learning offerings and how to better design future ones. Therefore, the following hypothesis H3 was set:

> ***Hypothesis H3****: Direct testing after an e-learning intervention followed by corrective feedback on errors will have an additional positive effect on overall learning outcomes within equal spaced learning schedules.*

Prior research demonstrated that learners perceive massed learning to be more efficient and effective than spaced learning or retrieval practice, although better retrieval performances were achieved in any spacing learning intervention (Bjork, 1999; Dunlosky et al., 2013; Morehead et al., 2015; Simon & Bjork, 2001; Vlach et al., 2019; Zechmeister & Shaughnessy, 1980). Falsified judgements of learning outcomes lead to ineffective study strategy selection such as massed instead of spaced studying and re-reading instead of self-testing or retrieval practicing (Ariel & Karpicke, 2018; Dunlosky & Rawson, 2015; Karpicke & Roediger, 2007a; Karpicke et al., 2009; Kornell & Bjork, 2007; Tullis et al., 2013). This appears to be even more critical in the field of work-based learning, in which managers and others in charge still largely follow approaches to learning and development which rely on traditional classroom learning methods inappropriate for work-based learning (Ellinger, 2004; Tauber et al., 2019). If individual learners as well as those designing learning interventions would be cognisant and confident that spaced learning conditions are better than massed learning conditions and benefit the overall learning process (Vlach et al., 2019), confidence in and application of spaced learning conditions is said to be enhanced (Ariel & Karpicke, 2018). It is plausible to assume that not only learners in simple, low-level knowledge spaced learning conditions rate massed learning superior to the first but also, and even more, learners in complex, high-level knowledge spaced learning conditions, found in real-life work-based learning environments. Thus, the following hypotheses H4 and H5 were set:

Hypothesis H4: *Most learners taking part in spaced e-learning interventions will report that they felt less self-confident with regards to predicted learning outcomes than learners in the massed e-learning intervention.*

Hypothesis H5: *Most learners taking part in spaced e-learning interventions will see a greater delta between perceived and real learning outcomes learning outcomes than learners in the massed e-learning intervention.*

Prior investigations into formal work-based learning call for transformative, adaptable learning strategies, which keep up with changing circumstances, leading to transferrable skills allowing for long-term application (Glaveski, 2019; Kane et al., 2018; Tauber et al., 2019). Transfer of knowledge and skills is key to competitive advantage for employees as well as organisations. Thus, developing and designing work-based learning interventions should inevitably consider biological and neuroscientific research findings, aiming at improving current and future job-related skills, whilst also creating opportunities for employees to incorporate learning into their day-to-day business and preventing that currently half of skills and knowledge taught are forgotten about one day after the intervention (Blanchard, 2013; Glaveski, 2019; Ryu & Moon, 2019; Tauber et al., 2019). If all the above would be considered in detail and transfer of what was learnt during formal work-based learning is actively reinforced, economic advantages resulting from education for individuals as well as organisations can be increased (Blume et al., 2010; Kirkpatrick, 1967). Yet, it was also shown that the majority of work-based learning interventions are developed and designed with little thought on what and how to train individuals (Tynjälä, 2008). Severe discrepancies exists between executive managers and learning and development departments, resulting in ineffective learning programmes being offered (Tauber et al., 2019) and furthermore, the holistic view of lifelong learning is disregarded (Beier, 2021; Dunlosky et al., 2013; Pashler et al., 2007; Tuijnman & Boström, 2002).

Research in the field of instructional design revealed that regardless of whether learning takes place online or offline, designers of learning interventions must adhere to human cognitive architecture (Mayer & Moreno, 2003; Sorden, 2005; Sweller et al., 2019). In doing so, cognitive load should be reduced and learners' interest should be fostered to free up working memory capacities (Issa et al., 2013; Mayer et al., 2004). As a result, meaningful learning is achieved (Mayer & Moreno, 1998).

According to this instructional design research, it is of further importance to base all decisions on how to design instruction on evidence-based procedures (Issa et al., 2013). Therefore, it appears highly relevant to gather insights into learner's satisfaction and preferences when it comes to work-based e-learning interventions, stemming from evidence-based practice. Eventually, these insights shall provide clarity on how to effectively design work-based e-learning interventions to maximise their impact. Based on the works of Mayer and Moreno (1998, 2002, 2003) on meaningful learning and multimedia instruction in e-learning, the following hypotheses H6 and H7, each with their three sub-hypotheses, were set:

Hypothesis H6:

a. *Most learners will report that their overall satisfaction with the e-learning experience decreases the more learning sessions they need to attend.*
b. *Most learners will report that their overall satisfaction with interactive learning formats is higher than with non-interactive ones.*
c. *Most learners will report that their overall satisfaction with guided learning formats is higher than with self-paced ones.*

Hypothesis H7:

a. *Most learners will report a preference for more spaced e-learning interventions in their work life.*
b. *Most learners will report a preference for interactive learning interventions over non-interactive ones.*
c. *Most learners will report a preference for guided learning interventions over self-paced ones.*

Finally, research has identified that the skills needed for companies and individuals to stay competitive in the future of work, such as digital, software, and technological literacy as well as cognitive, interpersonal, and self-leadership skills (Dondi et al., 2021) are predominantly to be assigned to the declarative and procedural knowledge categories. As outlined in chapter 2.1.1, facts and larger organisational structures built around them are declarative knowledge stored in the declarative memory, whereas cognitive and motoric skills are procedural knowledge, stored in the procedural memory (Roth, 2011). Declarative elements or information are learnt through conscious repetitive instruction and exposure, supported by cues to retrieve, and recall this theoretical knowledge from memory (Anderson, Krathwohl et al., 2001). To gain mastery in procedural knowledge, and to translate a theoretical 'what' into a practical 'how to', a lot of practice is needed (Anderson, Krathwohl et al., 2001).

To sum it up, declarative knowledge is based and derived from factual, theoretical statements, whereas procedural knowledge is formed and derived from practically doing something (Anderson, Krathwohl et al., 2001; Roth, 2011). As mentioned before, previous research in the field of spaced learning mainly explored its effects on the factual and, at times, on the conceptual knowledge dimensions (Carpenter et al., 2012; Kang & Pashler, 2012; Kapler et al., 2015; Kornell & Bjork, 2008; Rohrer, 2009; Rohrer & Taylor, 2006, 2007; Vlach et al., 2008). Studies examining effects of spaced learning on more application-oriented knowledge, such as procedural knowledge is not known. It is assumed that this is due to the fact that each type of knowledge is acquired differently: Whilst the first heavily depends on systematic, educational instruction and conscious cognitive attention of the learner, the second can also be learnt unconsciously by observing others, or through trial and error (Anderson, Krathwohl et al., 2001; Roth, 2011).

However, as noted above, it is not only factual- and conceptual-based skills like digital, software, and technological literacy that are lacking and needed to stay competitive in global competition and labour markets, but rather more procedural-based skills, such as cognitive, interpersonal, and self-leadership skills are of importance (Dondi et al., 2021). Since no study is known which examined if spaced learning is also applicable to the field of procedural knowledge, it would be interesting to understand, if any effects become evident, although it is assumed that this is not the case due to the different ways in which the two forms of knowledge are neurologically acquired.

Still, it is assumed that learners are going to assess relative learning outcomes higher for the training on procedural knowledge compared to the training on factual and conceptual knowledge, due to the better applicability and relation to real life, regardless of the instructional method used. This assumption results from the following two arguments: first, Schunk's (2012) definition of learning which states that learning is a change in behaviour resulting from experiences and practice and second, as procedural knowledge, as stated by Anderson, Krathwohl and associates (2001), is best learnt by trial and error and observation. By engaging in trial and error and practicing something that has just been learnt regularly, one could argue that learners' judge both their encoding and retrieval fluency (Finn & Tauber, 2015; Moore & Healy, 2008), as outlined in chapter 2.2.6, as high and thus will rate learning from practice and experience as better than learning from pure instruction only.

Therefore, and taking all of the above together, hypothesis H8 was formulated as follows:

> ***Hypothesis H8:*** *The effects of spaced learning are especially applicable in work-based learning interventions for factual and conceptual over those for procedural knowledge. Learners themselves, however, will perceive the training on procedural knowledge to have yielded a higher metacognitive learning success.*

Jointly the experimental review against these hypotheses will provide insights regarding the effectiveness of spaced learning and its' applicability for managerial, work-based, lifelong learning in the factual, conceptual, and procedural knowledge domains, thus aiming at answering the overarching research question. All the above will be addressed in two empirical experiments conducted for the research at hand and will be detailed in the following chapter 4.

4. Experimental research

To best address the research focus as just deducted, this research followed an applied evaluative multi-method quantitative approach. According to Saunders, Lewis, and Thornhill (2016) such type of research applies the research paradigm of positivism. For this research, two field experiments (Gerber, 2011) were conducted, gathering quantitative data in form of knowledge tests and surveys. This research concludes by giving pragmatic advice on how to integrate the findings in real-life management learning curricula.

Chapter 4.1 first summarises how the two experimental e-learnings conducted for this research were *designed*. Therein, chapter 4.1.1 focuses on the first *e-learning on "platform business models"* and chapter 4.1.2 focuses on the second *e-learning on "time management"*. From these experiments, profound recommendations shall be drawn for future work-based e-learning interventions. The subsequent chapters 4.2 and 4.3 represent the experimental reports of both experiments. Within chapter 4.4, the effects of both experiments are compared, aiming at answering hypothesis H8 of this research.

4.1 Design of experimental e-learnings

The two field experiments conducted were set-up, managed, and delivered via the e-learning platform of *University4Industry*. University4Industry (2021) is a training provider in the field of digitalisation that uses a blended learning format and offers a learning platform as well as content in different multimedia formats. University4Industry (2021) follows the learning approach of Learn, Explore, Discuss, and Act (LEDA). To be in line with the learning offerings of University4Industry, the e-learnings designed for this research were set up similarly, meaning self-paced learning content was provided on slides and with videos (*Learn*), at least one live session was scheduled and conducted for learners to recap and discuss what was taught during the self-paced learning sessions (*Discuss*), and a knowledge test was designed in the wake of acting (*Act*). Since it was not possible to generate a simulation of the learning content in this experiment, the *Explore* part of the learning approach was omitted. In the following, it is explained how content for the two experimental e-learnings was derived, designed, and planned.

4.1.1 E-learning on "platform business models"

The basic content of the e-learning "platform business models" originated from the lecture course "Digital Strategy" of the faculty of Management at Heinrich Heine University, Düsseldorf. The course on "platform business models" contained five sub-topics, which provided the structure for the to-be-developed e-learning. In addition to the provided slides, the faculty of Management at Heinrich Heine University, Düsseldorf, shot five videos which allowed learners to further explore the five individual topics by elaborating and repeating the content of the slides. Mainly factual as well as conceptual knowledge was taught within this course. Based on these materials, the following learning objectives of the e-learning programme were set:

1. Learners can recall the term "platform business models"
2. Learners can recall, interpret, summarise, compare, and explain different platform strategies and concepts
3. Learners can differentiate main concepts underlying platform business models
4. Learners can evaluate the appropriateness of different KPIs for platform business models

Afterwards, an assessment to test knowledge after the learning intervention was developed and set-up in Typeform (2021), an online survey tool (Appendix A). For spacing Group 4 an additional multiple choice test consisting of five questions was developed and set-up in ClassMarker (2021), an online testing website used for multiple choice testing (Appendix

B) Following previous research (e.g., Butler et al., 2007, 2008, 2013; Cogliano et al., 2019), this was used directly after the respective learning units and direct feedback was given on correct and incorrect answers. In Group 4, the aim was to test whether the immediate feedback had an additional influence on the learning success apart from the spacing itself.

The slides provided by the faculty were in English. As the videos shot by the faculty were in German and could not be shot again, all slides were translated into the German language and at the same time were redesigned to adhere to the multimedia design principles of cognitive load theory proposed by Mayer and Moreno (2003). Though, it must be mentioned that this e-learning was not set up to validate the theories of multimedia design or cognitive load theory. Rather those theories were used as guidelines to ease the learning process for participants and to minimise cognitive disruptors.

The live sessions for this e-learning were set-up in consultation with Prof. Dr Andreas Engelen. These were either scheduled for 60 minutes or twice for 30 minutes, depending on the experimental e-learning group. All sessions were set up the same way: in the first part of the session Prof. Dr Engelen repeated what was taught during the self-paced learning sessions, using another example to illustrate and deepen the concepts of platform business models again, and the second part of the sessions were open to participants to raise questions and discuss the topics. Prof. Dr Engelen was briefed prior to each session on the contents the individual groups had already seen, to be able to guide the session accordingly.

Once all contents and materials used for the e-learning programme on “platform business models” were agreed on and confirmed, the individual sessions per experimental e-learning group were spaced. First, RIs were set with two weeks for four groups and four weeks for two groups, following an equal spacing schedule. Based on the research findings of Cepeda and team (2008), ISIs of four days for the two week RI groups and seven days for the four week RI were defined. As the learning sessions were not supposed to take place on a Saturday, the ISI was shortened to three days in two cases for the two-week RI groups. The exact sequence of learning sessions held per group and date is shown in Appendix C.

To gather metacognitive judgments of learning, feedback surveys were developed via Typeform (Appendix D), containing JOL prompts such as: "*What percentage of the final test questions do you think you can answer correctly 2-4 weeks after the last training?*" for those surveys right after each learning intervention and "*How many of the last 15 test questions do you think you answered correctly?*", "*Do you think your specific training process had a positive effect on your personal learning success?*", and "*Which of the following types of training would be suitable for your professional environment?*" for the survey right after the final knowledge tests. This question format was chosen based on previous research (e.g., Cogliano et al., 2019; Tullis et al., 2013; Vlach et al., 2019). In the same surveys, questions were raised to gather insights on learners' satisfaction levels and preferences.

To allow for a smooth learning experience, all learning materials as well as after-learning-intervention surveys were accessible via the learning platform provided by University4Industry. Only the final knowledge test was sent to the learners as a link embedded in an email.

4.1.2 E-learning on "time management"

The idea to create an e-learning on "time management" came from observations of students at the chair of management, Heinrich Heine University, Düsseldorf, who had problems with their own personal time management. This resulted in a video series called "+beyond" to give students tools to address – among others – time management. In addition, a cooperation with the German Football Association took place, resulting in a seminar on "Time Management for Managers". Three of the +beyond videos on the topic of "time management", as well as the documents from a seminar with the German Football Association, served as the basis for the creation of the e-learning. Mainly procedural knowledge was taught during the training.

An additional introductory video was shot based on the seminar documentation and a set of slides was generated on the basis of each of the videos, again following the theories of Mayer and Moreno (2003) on multimedia design and cognitive load. However, as already mentioned for the first set of experiments, this e-learning too was not set-up to validate these theories. Rather those theories were used as guidelines to ease the learning process for participants and to minimise cognitive disruptors.

Based on the videos and slides, the following learning objectives of the e-learning programme were set:

1. Learners can recall different “time types”
2. Learners can recall and recognise different “time traps”
3. Learners can recognise the three “Ps” of time management theory
4. Learners can compare and differentiate different “time types”
5. Learners can explain which techniques to improve time management are attributed to which “time type” and to make judgements about the applicability of these

All the following steps were analogous to the first e-learning on the topic of “platform business models” (chapter 4.1.1):

- 12 assessment questions were developed based on the learning objectives and again set-up in Typeform (Appendix E)
- Spacing Group 4 was given an extra multiple-choice test, set-up in ClassMarker (Appendix F)
- Live sessions with Prof. Dr Engelen were set-up and conducted as in the first experiment
- The experimental groups were spaced, following an equal spacing schedule at the same inter-session and retention intervals as the first experimental e-learning (Appendix G)

- Surveys to gather metacognitive judgments of learning were developed and set-up in Typeform analogue to the questions raised in the first experiment (Appendix H)
- All learning content was made available via the learning platform of University4Industry except for the final knowledge test which was distributed via email.

4.2 Experiment 1: Factual and conceptual knowledge

The first field experiment of this research examined the effect the instructional method of spaced learning had on an e-learning programme on the topic of “platform business models”. Factual and conceptual knowledge (Anderson, Krathwohl et al., 2001) regarding the following five topics was taught: first, how platform business models differentiate from pipeline business models; second, how supply and demand interact for platform business models; third, how they are monetised; fourth, how economies of scale work for platform business models, and fifth, different success KPIs for platform business models for different growth stages. This topic was chosen as one example of factual and conceptual knowledge, as taught by the faculty of management, Heinrich Heine University, Düsseldorf. The experiment took place between September 27, 2021, and November 22, 2021.

4.2.1 Method

4.2.1.1 Sample

To determine the sample size of the experiment needed to be statistically significant, the optimal sample size was calculated with the following assumptions:

- Effect size: $\eta^2 = 0.06$; this medium effect was assumed as for this research to be of practical relevance and to be put into action any effect < 0.06 seems irrelevant for implementation (Cohen, 1988)

- Alpha level: 0.05
- Power: 0.8 (see Cepeda et al., 2006)

With seven groups, this results in a minimum number of participants of 32 people per group and 224 in sum (Hemmerich, 2018). A total of 226 participants signed up for the e-learning programme on “platform business models”, following self-selection sampling, i.e., those interested in the topic of “platform business models” enquired to participate. The training was advertised through several channels on LinkedIn, email communications at Heinrich Heine University, Düsseldorf, TU Dortmund, Ruhr University Bochum, University of Munster, and University of Applied Sciences Trier. The participants were then stratified-randomly assigned into one of seven experimental groups, whereby focus was put on all groups having roughly the same averages for age, gender, pre-experience with the taught topic, and pre-experience with e-learning programmes. This was necessary to ensure a meaningful comparison of all groups as no pre-test of the participants' abilities took place before the experiment. Table 9 summarises the characteristics of each group.

Table 9 – Characteristics of each experimental group, author's own compilation

Group	Gender [% female]	Gender [% male]	Age [years]	Work experience [years]	Pre-experience training topic [% yes]	Pre-experience e-learning [% yes]
	Characteristics (averages)					
Group 1	37	63	29	6	30	73
Group 2	43	57	29	6	37	84
Group 3	43	57	28	5	36	79
Group 4	48	52	30	7	39	82
Group 5	47	53	29	7	28	91
Group 6	37	63	29	6	40	97
Group 7	39	61	27	4	36	93
Average	42	58	29	6	35	85

Upon registration, participants were assured that participation in the training is voluntary, that they can drop out any time; and that all personal data collected will not be shared with anyone but the experiment team and are only used for randomisation purposes. Further, it was made clear that all data collected in the surveys and knowledge tests cannot be traced back to any individual. All of these ethical considerations were in line with the data protection policies of Heinrich Heine University, Düsseldorf (Appendix I).

213 participants finished all e-learning sessions as well as the knowledge test two weeks or four weeks after the last e-learning session and received a certificate of participation from the faculty of Management at Heinrich Heine University, Düsseldorf. 33 participants in Group 1, 32 participants each in Groups 2-5, 31 participants in Group 6, and 21 participants in Group 7 submitted the knowledge test after the e-learning intervention.

4.2.1.2 Design and materials

A between-subjects experimental design was used across the seven groups. The first group followed a massed learning condition, studying all learning content in one day. They then took one test after a retention interval of two weeks, serving as control group for Groups 2 to 5 and another test after a retention interval of four weeks, serving as control group for Groups 6 and 7. Groups 2 to 7 followed equal spaced learning conditions, whereby learning sequences, ISIs, and RIs varied. At the end of the experiment, learning outcomes at final test were compared between the massed group and the spacing groups as well as between the spacing groups.

Group 2-5 had an ISI of three, respectively four days each and a RI of two weeks. Group 2-4 had three study sessions in sum, whereby Group 2 started with a video for each of the five to-be-learnt topics in the first session, followed by five sets of slides (one per topic) in the second session and concluded with a live session covering all five topics, hosted by Prof. Dr Engelen. Group 3 followed a different learning sequence, starting with the five sets of slides, followed by the live session, and concluding with the five videos. Group 4 followed the same learning sequence as Group 2 with the difference that after each learning session, a multiple-choice test on the taught topics was conducted, on which immediate feedback was given. Group 5 had five study sessions in sum: the first session introduced the first and second topic via videos, in the second session the third video was introduced and the slides on topic one were made available. The third session entailed slides on sessions two and three and a live session was held on the topics one, two, and three. The fourth session introduced videos four and five and the learning programme concluded with a fifth session, entailing slides on topic four and five and a live session on these topics. Group 6 and 7 had an ISI of seven days and a RI of fourteen days. Thereby, Group 6 followed the same sequence as Group 2, and Group 7 followed the same sequence as Group 5. The design plan of the experiment is shown in Appendix C.

After each learning session, participants were asked to complete a short survey, gathering information on their judgement of learning, preference of learning media, and satisfaction with their learning experience (Appendix D). The final knowledge test conducted either two weeks or four weeks after the final learning session consisted of 15 questions. Furthermore, as part of the test, a survey was conducted, gathering information such as participant's judgement of learning, overall learning preferences, and satisfaction (Appendix A).

4.2.1.3 Procedure

After registration, participants received a "welcome email" to the e-learning programme and were informed about their learning schedule as well as their access details to the learning platform. The individual learning sessions were made available manually by the researcher in the mornings of the corresponding training days. Participants were informed by email that the learning session of the respective training days were available for 24 hours and were asked to complete the session within this time frame. Further, they were reminded about the upcoming training days. During the single training days, participants were allowed to complete the individual units of each session as they preferred. At the end of each session, participants were asked to complete a short, non-mandatory survey via the online tool Typeform (2021). During the survey, participants were asked to rate satisfaction ranging from very unsatisfied, unsatisfied, neither satisfied nor unsatisfied, satisfied and very satisfied. This data was transformed into a scale ranging from 1 to 5. Group 4 was further asked to take part in a five-question multiple-choice test via the online testing tool ClassMarker (2021). This test took on average three minutes to complete. After each test completion, the tool automatically fed back to the participants on which answers were right and wrong and gave explanations for each.

After every group's RI, participants received another email, asking them to take part in a final test, conducted via Typeform (2021). Within this test, 15 question were asked, four were multiple-choice questions assessing recognition and eleven were open questions assessing recall. Subsequently, participants were further asked to answer a survey in which 13 questions were asked on their JOL, learning satisfaction, preferences, and confidence.

Participants received one kind reminder to participate in the final test one day after the retention interval ended to ensure that as few participants as possible drop out of the field experiment and that statistical power per group was achieved. Since all tests were conducted anonymously and, for data security purposes, it was not possible to trace which participant took the final test, the final note of the survey asked participants to reach out to the researcher, informing her about their participation in the test. Upon this email, participants were thanked for their contribution to the learning research and were given an official certificate from the faculty of Management at Heinrich Heine University, Düsseldorf.

4.2.1.4 Analysis

All analyses presented in the following chapters 4.2.2.1-4.2.2.6 were calculated with the software "*The Jamovi project*", *Jamovi 2.2.5* (2021), using either a one-way or two-way analysis of variance (ANOVA), each with a confidence interval of 95%, corresponding to an alpha level of 0.05. In case of a statistically significant result, it was tested again with a post hoc test, using a Tukey-correction to determine which group caused the difference in means.

The results in section 4.2.2.7 are descriptive and were obtained from the evaluation of the final survey. In total, 213 knowledge tests including surveys were evaluated and used for data analysis. 15 participants left the experiment early.

4.2.2 Results

Within this chapter, results of the first experiment are presented. For a better overview, the following Table 10 summarises all hypotheses with the respective dependent and independent variables, as well as the corresponding results before each individual hypothesis is discussed.

Table 10 – Overview of research hypotheses and results experiment 1, author's own compilation

Hypothesis	Dependent variable	Independent variable(s)	Brief description of results
H1 – Massed learning versus spaced learning	▪ Points/ person at final test	▪ Instructional method (massed or spaced) ▪ RI (2 weeks and 4 weeks)	▪ Spaced learning leads to significantly better test results than massed learning ($p < 0.001$; $\eta^2 = 0.119$)
H2a/b – Spacing frequency	▪ Points/ person at final test	▪ 2a: RI (2 weeks and 4 weeks) ▪ 2b: Frequency of spaced learning sessions (3 or 5 sessions)	▪ 2a: Main effect of spaced learning leading to better test results than massed learning might only be valid at longer RIs, here: 4 weeks ▪ 2b: No statistically significant main effect detected; descriptive trend noted towards spaced learning with 5 sessions
H3 – Testing and immediate feedback	▪ Points/ person at final test	▪ Testing and feedback or no testing and feedback during learning process	▪ No statistically significant main effect detected ▪ Descriptive trend noted towards testing and feedback giving during learning process
H4 – Learner's self-confidence	▪ Confidence	▪ Number of lear-ning sessions (3 or 5 sessions) ▪ RI (2 weeks and 4 weeks)	▪ Confidence depends on RI ($p = 0.001$) ▪ Spaced learners with RI of 2 weeks reported highest confidence towards a future test ▪ Massed learners reported least confidence towards a future test
H5 – Learner's self-perception	▪ Delta real to perceived learning success	▪ Instructional method (massed or spaced) ▪ RI (2 weeks and 4 weeks)	▪ 2 significant main effects: instructional method ($p = 0.002$), RI ($p < 0.001$, $\eta^2 = 0.119$) ▪ Significant interaction effect: instructional method ($p < 0.001$, $\eta^2 = 0.009$) ▪ For learners with RI of 2 weeks: self-perception less positive for massed group than for spaced groups ▪ For learners with RI of 4 weeks: vice versa
H6 – Learner's satisfaction	▪ Satisfac-tion	▪ Instructional method (massed or spaced) ▪ Interactivity of learning session ▪ Guidance during learning session	▪ Satisfaction depends on instructional method ($p = 0.004$) ▪ RI also impacts satisfaction: the more sessions, the higher satisfaction ▪ Massed learners more satisfied with interactive, non-guided learning sessions ▪ Spaced learners more satisfied with non-interactive, guided learning sessions
H7 – Learner's preference	▪ Descriptive analysis, derived from the evaluation of survey after final test		▪ Overall, learners prefer spaced, interactive, guided learning sessions for future learning set-ups

4.2.2.1 Massed learning versus spaced learning (H1 and H2a)

Figure 18 shows the performance of participants on the final knowledge test (administered either two weeks or four weeks after the last learning session).

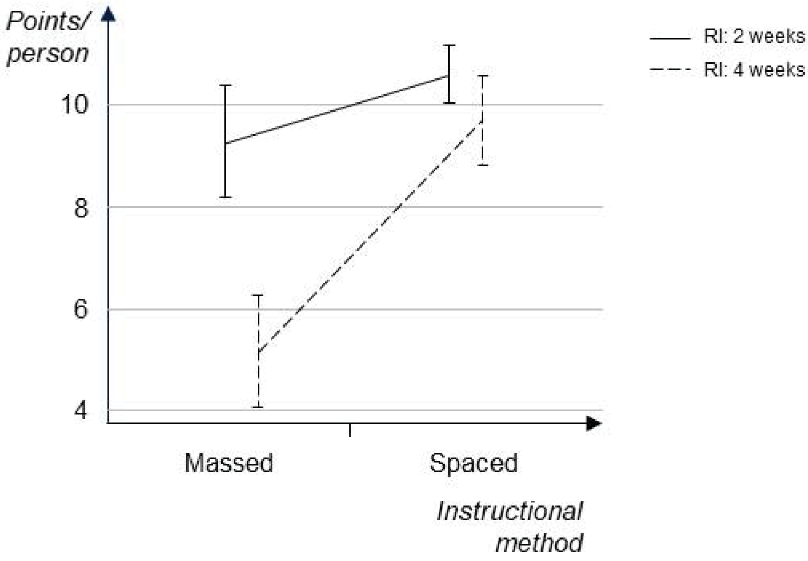


Figure 18 – Average performance of participants at final test; error bars represent 95% confidence intervals; adapted from Jamovi, 2022

As can be seen from the data in Table 11, the difference between spaced and massed learning is statistically significant with $p < 0.001$ and $\eta^2 = 0.119$ (average of 7.29 points/person for learners in the massed condition; average of 10.23 points/person for learners in the spaced condition). This effect size was calculated with a two-way ANOVA, using points/person at final test as dependent variable and instructional method and RI as independent variables.

Table 11 – Two-way ANOVA with points/person at final test as dependent variable, and instructional method and RI as independent variables, adapted from Jamovi, 2022

ANOVA – Points/person

	Sum of squares	df	Mean square	F	p	η^2
Instructional method	393	1	393.2	38.2	< 0.001	0.119
RI	287	1	287.0	27.8	< 0.001	0.087
Instructional method * RI	119	1	118.8	11.5	< 0.001	0.036

Yet, Figure 18 further shows that an increase in RI corresponds to a decrease in final test score (average of 9.77 points/person for spacing groups with a RI of four weeks) compared to the spacing groups with a RI of two weeks (average of 10.68 points/person). The benefit of spacing compared to massed learning at longer RIs is higher compared to shorter RIs (9.77 points/person vs. 5.18 points/person compared to 10.68 points/person vs. 9.39 points/person).

This interaction effect of instructional method and retention interval is disordinal (Table 12; see comparison massed 2 weeks vs. spaced 2 weeks and massed 4 weeks vs. spaced 4 weeks), meaning that the main effect of spaced learning leading to superior knowledge retention might only be valid under certain conditions, e.g., at longer RIs. Further, the Tukey-corrected post hoc test revealed a significant difference ($p_{tukey} < 0.001$) between massed learning with a RI of two weeks and massed learning with a RI of four weeks, with the learning outcomes of the former being superior to the learning outcomes of the latter. It further exposed that massed learning with a RI of four weeks is significantly different ($p_{tukey} < 0.001$) to spaced learning with both a RI of two and four weeks (Table 12).

*Table 12 – Tukey corrected post hoc test comparisons instructional method * RI, adapted from Jamovi, 2022*

Post hoc comparisons – Instructional method * RI

Comparison							
Instructional method RI	**Instructional method RI**	**Mean difference**	**SE**	**df**	**t**	**p_{tukey}**
Massed 2 w	vs. Massed 4 w	4.121	0.790	242	5.215	< 0.001
Massed 2 w	vs. Spaced 2 w	-1.322	0.627	242	-2.109	0.153
Massed 2 w	vs. Spaced 4 w	-0.428	0.714	242	-0.599	0.932
Massed 4 w	vs. Spaced 2 w	-5.443	0.627	242	-8.685	< 0.001
Massed 4 w	vs. Spaced 4 w	-4.549	0.714	242	-6.367	< 0.001
Spaced 2 w	vs. Spaced 4 w	0.849	0.528	242	1.694	0.329

Note: Comparisons are based on estimated marginal means

4.2.2.2 Spacing frequency (H2b)

Figure 19 represents the differences in test results at final tests of groups following the instructional method of spaced learning who had three learning sessions and groups following the same who had five learning sessions.

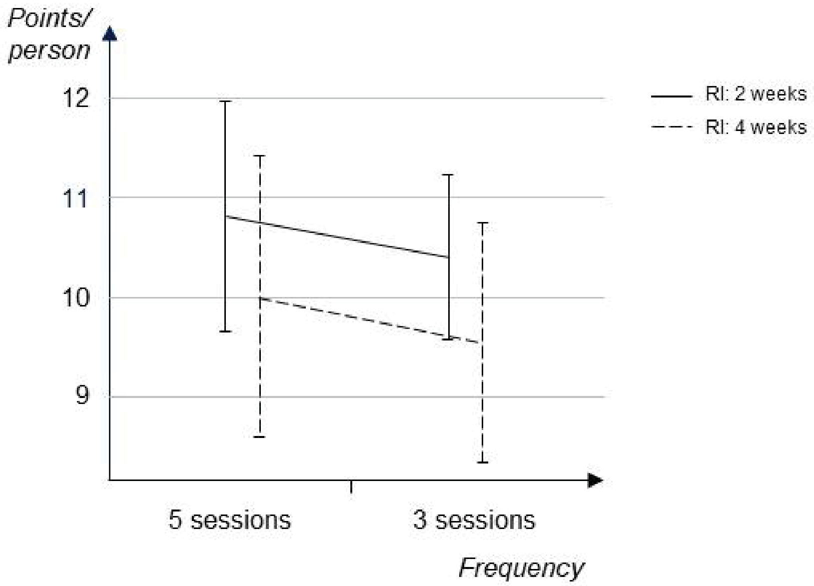


Figure 19 – Difference in test results of spaced learning groups with three and five learning sessions; error bars represent 95% confidence intervals; adapted from Jamovi, 2022

The statistical data of Table 13, again calculated with a two-way ANOVA using points/person at final test as dependent variable and frequency of learning sessions and RI as independent variables did not reveal a statistical main or interaction effect, with *p(frequency)* = 0.461, *p(RI)* = 0.152 and *p(frequency*RI)* = 0.969 each greater 0.05. However, with the given sample a descriptive trend is noted towards the spaced learning groups who had five learning sessions instead of three. Further, the data illustrates descriptively that the spacing effect of the groups with a RI of four weeks is lower than for those groups with a RI of two weeks.

Table 13 – Two-way ANOVA with points/person at final test as dependent variable and frequency of learning sessions and RI as independent variables, adapted from Jamovi, 2022

ANOVA – Points/person

	Sum of squares	df	Mean square	F	p	η^2
Frequency	5.8060	1	5.8060	0.54646	0.461	0.004
RI	22.0124	1	22.0124	2.07176	0.152	0.014
Frequency * RI	0.0162	1	0.0162	0.00153	0.969	0.000

4.2.2.3 Testing and immediate feedback (H3)

Figure 20 shows the difference in points per person at final test for spacing groups who were not tested right after a learning session and were given immediate feedback on the test and the group who was tested and was given immediate feedback.

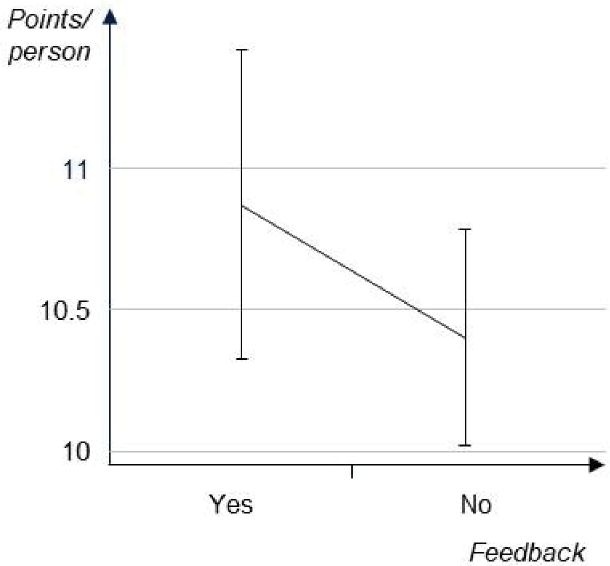


Figure 20 – Impact of direct testing and feedback on results on final test; error bars represent 95% confidence intervals; adapted from Jamovi, 2022

The statistical data of Table 14, calculated with a one-way ANOVA, using points/person at final test as dependent variable and direct testing and feedback or no testing and feedback as independent variables do not show any significant main effect ($p = 0.481$). Yet, the descriptive data of the plot above shows a positive trend that feedback might benefit the learning process.

Table 14 – One-way ANOVA with points/person at final test as dependent variable and feedback or no feedback as independent variables, adapted from Jamovi, 2022

ANOVA – Points/person

	Sum of squares	df	Mean square	F	p	η^2
Feedback	4.69	1	4.69	0.501	0.481	0.005

4.2.2.4 Learner's self-confidence (H4)

Figure 21 represents the confidence learners reported after their final learning session towards taking a knowledge test within the next two or four weeks.

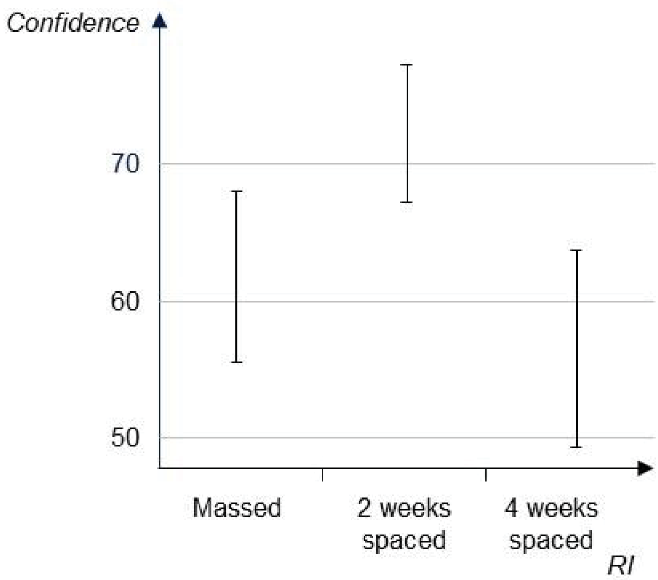


Figure 21 – Participant's reported confidence after each learning session; differences in confidence intervals might be explained by different group sizes/participants per learning session; error bars represent 95% confidence intervals; adapted from Jamovi, 2022

As can be seen from the data in Table 15, calculated with a one-way ANOVA using confidence as the dependent variable and RI as the independent variable, there is a statistically significant difference between the confidence of different groups with $p = 0.001$ and $\eta^2 = 0.069$.

Table 15 – One-way ANOVA using confidence as dependent variable and RI as independent variable, adapted from Jamovi, 2022

ANOVA – Confidence

	Sum of squares	df	Mean square	F	p	η^2
RI	7927	2	3964	6.88	0.001	0.069

A Tukey-corrected post hoc test (Table 16) further revealed that this statistical significance derives from differences between the two spacing groups, i.e., those with a RI of two weeks and those with a RI of four weeks, with $p_{tukey} = 0.002$. Between the two weeks spaced and massed groups, only a marginal significant difference can be reported. It could be the case that there are effects that are not visible in the small group size available and could potentially be if the groups would have had more participants.

Table 16 – Tukey-corrected post hoc comparisons RI, adapted from Jamovi, 2022

Post hoc comparisons – RI

Comparison						
RI	RI	Mean difference	SE	df	t	p_{tukey}
Massed	vs. 2 w spaced	-9.47	4.06	187	-2.33	0.054
Massed	vs. 4 w spaced	6.21	4.85	187	1.28	0.408
2 w spaced	vs. 4 w spaced	15.68	4.45	187	3.52	0.002

Note: Comparisons are based on estimated marginal means

4.2.2.5 Learner's self-perception (H5)

Figure 22 shows the delta of the perceived learning success and the actual learning success of the participants. From the plot for a RI of two weeks, the delta of perceived versus actual learning success is more negative for the massed group whereas it is more positive for the spaced groups. For the RI of four weeks, a different picture can be seen: the massed group assessed themselves much better than they actually are whereas the spaced group is again negative, yet very close to how their spaced counterparts with a RI of two weeks rated themselves.

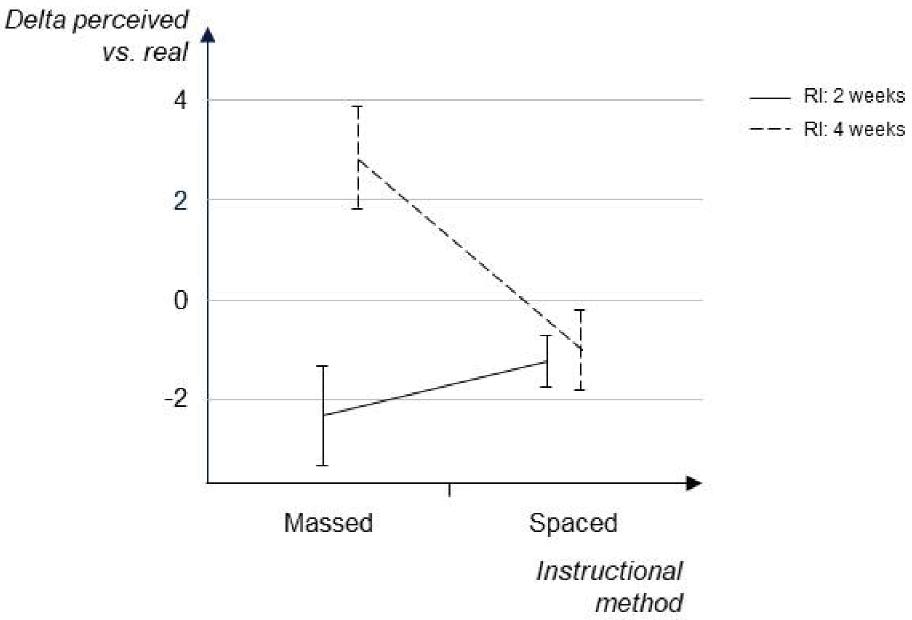


Figure 22 – Delta of perceived and actual learning success of participants; error bars represent 95% confidence intervals; adapted from Jamovi, 2022

Table 17 verifies this picture with two significant main effects: first, instructional method with $p = 0.002$ and $\eta^2 = 0.031$ and second, RI with $p < 0.001$ and $\eta^2 = 0.119$. Furthermore, a significant interaction effect between instructional method and RI, with $p < 0.001$ and $\eta^2 = 0.099$ was revealed. The data was submitted into a two-way ANOVA, using delta real-perceived learning success at final test as dependent variable and instructional method as well as RI as independent variables.

Table 17 – Two-way ANOVA with delta perceived vs. real learning success at final test as dependent variable and instructional method as well as RI as independent variables, adapted from Jamovi, 2022

ANOVA – Delta perceived vs. real learning success

	Sum of squares	df	Mean square	F	p	η^2
Instructional method	85.6	1	85.62	9.78	0.002	0.031
RI	329.1	1	329.14	37.59	< 0.001	0.119
Instructional method * RI	274.6	1	274.59	31.36	< 0.001	0.099

A Tukey-corrected post hoc test for the main and interaction effects mentioned above confirmed the significant effects as follows (for complete analysis refer to Appendix J): for instructional method (massed vs. spaced) with $p_{tukey} = 0.002$, for RI (two weeks vs. four weeks) $p_{tukey} < 0.001$ and three interaction effects between instructional method and RI. The three statistically significant interaction effects are first, between massed learning two weeks versus spaced learning four weeks with $p_{tukey} < 0.001$, second, massed learning four weeks versus spaced learning two weeks: $p_{tukey} < 0.001$, and third, massed learning four weeks versus spaced learning four weeks: $p_{tukey} < 0.001$.

In addition to this, a three-way repeated measures ANOVA (instructional method, RI, learning success (perceived results vs. actual learning success)) was calculated which revealed similar results (Appendix J). A descriptive underlining of these results can be found in Figure 23.

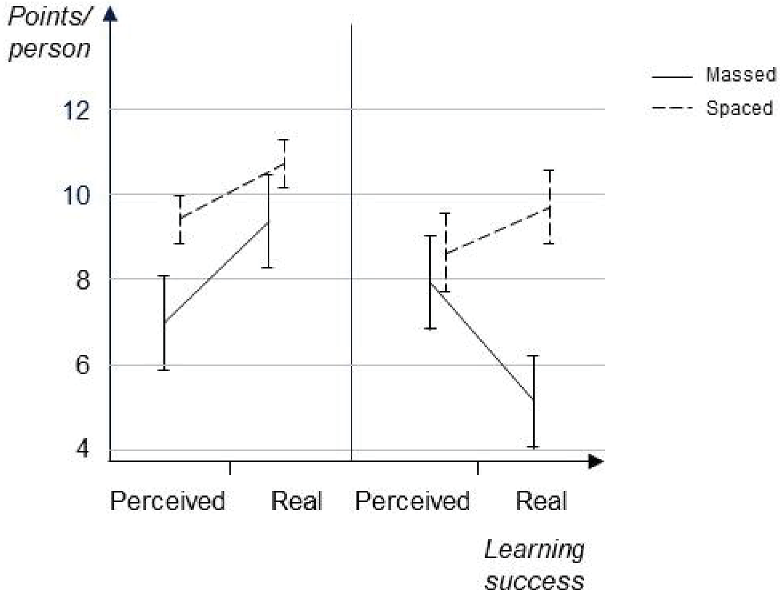


Figure 23 – Learning success (actual vs. perceived) of the participants with additional independent variables instructional method and retention interval; error bars represent 95% confidence intervals; adapted from Jamovi, 2022

4.2.2.6 Learner's satisfaction (H6)

Figure 24 presents the cumulated satisfaction ratings participants gave after each learning session they attended. The data in the figure below show that participants in the spaced learning groups attending five learning sessions recorded highest satisfaction levels and participants in the massed learning group recorded lowest satisfaction scores.

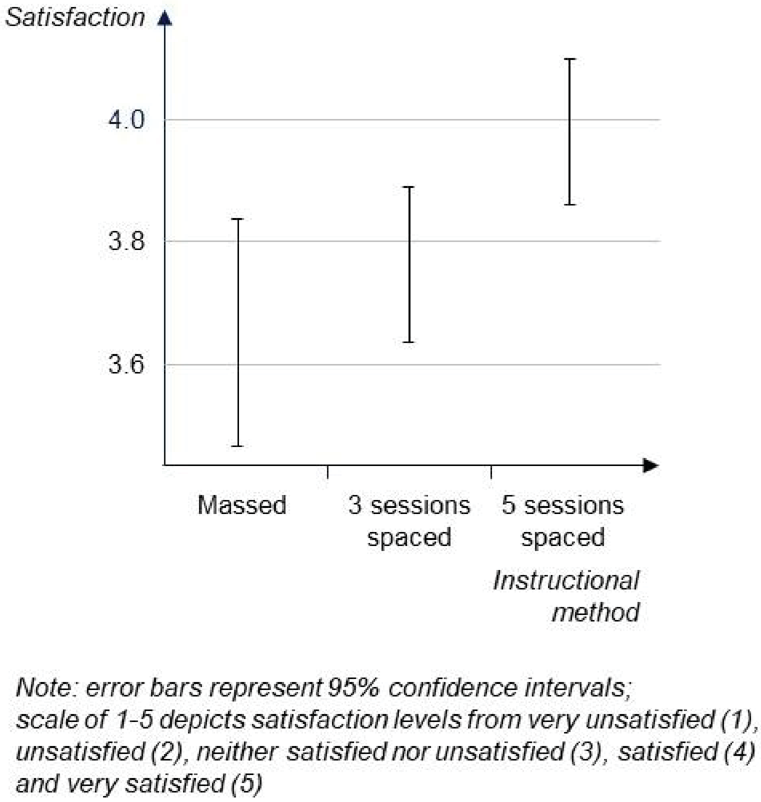


Figure 24 – Cumulated satisfaction ratings after each learning session; adapted from Jamovi, 2022

The data in Table 18 confirms the descriptive plot above: satisfaction of participants taking part in the experiment was dependent on the instructional method and number of learning sessions of their respective group, with p = 0.004. The data was analysed with a one-way ANOVA, using satisfaction as dependent variable and instructional method as independent variable. Yet, this statistical main effect merely states that all three groups do not have the same mean value of satisfaction.

Table 18 – One-way ANOVA with satisfaction as dependent variable and instructional method as independent variable, adapted from Jamovi, 2022

ANOVA – Satisfaction

	Sum of squares	df	Mean square	F	p	η^2
Instructional method	11.8	2	5.90	5.50	0.004	0.016

A Tukey-corrected post hoc test within this ANOVA (see Table 19) shows that both the difference between the massed groups and the five-session spaced learning groups is significant with p = 0.009 and the difference between the three-session spaced learning groups and the five-session spaced learning groups is statistically significant with p = 0.035.

Table 19 – Tukey-corrected post hoc test of experimental groups, adapted from Jamovi, 2022

Post hoc comparisons – Instructional method

Comparison							
Instructional method		Instructional method	Mean difference	SE	df	t	p_{tukey}
Massed	vs.	Spaced 3 sessions	-0.112	0.1142	679	-0.984	0.587
Massed	vs.	Spaced 5 sessions	-0.330	0.1118	679	-2.952	0.009
Spaced 3 sessions	vs.	Spaced 5 sessions	-0.218	0.0876	679	-2.484	0.035

Note: Comparisons are based on estimated marginal means

Two further influences on the participant's satisfaction levels were assessed: The interactivity of the course and whether a learning session was self-paced or guided. In that sense, two one-way ANOVA were undertaken, one with interactivity as independent variable and one with guidance as independent variable. Since all interactive learning interventions were also the guided ones, all calculations listed here for interactivity apply equally to guidance. For further analysis refer to Appendix L. Both investigations did not reveal any statistical relevant results, as shown by the example of interactivity, with p = 0.902 (Table 20).

Table 20 – One-way ANOVA with satisfaction as dependent variable and interactivity as independent variable, adapted from Jamovi, 2022

ANOVA – Satisfaction

	Sum of squares	df	Mean square	F	p	η^2
Interactivity	0.0164	1	0.0164	0.0151	0.902	0.000

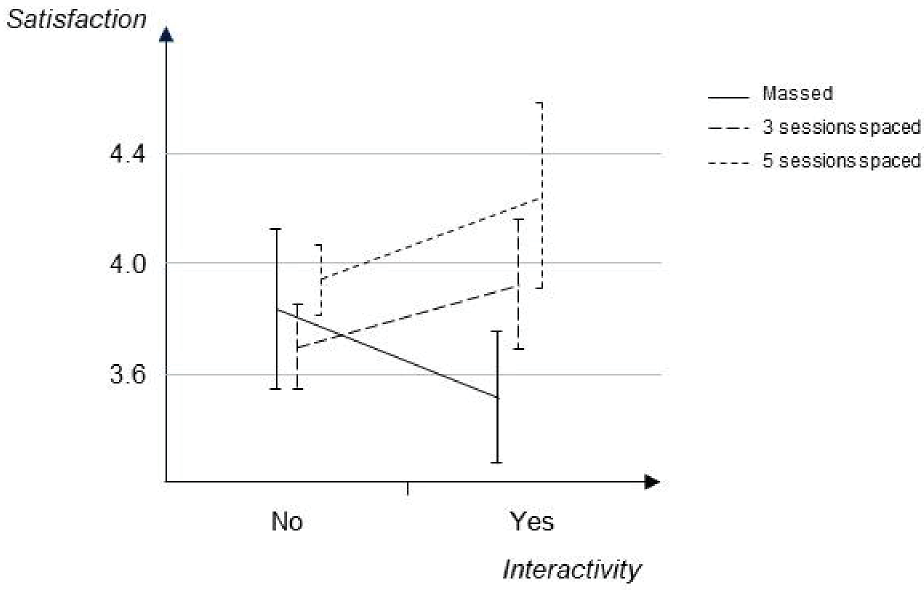
Yet, when conducting a two-way ANOVA, using satisfaction as the dependent variable and instructional method and interactivity as the independent variables, a statistically significant interaction effect of instructional method and interactivity with $p = 0.031$ was revealed (Table 21). Further, the main effect of instructional method as discussed earlier was confirmed with $p = 0.005$.

Table 21 – Two-way ANOVA with satisfaction as dependent variable and instructional method and interactivity as independent variables, adapted from Jamovi, 2022

ANOVA – Satisfaction

	Sum of squares	df	Mean square	F	*p*	η^2
Instructional method	11.509	2	5.754	5.405	0.005	0.016
Interactivity	0.469	1	0.469	0.441	0.507	0.001
Instructional method * Interactivity	7.405	2	3.703	3.478	0.031	0.010

Figure 25 below illustrates these results: participants in the massed learning group were least satisfied with interactive learnings sessions, whereas participants in the spaced learning groups and especially those having five learning sessions were most satisfied with interactive learning sessions, followed by those learners in the three-session spaced learning groups. Participants in the three-session spaced learning groups were (on a descriptive level) least satisfied with non-interactive learning sessions, yet no statistically significant effect has been found in terms of satisfaction with non-interactive learning sessions.



Note: error bars represent 95% confidence intervals; scale of 1-5 depicts satisfaction levels from very unsatisfied (1), unsatisfied (2), neither satisfied nor unsatisfied (3), satisfied (4) and very satisfied (5)

Figure 25 – Satisfaction of participants attending interactive learning formats; adapted from Jamovi, 2022

The same two-way ANOVA was conducted using guidance as well as instructional method as independent variables and satisfaction as the dependent variable. All effects were the same and are shown in Appendix L. Both statistically significant interaction effects were confirmed by two Tukey-corrected post hoc tests (Appendix L), which showed significant differences between the following groups (and can be applied also to guidance):

- Massed learning group, receiving an interactive learning session vs. spaced learning groups with five learning sessions in total, receiving no interactive learning sessions ($p_{tukey} = 0.025$)
- Massed learning group, receiving an interactive learning session vs. spaced learning groups with five learning sessions in total, receiving interactive learning sessions ($p_{tukey} = 0.007$)

- Spaced learning groups with three learning sessions in total, receiving no interaction vs. spaced learning groups with five learning sessions in total, receiving interactive learning sessions ($p_{tukey} = 0.040$)

4.2.2.7 Learner's preference (H7)

Before evaluating this hypothesis, it should be noted that when adding up the percentages, the result is not 100 percent because there were participants who either did not give an answer or did not stay until the end of the experiment. This is the case for all the analyses below.

When asked about whether participants prefer a spaced learning approach ("*more shorter sessions*") over a massed one ("*one longer session*"), both the massed and spacing groups believed a spaced learning approach would be more effective in terms of long-term knowledge retention (on average 4.11 vs. 2.42 and a delta 1.09-1.97 points, each on a 5-scale) compared to a massed one. Further, when asked about their desire for future work-based learning setups, participants clearly voiced preference for more spaced learning sessions. Only 46 percent see it as their current practice against 78 percent who would prefer it. This is mirrored by a low desire for one longer learning session (= massed). Here, 14 percent voiced preference for it, against 40 percent who are currently learning in this way. The descriptive data is shown in Figure 26.

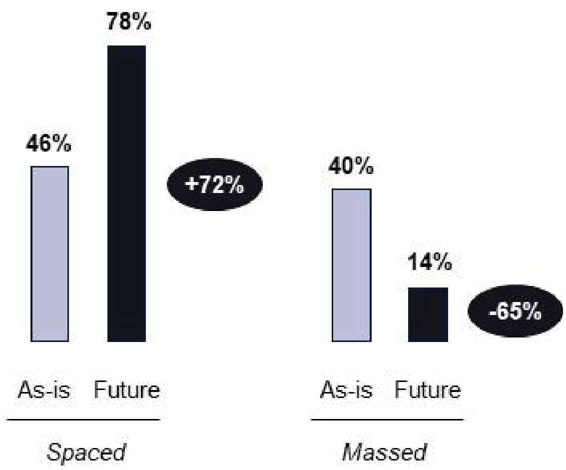


Figure 26 – Participants' preference of spaced learning versus massed learning, author's own compilation

When asked about whether participants would prefer interactive or non-interactive learning sessions in the future, a clear desire for interactive sessions was stated with 66 percent of participants asking for it. This is an increase of 72 percent compared to 38 percent of participants receiving mainly interactive learning interventions currently. Only 4 percent asked for non-interactive learning sessions in their future learning, which is a drop of 80 percent compared to 19 percent of participants receiving mainly non-interactive learning interventions currently. The descriptive data is shown in Figure 27.

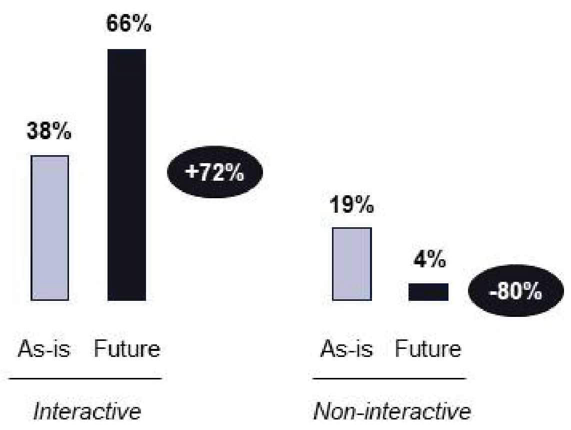


Figure 27 – Participants' preference of interactive versus non-interactive learning formats, author's own compilation

When asked about their preference of guided learning sessions versus self-paced learning sessions, participants recorded that current learning setups encourage self-paced learning (48 percent) over guided ones (26 percent). However, when asked about how they envision future learning setups to look, they preferred guided learning sessions over self-paced ones (preference for self-paced learning interventions dropped by 12 percent to 42 percent and preference for guided learning interventions increased by 16 percent to 30 percent). The descriptive data is shown in Figure 28.

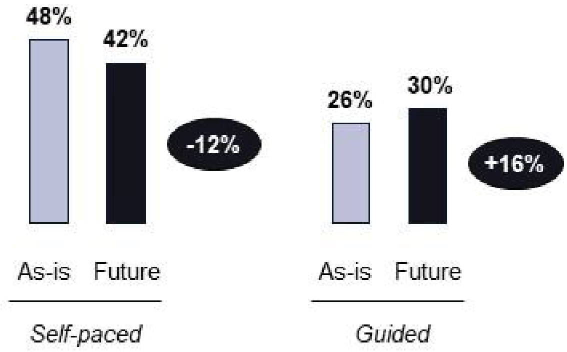


Figure 28 – Participants' preference of self-paced versus guided learning formats, author's own compilation

4.2.3 Discussion

The purpose of this experiment was to answer the leading research question whether the instructional method of spaced learning causes better learning in a work-based e-learning environment. Results indicate that work-based relevant factual and conceptual knowledge is associated with better learning compared to massed learning in an e-learning environment, in accordance with the overarching hypothesis H1. Therefore, it appears that spaced learning is not only applicable to laboratory studies, focusing on verbal and simple fact learning (e.g., Bird, 2010; Cepeda et al., 2008; Delaney et al., 2010; Dempster, 1989; Melton, 1970; Vlach et al., 2019) but also in real work-based learning environments that focus on complex concept learning. Especially at increasing RIs relevant for lifelong learning, spaced learning showed a large effect on knowledge retention, which is in line with hypothesis H2a.

Within this experiment, no statistical evidence has been found to claim that testing and immediate feedback as well as a higher frequency of learning sessions leads to even better knowledge retention as found in previous experiments (e.g., Cogliano et al., 2019; Dunlosky et al., 2013; Roediger & Karpicke, 2006a/b). Yet, descriptive trends towards confirming these hypotheses are given. It appears that these effects might be small in comparison to the effect of spaced learning (vs massed learning). If so, larger sample sizes might be necessary to assess the statistical significance of these effects and further investigations are needed to approve or disapprove both hypothesis H2b and hypothesis H3.

Interestingly, this experiment yielded different results in terms of learners' confidence than previous experiments did (e.g., Bjork, 1999; Dunlosky et al., 2013; Morehead et al., 2015): learners in the spaced learning intervention with a RI of two weeks felt more confident towards a future test than their counterparts with a RI of four weeks or those in a massed learning intervention. Even though no statistically significant difference exists between the spaced learning groups and the massed learning group, descriptive trends indicate this. Therefore, and as discussed before, larger sample sizes might be necessary to examine this effect and to understand why this experiment yielded the exact opposite finding of previous research and what influence the RI has on learners' confidence towards a future test. At this stage, hypothesis H4 cannot be supported by the research at hand.

Also, different results in terms of learners' self-perception on their learning success were yielded than previous research (e.g., Simon & Bjork, 2001; Vlach et al., 2019; Zechmeister & Shaughnessy, 1980): the spacing groups with the shorter RI of two weeks had a higher self-perception with regards to learning outcomes than the massed control group, still underconfident. For the RI of four weeks, the same picture as in previous experiments showed: the spacing groups judged their own learning as inferior compared to the massed control learning group. Still, their judgement of learning was very close to actual test results. Even though hypothesis H5 cannot be supported by this investigation, it is proposed to further examine the effects demonstrated in this research by differentiating between different RIs.

Within this experiment, participants attending the most learning sessions, i.e., those participants in spacing groups with five learning sessions, were the most satisfied ones. Participants in the massed learning groups where least satisfied. Both as opposed to hypothesis H6a. Further, it could not be proven that either interactive or guided learning sessions had an additional positive effect on participants' satisfaction, as captured in the works of Mayer and Moreno (1998, 2002, 2003). Thus, no evidence for the existence of a main effect of hypotheses H6a, H6b, and H6c could be shown. Yet, what could be shown is that a more differentiated view is more applicable as it appears that satisfaction levels might be dependent on the interaction of both instructional method and learning format: massed learners where more satisfied with non-interactive and self-paced learning interventions whereas the spacing groups scored satisfaction higher for interactive and guided learning interventions.

Lastly, it can be summarised that participants of the given experiment generally prefer spaced, interactive as well as guided learning interventions in their work life, which is in line with hypothesis H7. Even though preferences for self-paced learning formats were higher than for guided sessions, participants asked for less self-pacing and more guidance in their preferred future learning.

Therefore, and taking all the above together, it is assumed that the instructional method of spaced learning and interactive and guided forms of learning would be well accepted by learners, especially if, as in this experiment, they are participants who have an interest in the subject matter. Spaced learning in itself does not only have a positive impact on long-term knowledge retention but is also perceived by learners as a better instructional method. This should be fully considered in the future design of work-based learning interventions.

In sum, the given experiment concludes the following:

- Hypothesis H1 as well as hypothesis H2a is supported by the investigation.
- Hypotheses H2b and H3 are lacking strong evidence within the data but seem plausible on a descriptive level. Further investigations herein are needed.
- Hypothesis H4 is not supported by the investigation. Further investigations herein are needed, especially how different RIs impact confidence.

- Hypothesis H5 is not supported by the investigation and needs to be considered in greater detail, differentiating between different RIs.
- Hypothesis H6 cannot be supported on the level of main effects, but a more differentiated view seems appropriate. Again, further investigations are needed.
- Hypothesis H7 is supported by the investigation.

4.3 Experiment 2: Procedural knowledge

The second field experiment of this research examined the effect the instructional method of spaced learning had on an e-learning programme on the topic of "time management". Procedural knowledge (Anderson, Krathwohl et al., 2001) on how to use one's time better and more effectively, depending on which "time person" someone is, was taught. Thereby, deep dives took place into common time-traps and how to overcome these depending on each of the three time-types. This topic was chosen as one example of procedural knowledge, as taught by the faculty of management, Heinrich Heine University, Düsseldorf. The experiment took place between May 2, 2022, and June 27, 2022.

4.3.1 Method

4.3.1.1 Sample

As for the first experiment, the optimal sample size for this experiment was determined based on these assumptions:

- Effect size: $\eta^2 = 0.06$; this medium effect was assumed as for this research to be of practical relevance and to be put into action any effect < 0.06 seems irrelevant for implementation (Cohen, 1988)
- Alpha level: 0.05
- Power: 0.8 (see Cepeda et al., 2006)

Also resulting in a minimum number of 32 participants per group for seven groups (Hemmerich, 2018). However, the required 224 participants for the experiment were not achieved. A total of 211 participants signed up for the e-learning programme on “time management”, again following self-selection sampling. The participants were then stratified-randomly assigned into one of seven experimental groups, whereby focus was put on all groups having roughly the same averages for age, gender, previous experience with the taught topic, and previous experience with e-learning programmes. This was necessary to ensure a meaningful comparison of all groups as no pre-test took place. Table 22 summarises the characteristics of each group.

Table 22 – Characteristics of each experimental group, author's own compilation

Group	Characteristics (averages)					
	Gender [% female]	Gender [% male]	Age [years]	Work experience [years]	Pre-experience training topic [% yes]	Pre-experience e-learning [% yes]
Group 1	46	54	34	11	58	92
Group 2	52	48	25	3	45	94
Group 3	50	50	26	4	47	91
Group 4	47	53	27	4	50	91
Group 5	45	55	29	6	48	93
Group 6	58	42	35	12	58	96
Group 7	42	58	30	7	45	94
Average	49	51	29	6	50	93

192 participants finished all e-learning sessions as well as the knowledge test two weeks or four weeks after the last e-learning session and received a certificate of participation from the faculty of Management at Heinrich Heine University, Düsseldorf. 28 participants in Group 1, 27 participants in each Groups 2-6, and 29 participants in Group 7 submitted the knowledge test after the e-learning intervention.

4.3.1.2 Design and materials

The design of the second experiment was adapted from the first experiment, as it was aimed to understand if what holds true for factual and conceptual knowledge is also applicable for procedural knowledge. The design plan of this experiment is shown in Appendix G.

As was the case in the first experiment, participants were asked to complete a short survey, gathering information on their judgement of learning, preference of learning media, and satisfaction with their learning experience after each session. In addition, participants were also asked about their level of motivation towards the training to gain better understanding on why some participants dropped out of the learning programme. In the survey participants were again asked to rate satisfaction ranging from very unsatisfied, unsatisfied, neither satisfied nor unsatisfied, satisfied and very satisfied (Appendix H). This data was transformed into a scale ranging from 1 to 5 (1 = very unsatisfied, 5 = very satisfied). The final knowledge test conducted either two or four weeks after the final learning session consisted of 12 questions. Again, as part of the test, a survey was conducted, gathering information such as participants' judgement of learning, overall learning preferences, and satisfaction (Appendix E).

4.3.1.3 Procedure

This experiment mainly followed the same procedure as experiment 1 as outlined in chapter 4.2.1.3. Two deviations from experiment 1 are to mention: firstly, the survey conducted after each learning session included a question on how motivated participants were with regards to the learning session they attended. They rated motivation ranging from very motivated, motivated, neither motivated nor unmotivated, unmotivated, and very unmotivated. Secondly, the final knowledge test consisted of 12 questions instead of 15. Out of these 12 questions, four were multiple-choice assessing recognition and eight were open questions assessing recall.

4.3.1.4 Analysis

This experiment followed the same analysis as experiment 1 on factual knowledge as outlined in chapter 4.2.1.4. In total, 192 knowledge tests including surveys were evaluated and used for data analysis. 19 participants left the experiment early.

4.3.2 Results

Within this chapter, results of the second experiment are presented. For a better overview, the following Table 23 summarises all hypotheses with the respective dependent and independent variables, as well as the corresponding results, before each individual hypothesis is discussed.

Table 23 – Overview of research hypotheses and results experiment 2, author's own compilation

Hypothesis	Dependent variable	Independent variable(s)	Brief description of results
H1 – Massed learning versus spaced learning	▪ Points/ person at final test	▪ Instructional method (massed or spaced) ▪ RI (2 weeks and 4 weeks)	▪ No statistically significant difference / main effect detected
H2a/b – Spacing frequency	▪ Points/ person at final test	▪ 2a: RI (2 weeks and 4 weeks) ▪ 2b: Frequency of spaced learning sessions (3 or 5 sessions)	▪ 2a: No statistically significant difference / main effect detected ▪ 2b: No statistically significant difference / main effect detected
H3 – Testing and immediate feedback	▪ Points/ person at final test	▪ Testing and feedback or no testing and feedback during learning process	▪ No statistically significant difference / main effect detected
H4 – Learner's self-confidence	▪ Confidence	▪ Number of lear-ning sessions (3 or 5 sessions) ▪ RI (2 weeks and 4 weeks)	▪ No statistically significant difference / main effect detected ▪ Descriptive trend towards spaced learning groups with RI of 2 weeks most confident plausible
H5 – Learner's self-perception	▪ Delta real to perceived learning success	▪ Instructional method (massed or spaced) ▪ RI (2 weeks and 4 weeks)	▪ No statistically significant main effect detected, however statistical significant interaction effect (instructional method and RI (p = 0.050)) ▪ All learners underestimated real learning strongly
H6 – Learner's satisfaction	▪ Satisfac-tion	▪ Instructional method (massed or spaced) ▪ Interactivity of learning session ▪ Guidance during learning session	▪ No statistically significant difference / main effect detected ▪ Descriptive trend towards massed learners being least satisfied plausible
H7 – Learner's preference	▪ Descriptive analysis, derived from the evaluation of survey after final test		▪ Overall, learners prefer spaced, interactive, guided learning sessions for future learning set-ups

4.3.2.1 Massed learning versus spaced learning (H1 and H2a)

Figure 29 shows the performance of participants on the final knowledge test (administered either two weeks or four weeks after the last learning session). Based on the figure, no significant difference between the variables can be observed.

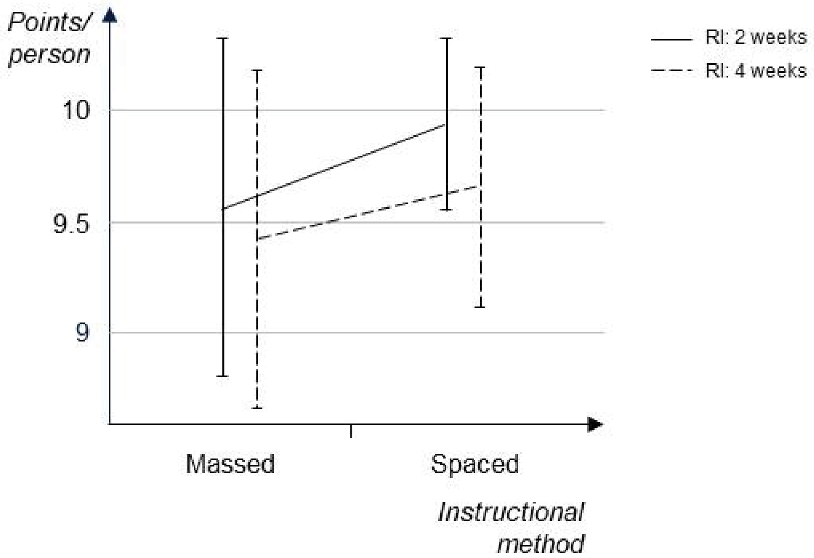


Figure 29 – Average performance of participants at final test; error bars represent 95% confidence intervals; adapted from Jamovi, 2022

This is also verified by the data in Table 24, calculated with a two-way ANOVA, using points/person at final test as the dependent variable and instructional method and RI as independent variables, which does not reveal any statistical main or interaction effect, with *p(instructional method)* = 0.348, *p(RI)* = 0.508 and *p(instructional method*RI)* = 0.827, which is each greater than 0.05. Based on the data gathered, it appears that the hypothesised effects that spaced learning leads to better learning than massed learning and that this effect is pronounced with longer RIs do not exist.

Table 24 – Two-way ANOVA with points/person at final test as dependent variable, and instructional method and RI as independent variables, adapted from Jamovi, 2022

ANOVA – Points/person

	Sum of squares	df	Mean square	F	p	η^2
Instructional method	3.716	1	3.716	0.8861	0.348	0.004
RI	1.847	1	1.847	0.4403	0.508	0.002
Instructional method * RI	0.201	1	0.201	0.0480	0.827	0.000

4.3.2.2 Spacing frequency (H2b)

Figure 30 represents the differences in test results at final tests of groups following the instructional method of spaced learning who had three learning sessions and groups following the same who had five learning sessions.

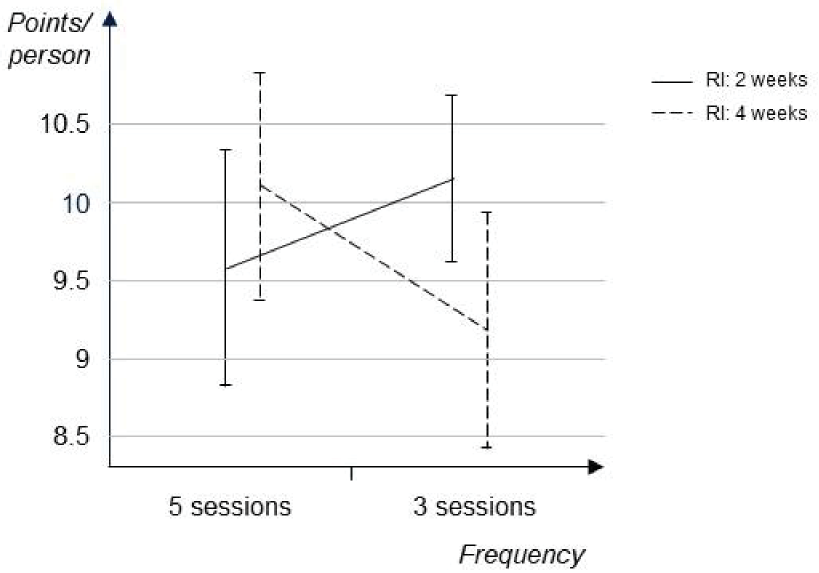


Figure 30 – Difference in test results of spaced learning groups with three and five learning sessions; error bars represent 95% confidence intervals; adapted from Jamovi, 2022

The statistical data of Table 25, again calculated with a two-way ANOVA using points/person at final test as the dependent variable and frequency of learning sessions and RI as independent variables did not reveal a main effect, with *p(frequency)* = 0.608, *p(RI)* = 0.525, both of which are greater than 0.05. Yet, a statistically significant interaction effect between frequency and RI was detected, with $p = 0.039$ and $\eta^2 = 0.032$. From the available data it appears that even though, on average it makes no difference whether one increases the RI from two to four weeks or has three or five learning sessions, there is a correlation in a common observation: for a RI of four weeks, it appears that five repetitions are better and for a RI of two weeks three repetitions seem to lead to higher impact.

Table 25 – Two-way ANOVA with points/person at final test as dependent variable and frequency of learning sessions and RI as independent variables, adapted from Jamovi, 2022

ANOVA – Points/person

	Sum of squares	df	Mean square	F	p	η^2
Frequency	1.04	1	1.04	0.264	0.608	0.002
RI	1.61	1	1.61	0.410	0.523	0.003
Frequency * RI	17.09	1	17.09	4.354	0.039	0.032

However, a Tukey-corrected post hoc test (Table 26) did not reveal any significant differences with each p_{tukey} greater than 0.05, and therefore, no interpretation can be drawn from the data.

*Table 26 – Tukey-corrected post hoc test comparisons frequency * RI, adapted from Jamovi, 2022*

Post hoc comparisons – Frequency * RI

Comparison									
Frequency	**RI**		**Frequency**	**RI**	**Mean difference**	**SE**	**df**	**t**	p_{tukey}
5 sessions	2 w	vs.	5 sessions	4 w	-0.5109	0.530	133	-0.9641	0.770
5 sessions	2 w	vs.	3 sessions	2 w	-0.5556	0.467	133	-1.1896	0.634
5 sessions	2 w	vs.	3 sessions	4 w	0.4074	0.539	133	0.7555	0.874
5 sessions	4 w	vs.	3 sessions	2 w	-0.0447	0.456	133	-0.0980	1.000
5 sessions	4 w	vs.	3 sessions	4 w	0.9183	0.530	133	1.7330	0.311
3 sessions	2 w	vs.	3 sessions	4 w	0.9630	0.467	133	2.0620	0.171

Note: Comparisons are based on estimated marginal means

4.3.2.3 Testing and immediate feedback (H3)

Figure 31 shows the difference in points per person at final test for spacing groups who were not tested right after a learning session and were given immediate feedback upon the test and the group who was tested and was given immediate feedback.

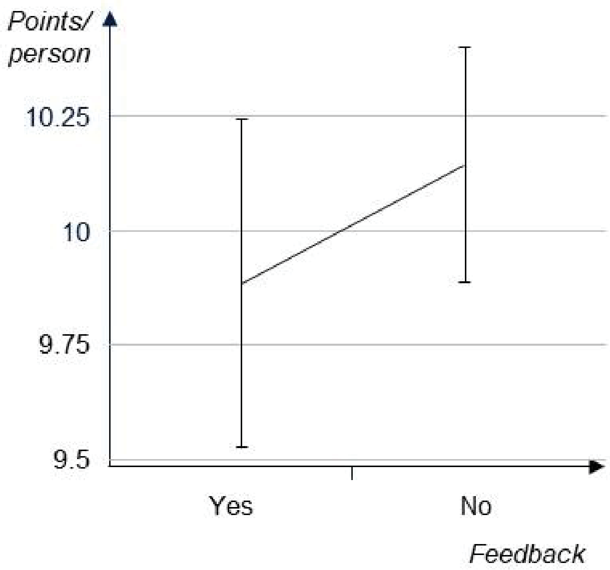


Figure 31 – Impact of direct testing and feedback on results on final test; error bars represent 95% confidence intervals; adapted from Jamovi, 2022

The statistical data of Table 27, calculated with a one-way ANOVA, using points/person at final test as dependent variable and direct testing and feedback or no testing and feedback as independent variables do not show any significant main effect with $p = 0.559$. Again, based on the data gathered, it appears that the hypothesised effect that testing and direct feedback leads to better learning than no testing and direct feedback does not exist.

Table 27 – One-way ANOVA with points/person at final test as dependent variable and feedback or no feedback as independent variables, adapted from Jamovi, 2022

ANOVA – Points/person

	Sum of squares	df	Mean square	F	*p*	η^2
Feedback	1.21	1	1.21	0.344	0.559	0.004

4.3.2.4 Learner's self-confidence (H4)

The calculations of the data in a one-way ANOVA using confidence after last learning session as the dependent variable and RI as the independent variable, did not reveal any significant main effect, with $p = 0.230$ (see Table 28).

Table 28 – One-way ANOVA with confidence after last learning session test as dependent variable and RI as independent variable, adapted from Jamovi, 2022

ANOVA – Confidence

	Sum of squares	df	Mean square	F	*p*	η^2
	979	2	488	1.48	0.230	0.020

This is also confirmed by the descriptive plots in Figure 32: no significant differences can be seen in confidence between the different groups. However, it could be argued that a trend towards spaced learning groups and especially those with a RI of two weeks being more confident towards a future test might occur.

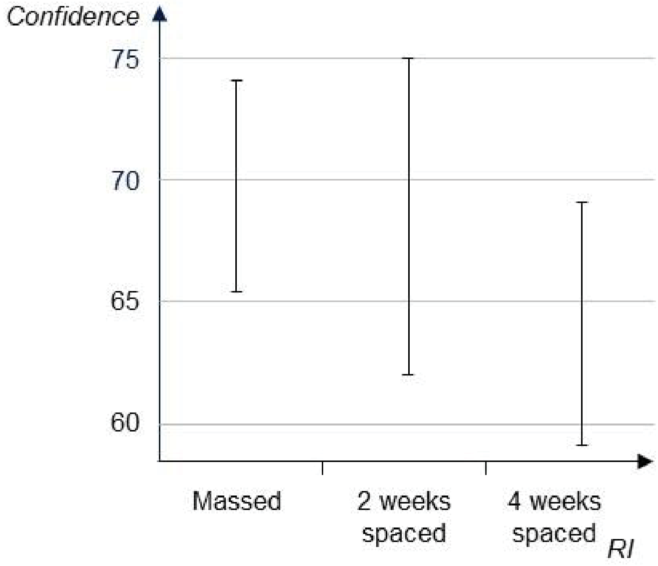


Figure 32 – Participant's reported confidence after final learning session; error bars represent 95% confidence intervals; adapted from Jamovi, 2022

4.3.2.5 Learner's self-perception (H5)

The calculations of the data in a two-way ANOVA using the delta of perceived vs. real outcomes at test as the dependent variable and instructional method as well as RI as independent variables (see Table 29), did not reveal any significant main effect, with *p(instructional method)* = 0.992 and *p(RI)* = 0.068, both of which are greater than 0.05. However, a statistically significant interaction effect between instructional method and RI was detected, with $p = 0.050$. Yet, since this significance is right on the borderline of significance, conclusions about possible interpretations should be drawn with caution.

Table 29 – Two-way ANOVA with the delta of perceived vs. real outcomes at test as dependent variable and instructional method as well as RI as independent variables, adapted from Jamovi, 2022

ANOVA – Delta perceived vs. real learning success

	Sum of squares	df	Mean square	F	p	η^2
Instructional method	3.62e-4	1	3.62e-4	1.13e-4	0.992	0.000
RI	10.8	1	10.75	3.36	0.068	0.015
Instructional method * RI	12.4	1	12.40	3.88	0.050	0.017

A Tukey-corrected post hoc test (Table 30) could not reveal any significant comparison with each p_{tukey} greater than 0.05, and therefore, no conclusion should be drawn from this interaction.

*Table 30 – Tukey-corrected post hoc test comparisons instructional method * RI, adapted from Jamovi, 2022*

Post hoc comparisons – Instructional method * RI

Comparison									
	RI		Instructional method	RI	Mean difference	SE	df	t	p_{tukey}
Massed	2 w	vs.	Massed	4 w	1.0679	0.478	215	2.235	0.117
Massed	2 w	vs.	Spaced	2 w	0.5560	0.380	215	1.465	0.461
Massed	2 w	vs.	Spaced	4 w	0.5179	0.414	215	1.251	0.595
Massed	4 w	vs.	Spaced	2 w	-0.5119	0.380	215	-1.349	0.533
Massed	4 w	vs.	Spaced	4 w	-0.5500	0.414	215	-1.329	0.546
Spaced	2 w	vs.	Spaced	4 w	-0.0381	0.295	215	-0.129	0.999

Note: Comparisons are based on estimated marginal means

This is also confirmed by the descriptive plots in Figure 33. No significant differences between the groups can be seen in how learners perceived they performed and how they actually performed at their tests.

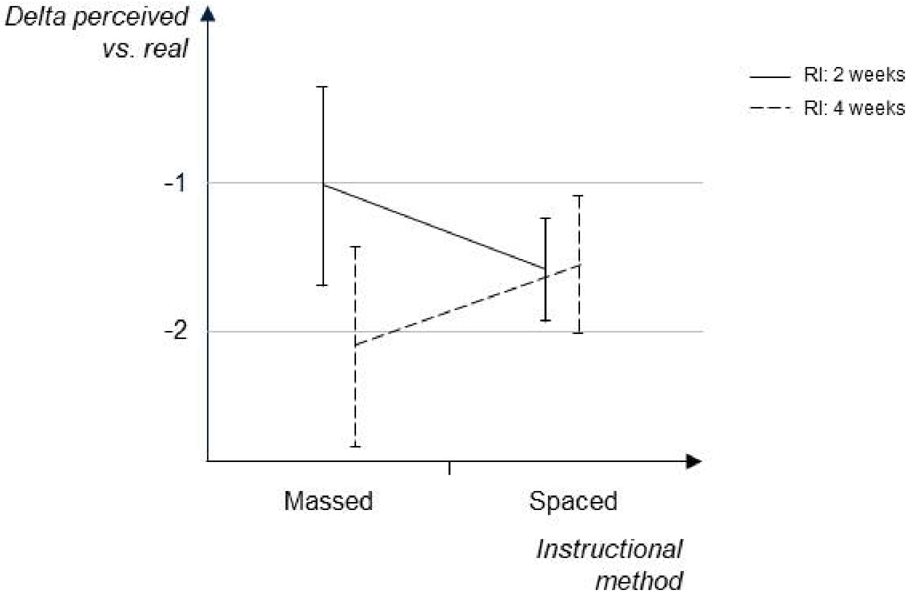


Figure 33 – Delta of perceived and actual learning success of participants; error bars represent 95% confidence intervals; adapted from Jamovi, 2022

As done in experiment 1, a repeated measures ANOVA was conducted (see Table 31).

Table 31 – Repeated measures ANOVA on delta of perceived and actual learning success of participants, adapted from Jamovi, 2022

Repeated measures ANOVA learner's self-perception (within subjects effects)

	Sum of squares	df	Mean square	F	*p*
Self-perception	195.69	1	195.69	122.42	< 0.001
Self-perception * Instructional method	1.81e-4	1	1.81e-4	1.13e-4	0.992
Self-perception * RI	5.38	1	5.38	3.36	0.068
Self-perception * Instructional method * RI	6.20	1	6.20	3.88	0.050

Note: Type 3 sums of squares

The repeated measures ANOVA revealed a significant main effect ($p < 0.001$) of the delta of perceived and actual learning success of participants, confirming the descriptive plots in above-shown Figure 33, namely that the delta of learners perceived, and actual learning success is significantly different from zero, which means that learners underestimate their actual learning success on average. Furthermore, a statistically significant interaction effect of self-perception, instructional method, and RI was revealed with $p = 0.050$, delineating the same picture as the previously calculated two-way ANOVA (with delta as dependent variable). Further analysis is shown in Appendix K.

4.3.2.6 Learner's satisfaction (H6)

Figure 34 represents the cumulated satisfaction ratings participants gave after each learning session they attended. The data in the figure shows that participants in the spaced learning groups attending three learning sessions recorded highest satisfaction levels and participants in the massed learning group recorded least lowest satisfaction scores.

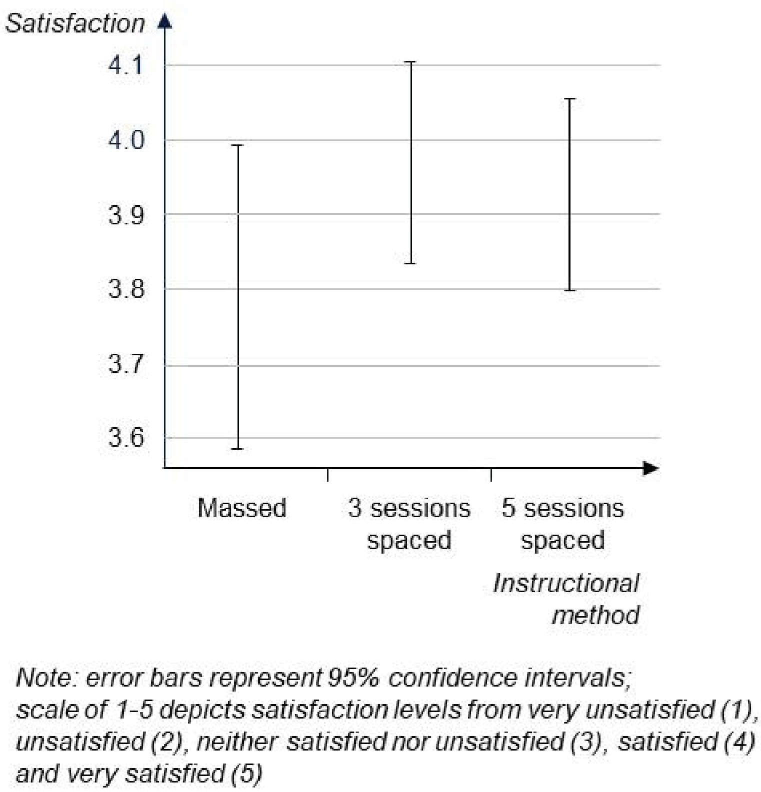


Figure 34 – Cumulated satisfaction ratings after each learning session; adapted from Jamovi, 2022

The data in Table 32 however, does not confirm a significant main effect, with $p = 0.345$. The data was analysed with a one-way ANOVA, using satisfaction as dependent variable and instructional method as the independent variable.

Table 32 – One-way ANOVA with satisfaction as dependent variable and instructional method as independent variable, adapted from Jamovi, 2022

ANOVA – Satisfaction

	Sum of squares	df	Mean square	F	*p*	η^2
Instructional method	2.38	2	1.19	1.07	0.345	0.004

Two further influences on the participant's satisfaction levels were assessed: The interactivity of the course and whether a learning session was self-paced or guided. In that sense, two one-way ANOVAs were undertaken, one with interactivity as the independent variable and the other with guidance as the independent variable. Since all learning interventions which were interactive in nature were also the guided ones, all calculations listed here for interactivity apply equally to guidance. For further analysis, refer to Appendix M. Both investigations did not reveal any statistical relevant results, as shown by the example of interactivity with $p = 0.687$ (Table 33).

Table 33 – One-way ANOVA with satisfaction as dependent variable and interactivity as independent variable, adapted from Jamovi, 2022

ANOVA – Satisfaction

	Sum of squares	df	Mean square	F	*p*	η^2
Interactivity	0.181	1	0.181	0.162	0.687	0.000

Even when conducting a two-way ANOVA, using satisfaction as the dependent variable and instructional method and interactivity as independent variables, no significant interaction effect of instructional method and interactivity was observed, with *p(instructional method)* = 0.394, *p(interactivity)* = 0.332 and *p(instructional method*interactivity)* = 0.365 (Table 34).

Table 34 – Two-way ANOVA with satisfaction as dependent variable and instructional method and interactivity as independent variable, adapted from Jamovi, 2022

ANOVA – Satisfaction

	Sum of squares	df	Mean square	F	*p*	η^2
Instructional method	2.08	2	1.04	0.933	0.394	0.003
Interactivity	1.10	1	1.10	0.982	0.322	0.002
Instructional method * Interactivity	2.25	2	1.13	1.010	0.365	0.003

Yet, since it can be seen from Figure 34 above that the mean values of the massed group are lower than those of the spaced group, it could be the case that there are effects that are not visible in the small group size available and could potentially be if the groups would have had more participants.

4.3.2.7 Learner's preference (H7)

Before evaluating this hypothesis, it should be noted that when adding up the percentages, the result is not 100 percent because there were participants who either did not give an answer or did not stay until the end of the experiment. This is the case for all the analyses below.

When asked about whether participants prefer a spaced learning approach ("more shorter sessions") over a massed one ("one longer session"), both the massed and spacing groups believed a spaced learning approach would be more effective in terms of long-term knowledge retention (on average 4.19 vs. 2.86 and a delta 0.85-1.90 points, each on a 5-scale) compared to a massed learning approach. Further, when asked about their desire for future work-based learning setups, participants clearly voiced preference for more shorter learning sessions (= spaced), with 84 percent of participants asking for it as compared to only 45 percent using this as current practice. This is mirrored by a low desire for one longer learning session (= massed). Here, 11 percent voiced preference for it, compared to 40 percent who currently only have these type of learning setups. The descriptive data is shown in Figure 35.

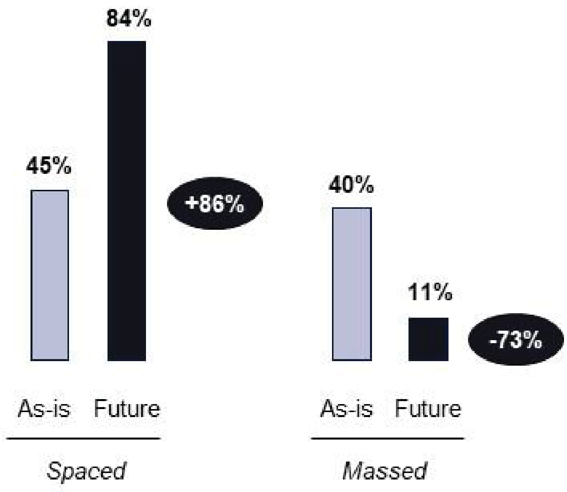


Figure 35 – Participants' preference of spaced learning versus massed learning, author's own compilation

When asked about whether participants would prefer interactive or non-interactive learning sessions in future, a clear desire for interactive sessions was stated, with 70 percent of participants asking for interactive ones. This is an increase of 82 percent compared to 38 percent of participants receiving mainly interactive learning interventions currently. Only 5 percent asked for non-interactive learning sessions in their future learning, which is a drop of 76 percent compared to 20 percent of participants receiving mainly non-interactive learning interventions currently. The descriptive data is shown in Figure 36.

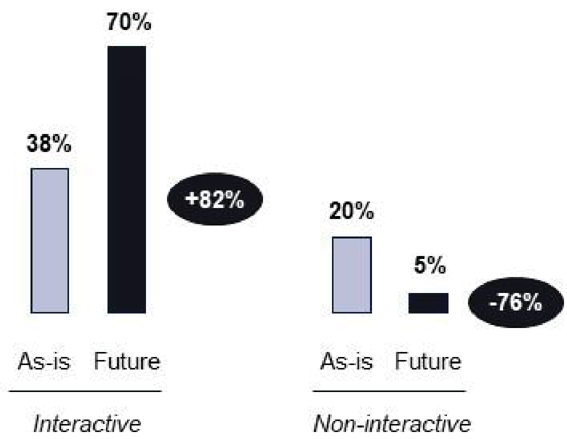


Figure 36 – Participants' preference of interactive versus non-interactive learning formats, author's own compilation

When asked about their preference of guided learning sessions versus self-paced learning sessions, participants recorded that current learning setups encourage self-paced learning (52 percent) over guided ones (24 percent). However, when asked on how they envision future learning setups to look, they preferred guided learning sessions over self-paced ones (preference for self-paced learning interventions dropped by 14 percent to 45 percent and preference for guided learning interventions increased by 54 percent to 36 percent). The descriptive data is shown in Figure 37.

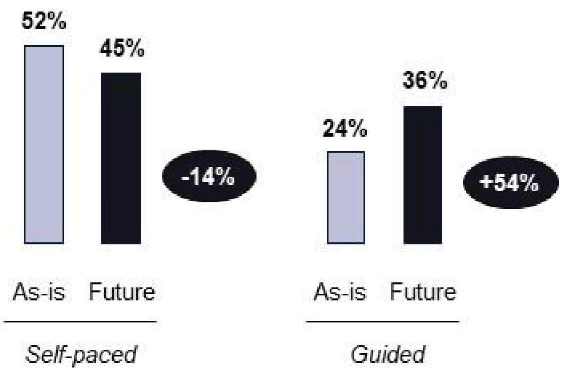


Figure 37 – Participants' preference of self-paced versus guided learning formats, author's own compilation

4.3.3 Discussion

The purpose of this experiment was to answer the leading research question whether the instructional method of spaced learning causes better learning in a work-based e-learning environment. Results indicate that for the case of work-based relevant procedural knowledge it is not associated with better learning compared to massed learning in an e-learning environment. Therefore, it appears that spaced learning is not applicable to procedural knowledge which can be easily implemented in daily life but rather for factual knowledge as shown in the previous experiment. As a result, the overarching hypothesis H1 and hypothesis H2a cannot be approved by this study.

Furthermore, within this experiment, no statistically significant evidence has been found to claim that testing and immediate feedback as well as a higher frequency of learning sessions leads to even better knowledge retention when teaching procedural knowledge as found in previous experiments teaching mainly factual knowledge (e.g., Cogliano et al., 2019; Dunlosky et al., 2013; Roediger & Karpicke, 2006a/b). Rather, other influences than testing and direct feedback as well as the frequency of learning sessions should be considered when dealing with procedural knowledge. From this investigation, both hypothesis H2b and hypothesis H3 are not supported.

Furthermore, this experiment yielded different results in terms of learners' self-perception on their learning success than previous experiments did (e.g., Bjork, 1999; Dunlosky et al., 2013; Morehead et al., 2015): no statistically significant differences have been found regarding how learners perceived their test performance and how they actually performed at test. Descriptively, and confirmed by a repeated measures ANOVA, all groups underestimated their learning. Thereby, for a RI of two weeks, massed learners showed a lower deviation to actual learning performance than spacers did. However, for the RI of four weeks the picture changed and the JOL of massed learners was more underconfident than those of the spaced learners. All in all, hypothesis H5 is not supported by this investigation and further research is needed on what drives learner's own judgement of learning, especially differentiating between different instructional methods and RIs.

Within this experiment, no support was found for learners in spaced learning interventions to feel less confident towards a future knowledge test than massed learners as shown by previous research (e.g., Simon & Bjork, 2001; Vlach et al., 2019; Zechmeister & Shaughnessy, 1980). From the given data, no statistical significance could be derived which underlines this finding and could support hypothesis H4. Another interesting finding was that participants in spaced learning groups attending three learning sessions were the most satisfied ones. Participants in the massed learning groups where least satisfied, contrary to postulations by Mayer and Moreno (1998, 2002, 2003). Yet, none of the statistical analyses revealed any significant main or interaction effect supporting hypotheses H6a, H6b or H6c, thus it is also not supported by the investigation. However, since for both analyses (H4 and H6) descriptive trends could be suspected, even if these trends are directed in the opposite direction to the original hypothesis, one could argue that the sample size of the given experiment was too small for potential opposite effects to become apparent. Therefore, further investigations independent from both hypotheses to explore these effects are needed.

Yet, it can be summarised that participants of the given experiment generally prefer spaced, interactive as well as guided learning interventions in their work life, which is in line with hypothesis H7. Even though preferences for self-paced learning formats were higher than for guided sessions, participants asked for less self-pacing and more guidance in their preferred future learning. Although the instructional method of spaced learning and related phenomena such as testing, feedback, and associated metacognitive effects, appear not to impact long-term knowledge retention when teaching procedural knowledge, it is yet perceived by learners to be the more efficient and effective instructional method, when compared to the instructional method of massed learning. Furthermore, learners interested in the subject matter would well accept interactive and guided forms of learning, especially within work-based learning offerings. Thus, this should be fully considered in the future design of work-based learning interventions.

In sum, the given experiment concludes the following:

- Hypotheses H1, H2, H3, H4, H5 and H6 are not supported by the investigation
- Hypothesis H7 is supported by the investigation

4.4 Comparison of effects between factual/conceptual and procedural knowledge

To find evidence for hypothesis H8, two steps were taken: First, a three-way ANOVA was performed between the experimental groups of both experiments 1 and 2. Relative points per person on the final test (0-100 percent) was chosen as the dependent variable, and the instructional method (massed or spaced), RI (two or four weeks), and the type of knowledge (factual/conceptual or process knowledge) were chosen as independent variables. As shown in Table 35, the analysis of the three-way ANOVA revealed three (i.e., all) statistically significant main effects between the two groups, namely for instructional method with $p < 0.001$ and $\eta^2 = 0.040$, for RI with $p < 0.001$ and $\eta^2 = 0.028$ and for the type of knowledge taught with $p < 0.001$ and $\eta^2 = 0.170$. Whereby the type of knowledge taught had the strongest impact on relative points. Furthermore, four (i.e., all) statistically significant interaction effects were discovered: first, between instructional method and RI with $p = 0.020$ and $\eta^2 = 0.008$, second, between instructional method and the type of knowledge taught with $p < 0.001$ and $\eta^2 = 0.025$, third, between RI and the type of knowledge taught with $p < 0\ .001$ and $\eta^2 = 0.019$ and fourth, between instructional method, RI and the type of knowledge taught with $p = 0.007$ and $\eta^2 = 0.011$ (also see Table 35).

Table 35 – Three-way ANOVA with relative points/person as dependent variable and instructional method, RI, and type of knowledge as independent variables, adapted from Jamovi, 2022

ANOVA – Relative points

	Sum of squares	df	Mean square	F	*p*	η^2
Instructional method	1.000	1	0.9996	26.09	< 0.001	0.040
RI	0.693	1	0.6932	18.09	< 0.001	0.028
Knowledge type taught	4.207	1	4.2074	109.80	< 0.001	0.170
Instructional method * RI	0.210	1	0.2102	5.49	0.020	0.008
Instructional method * Knowledge type taught	0.628	1	0.6282	16.39	< 0.001	0.025
RI * Knowledge type taught	0.481	1	0.4808	12.55	< 0.001	0.019
Instructional method * RI * Knowledge type taught	0.279	1	0.2788	7.28	0.007	0.011

The below Figure 38 illustrates this data further: participants in experiment 2 who were taught procedural knowledge scored better at final test than their counterparts in experiment 1 did, who were taught factual and conceptual knowledge. Furthermore, within both experimental groups, differences are obvious with regards to the relative points per person at final test depending on which instructional method they followed. However, it is also evident that the spaced learning effect becomes especially apparent for experimental group 1, who was taught factual and conceptual knowledge and is more pronounced with longer RIs.

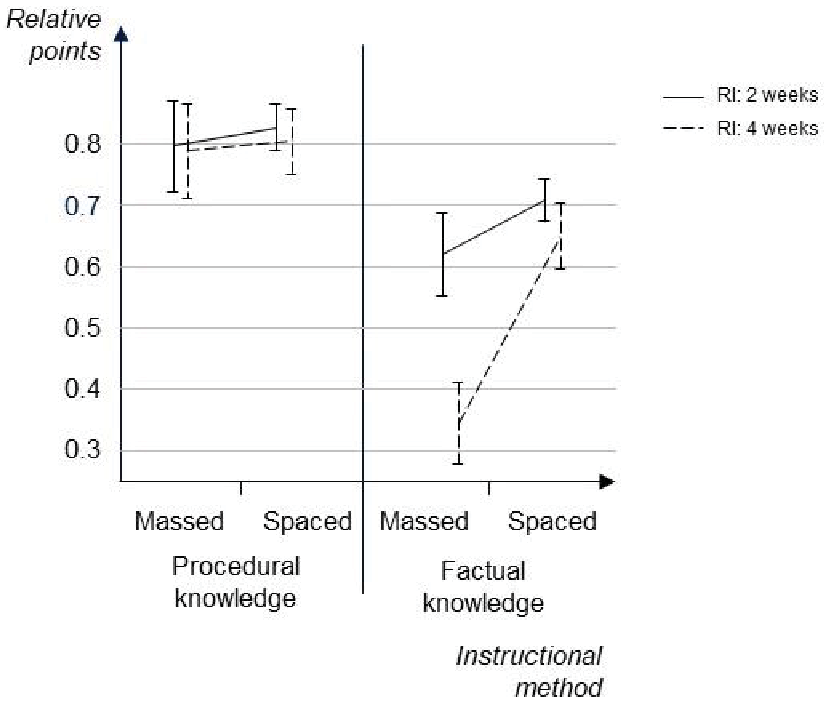


Figure 38 – Average performance of participants at final test for experiment 1 on factual knowledge (right-hand side) and for experiment 2 on procedural knowledge (left-hand side); error bars represent 95% confidence intervals; adapted from Jamovi, 2022

Furthermore, a Tukey-corrected post hoc test (Appendix N) confirmed the statistical significance of the expression of the spaced learning effect when comparing test results of learners who have been mainly learning factual and conceptual knowledge and those who have been mainly learning procedural knowledge. For the massed learning group with a RI of two weeks, being taught procedural knowledge three statistically significant differences have been revealed: first, against the massed learning group with a RI of two weeks, being taught factual knowledge with $p_{tukey} = 0.011$. Second, against the massed learning group with a RI of four weeks, being taught factual knowledge with $p_{tukey} < 0.001$. And third, against the spaced learning groups with a RI of four weeks, being taught factual knowledge with $p_{tukey} = 0.031$.

For the massed learning group with a RI of two weeks, being taught factual knowledge four statistically significant differences have been uncovered: First, versus the massed learning group with a RI of four weeks, being taught procedural knowledge with $p_{tukey} = 0.025$. Second, versus the massed learning group with a RI of four weeks, being taught factual knowledge with

$p_{tukey} < 0.001$. Third, versus the spaced learning groups with a RI of two weeks, being taught procedural knowledge with $p_{tukey} < 0.001$. And fourth, versus the spaced learning groups with a RI of four weeks, being taught procedural knowledge with $p_{tukey} < 0.001$.

For the massed learning group with a RI of four weeks, being taught procedural knowledge one statistically significant difference has been revealed, namely, versus the massed learning group with a RI of four weeks, being taught factual knowledge with $p_{tukey} < 0.001$.

For the massed learning group with a RI of four weeks, being taught factual knowledge four statistically significant differences have been found: First, against the spaced learning groups with a RI of two weeks, being taught procedural knowledge with $p_{tukey} < 0.001$. Second, against the spaced learning groups with a RI of two weeks, being taught factual knowledge with $p_{tukey} < 0.001$. Third, versus the spaced learning groups with a RI of 4 weeks, being taught procedural knowledge with $p_{tukey} < 0.001$. And fourth, versus the spaced learning groups with a RI of four weeks, being taught factual knowledge with $p_{tukey} < 0.001$.

Furthermore, for the spaced learning groups with a RI of two weeks, being taught procedural knowledge two statistically significant differences have been discovered: First, against the spaced learning groups with a RI of two weeks, being factual knowledge with $p_{tukey} < 0.001$ and second, against the spaced learning groups with a RI of four weeks, being taught factual knowledge with $p_{tukey} < 0.001$.

For the spaced learning groups with a RI of two weeks, being taught factual knowledge one statistically significant difference has been revealed versus the spaced learning groups with a RI of four weeks, being taught procedural knowledge with $p_{tukey} < 0.045$. Finally, for the spaced learning groups with a RI of four weeks, being taught procedural knowledge a statistically significant difference has been revealed against the spaced learning groups with a RI of four weeks, being taught factual knowledge with $p_{tukey} = 0.001$.

The results from this quantitative analysis confirm the first part of hypothesis H8 as outlined in chapter 3.2, arguing that the effects of spaced learning are especially applicable in work-based learning interventions for factual and conceptual rather for those for procedural knowledge.

As a second step, a metacognitive survey was set-up in Typeform (Appendix N) and sent to 88 participants who took part in both experiment 1 and experiment 2. Of these, 52 participants completed the survey. The aim of the survey was to gather information such as if and how differences between both learning topics were perceived, participants' thoughts on influences on learning outcomes, and participants' judgement of learning related to the two topics. The survey was based on the evaluation of the two metacognitive surveys after the final knowledge test of each learning experiments, where a thematic coding analysis on what particular elements the learners like the most revealed several differences between the two experiments, mainly on the topic of applicability (18 percent for "time management" vs. 0 percent for "platform business models"; see Appendix O for details). Results drawn from the survey are to underline the quantitative analysis of the statistical comparison of both experimental groups.

Through the metacognitive survey, the following tendencies emerged:

1. 69 percent of all participants found the training on "time management" more interesting in terms of content than the training on "platform business models"; likewise, 69 percent of participants stated that they knew more about the topic after the training on "time management" than after the training on "platform business models".
2. 83 percent of all participants who found the training on "time management" more interesting did so because of the direct applicability of what was learnt; the participants who found the training on "platform business models" better did so primarily because of the interesting content (69 percent).

3. 75 percent of the respondents found the training structure of the training on “time management” more effective than the training structure of the training on “platform business models”, although the structures of both trainings were identical (with slightly different divisions of the participants into the experimental groups).
4. 83 percent of respondents felt that the content of the training was more relevant to their perceived learning success than a more effective structure.
5. Personal interest and applicability (78 percent each, multiple answers possible) contributed most to the metacognitively assumed increase in knowledge; this was followed by the presentation of the media with 54 percent; it is interesting to note that participants who found the training on ‘time management’ more interesting in terms of content emphasised applicability with 86 percent, and among the participants who found the training on “platform business models” more interesting, the presentation of materials (63 percent) was almost as high as the interest on content (75 percent).
6. 90 percent of all participants found the training on “time management” more relevant for their everyday work than the training on “platform business models” (rating ‘agree’ or ‘fully agree’); 73 percent also stated that they can better remember the contents of the training on “time management”; 67 percent stated that they can “generally better remember topics that are more relevant for everyday life”; 77 percent of all participants stated that several repetition units help especially strongly with fact-based knowledge (only 48 percent stated this for the training on “platform business models”).

More details regarding the survey results can be found in Appendix N.

In summary, the results of this survey confirm that participants believe they learn easier and better when procedural knowledge is taught due to its applicability to their daily life. Applicability of the learnt by daily usage could be equated with constantly retrieving the learnt information from memory and hence, memory traces are strengthened and learning and knowledge retention increases (just like as it does with the testing effect). In addition, learners attribute their perceived learning success mainly to the content and applicability of what is learnt rather than the specific setup of their learning interventions and number of repetitions, i.e., spaced learning. These results are in line with what was postulated for hypothesis H8 in chapter 3.2, namely, that learners will judge their mastery in procedural knowledge higher due to easier encoding and retrieval fluency resulting from practice (Anderson, Krathwohl et al., 2001; Finn & Tauber, 2015; Moore & Healy, 2008). This is nicely described in one particular quote of a participant of the survey: "*In the case of [the training on] "platform business models", the repetition units were mainly relevant because the topic had less to do with one's own everyday life and one therefore internalised it less quickly*."

Taking all of the quantitative and qualitative analysis together, hypothesis H8 is thus confirmed by this investigation.

5. Summary of research

The final chapter of this research firstly builds upon the key findings drawn from the two experiments conducted for this research, summarises those, and *discusses in general the overarching research question as well as all hypotheses set* (chapter 5.1). Following is an elaboration on the *limitations of the two experiments* conducted to answer the research question (chapter 5.2). Finally, recommendations are made on first, *how to better design work-based e-learning interventions* in the future (chapter 5.3) and second, on where *future research* can continue (chapter 5.4).

5.1 General discussion of key findings

Although the spaced learning effect has created a large body of literature, prior research failed to examine spacing effects in more life-relevant learning environments such as work-based learning. In this research, to answer the question on positive learning effects in work-based e-learning environments due to spaced learning, two field experiments were conducted, both teaching work-based relevant contents aiming at closing this research gap, with the first experiment focusing on factual and conceptual knowledge (topic “platform business models”) and the second one focusing on procedural knowledge (topic “time management”).

When participants learnt factual and conceptual knowledge, the spaced learning effect became evident, especially for longer inter-session and retention intervals (herein seven days and four weeks, compared to a shorter interval of three days and two weeks). Altogether, participants recalled and retrieved more after the spaced learning intervention than participants in the massed learning intervention did (resembling an effect size of $\eta^2 = 0.119$). A different pattern became apparent for participants who have learnt procedural knowledge since the spaced learning effect did not become evident there.

Both results were independent of other factors previously investigated and said to enlarge the spaced learning effect, such as direct testing and feedback (e.g., Cogliano et al., 2019; Dunlosky et al., 2013; Morehead et al., 2015; Roediger & Karpicke, 2006a/b), as well as the number of learning sessions (e.g., Cepeda et al., 2006). The first experiment on work-based relevant factual knowledge revealed some descriptive trends towards spacing frequency, testing and immediate feedback, all of which positively impacts learning success, which clearly warrants further research. In contrast, the second experiment on work-based relevant procedural knowledge did not reveal any obvious systematic evidence that either spacing frequency or testing or immediate feedback impacts learning success at all.

Both experimental investigations failed to confirm previous findings on learner's confidence, self-perception, and satisfaction (e.g. Bjork, 1999; Dunlosky et al., 2013; Morehead et al., 2015; Simon & Bjork, 2001; Vlach et al., 2019; Zechmeister & Shaughnessy, 1980). Clear statements can neither be made that the learners of the massed condition always overestimated their own learning (both towards a future test and right after a distributed test has taken place) nor were they always more satisfied than the participants of the spaced condition. Rather effects directed in the opposite direction of the original hypotheses were detected and thus, other influences impacting learner's confidence, self-perception and satisfaction and the role different RIs play must be acknowledged, examined, and investigated in the future.

What can be noted however from both experiments is that regardless of which knowledge type was taught (factual and conceptual or procedural knowledge), participants universally agreed on the same learning preferences, asking for spaced, interactive, and guided work-based learning interventions (Table 36).

Table 36 – Changes in learner's preferences on learning design for experiment 1 and experiment 2; author's own compilation

	Preferred changes from current to future state	
	Experiment 1 (factual knowledge)	Experiment 2 (procedural knowledge)
Spaced	+72%	+86%
Massed	-65%	-73%
Interactive	+72%	+82%
Non-interactive	-80%	-76%
Self-paced	-12%	-14%
Guided	+16%	+54%

When answering the leading research question, if the instructional method of spaced learning causes better learning than massed learning in an e-learning environment in which work-related complex factual and procedural knowledge is taught, then the question must be answered in dependence of the knowledge type taught and until when knowledge needs to be retained. Spaced learning does not always lead to better learning in complex knowledge work-based e-learning environments, however, a more differentiated view is needed.

Whilst the spaced learning effect could be demonstrated clearly for work-based relevant factual and conceptual knowledge and long RIs, but not for work-based relevant procedural knowledge, a comparison of the spaced learning effects within the two experiments clearly showed that the spaced learning effect especially causes better learning for factual-based knowledge topics, which are unrelated to daily usage and should be retained for longer periods. When comparing the results of the overarching hypothesis H1 in both experiments, an effect size of $\eta^2 = 0.170$ was revealed for the impact the type of knowledge being taught had on the spaced learning effect. Hence, the spaced learning effect can be observed especially in the context of factual and conceptual knowledge. This quantitative finding was further supported by a metacognitive survey among all participants who took part in both experiments.

Given the practical importance of well-designed work-based learning interventions resulting in long-term knowledge retention, research is still widely lacking on how the spaced learning effect can be applied to work-based learning interventions. This research provides strong empirical evidence which should be investigated further, especially with regards to differentiating between types of knowledge being taught and for long-term desired learning success and retention.

A complete overview of all experimental results of the research at hand can be found in Table 37.

Table 37 – Overview of all experimental results of the research at hand; author's own compilation

Hypothesis	Results of experiment 1: factual knowledge	Results of experiment 2: procedural knowledge
H1 – Massed learning versus spaced learning	▪ Supported by the investigation	▪ Not supported by the investigation
H2a/b – Spacing frequency	▪ 2a: Supported by the investigation ▪ 2b: Lacking strong evidence within the data but seem plausible on a descriptive level ▪ Further research needed	▪ 2a: Not supported by the investigation ▪ 2b: Not supported by the investigation
H3 – Testing and immediate feedback	▪ Lacking strong evidence within the data but seem plausible on a descriptive level ▪ Further research needed	▪ Not supported by the investigation
H4 – Learner's self-confidence	▪ Not supported by the investigation ▪ Further research needed	▪ Not supported by the investigation
H5 – Learner's self-perception	▪ Not supported by the investigation ▪ Further research needed	▪ Not supported by the investigation
H6 – Learner's satisfaction	▪ Not supported by the investigation ▪ Further research needed	▪ Not supported by the investigation
H7 – Learner's preference	▪ Supported by the investigation	▪ Supported by the investigation
H8 – Knowledge categories	▪ Supported by the investigation	

5.2 Limitations of the research at hand

While this research successfully showed that spaced learning causes better learning in e-learning environments teaching complex factual and conceptual knowledge than common massed learning, there were some limitations to the internal and external validity of the research at hand.

First, as the literature review pointed out, there have been no real-life spaced learning (field) experiments in the knowledge areas of factual knowledge, conceptual knowledge, or procedural knowledge. Therefore, there was no valid prior literature on which to base this specific experimental design. Although one of the leading researchers in spaced learning research was consulted on the topic of experimental design (Appendix P) and found it reasonable, it was only possible in the design of the experiments to refer to literature that had 'simpler' sequences.

Second, since both experiments had self-selection as the primary selection method of study participants, there is an underlying sampling bias in both experiments. Thus, the results of both experiments cannot be fully generalised, but only refer to the population of learners who are interested in a specific teaching topic. Furthermore, the number of participants who left the experiment prior to finishing the knowledge test should be mentioned. This was especially relevant for groups with a RI of four weeks for the first experiment and due to the higher drop out of participants in these groups, the dependent variable of the knowledge tests could not be observed for all participants. It is assumed that this drop out could be due to diminished motivation of the respective participants or the self-assessment that they knew too little for a test. However, this cannot be fully confirmed with the present experiments. Therefore, it is recommended to direct further research on how to keep learners engaged over a longer period in a learning intervention.

Third, since both experiments were conducted as field experiments and not in the laboratory, there might have been unknown influences that could have been controlled in a laboratory environment. Furthermore, it was only possible to instruct the participants in advance not to take notes and transcripts and not to repeat the contents between the learning units and between the last learning unit and the knowledge test. However, it was not possible to control this within the setting of these experiments. It was also not possible to prevent a possible exchange of the study participants, who were in different learning groups, which could have resulted in a reduction of motivation to participate in the learning experiment.

Fourth, the required number of participants of 32 per group was not reached in all experimental groups and even in the groups in which 32 participants took part, the number of people was not exceeded. Since the statistical power of experiments depends on the effect size that one wants to detect with an experiment, one has a higher power for larger effects and vice versa. Accordingly, more test subjects are needed if smaller effects with a higher power shall be detected. Assuming that the effects of the hypotheses that were not confirmed by these experiments are smaller than expected, the power of both experiments was too low for the given sample size.

Finally, as this research followed the research data guideline of the Heinrich Heine University, Düsseldorf (Appendix I), the personal data on the participants' feelings and metacognition after the individual learning sessions and on the learning success of the study participants were anonymised according to the relevant research standards. This meant that it was not possible to assign the knowledge tests carried out to a specific person. However, this also meant that correlations between the participants' surveys after each learning unit and the results of the knowledge tests were not possible. Thus, each survey response and each test had to be considered individually with no option to assign these to specific individuals.

5.3 Practical recommendations for future work-based e-learnings

The empirical findings derived from the two experiments conducted for this research provide meaningful and beneficial insights for any designer and provider of work-based e-learning interventions as well as for managers and executives in charge seeking long-term knowledge retention of individuals taking part in the interventions.

First, the results of the empirical field experiments demonstrate that designers and providers of work-based e-learning interventions aiming at teaching factual as well as conceptual knowledge should make use of the instructional method of spaced learning.

Especially for those interventions in which long-term knowledge retention is key, longer ISIs as well as RIs should be considered compared to having a single e-learning session. Cepeda and team's (2008) analysis on the optimum gap for restudy could be used as a baseline for scheduling learning sessions and adapted where needed.

Second, the findings across both experiments show that irrespective of which knowledge type was taught, learners prefer spaced learning experiences (indicating a stronger motivation to learn). Thus, even though the spaced learning effect could not be proven for an e-learning intervention teaching procedural knowledge, designers and providers of work-based e-learning interventions should still consider the instructional method of spaced learning based on learner's preference. Also, learners unanimously preferred interactive and guided learning sessions over non-interactive, self-paced learning sessions. Therefore, it is recommended to design work-based e-learning interventions in a spaced, interactive, and guided form. This is all based on the premise that, under the given results, an actual increase in knowledge retention only occurs for factual and conceptual knowledge.

Third, the empirical findings at hand should encourage managers and executives in charge to rethink the way work-based learning is conducted, implemented, how they prioritise human resources and learning departments on their organisational agenda as well as how they invest their learning and development budgets. This research has clearly shown that approaches to learning and development relying on traditional instructional methods such as massed learning, non-interactive, self-paced learning formats are to be rejected. The following three reasons, derived from this research, shall be listed as justification:

- They do not lead to sustainable knowledge retention, especially for long lasting intervals and interventions teaching factual and conceptual knowledge.
- They are most likely to be incorrect investments of learning budgets.

- They are not the preferred choice of learners and hence might lead to job dissatisfaction, resulting in employees changing employers.

By getting learning right and supporting human resources as well as learning departments accordingly, organisations can sustain and enhance their competitive advantage by making sure the right people are kept within the organisation through investing in the right learning and skill-building measures.

5.4 Theoretical recommendations for future research

The present research is the first known empirical effort to investigate the effect of spaced learning on real-life work-based e-learning interventions and as such contributes to academic research in the field of learning in general and management learning in particular. It offers insights about how spaced learning can be applied to complex, non-laboratory, work-based e-learning interventions. Thereby, the research followed previously mentioned calls for applied real-life relevant research of the spaced learning effect (e.g., from Carpenter et al., 2012; Dempster, 1988; Kapler et al., 2015; Karpicke et al., 2016; Larsen, 2018; Mettler et al., 2016; Seabrook et al., 2005; Sobel et al., 2011).

The results from the first experiment of this research led to the conclusion that spaced learning interventions cause considerably better learning outcomes compared to massed learning interventions, in line with previous research (e.g., Carpenter et al., 2012; Cepeda et al., 2006; Delaney et al., 2010; Dempster, 1989; Gerbier & Toppino, 2015; Vlach et al., 2019), especially with increasing RI (which is of specific relevance for lifelong learning). However, the benefit of spaced learning interventions compared to massed learning interventions have not been validated by the second experiment of the research at hand. This lack of benefit is interesting as participants in the second experiment performed much better on average across groups in the final knowledge test than those of the first experiment, even though no spaced learning effect could be shown. One potential reasoning for this finding could be that the occurrence of the spaced learning effect does not solely depend on how learning interventions are designed technically, e.g., schedules, existence of feedback, number of sessions, but also on the knowledge type taught, i.e., either factual/conceptual or procedural. Since past learning research within the field of spaced learning has been examined in depth in the laboratory (Wiseheart et al., 2019), thereby mainly focusing on verbal or trivia factual learning (e.g., Carpenter et al., 2012; Kapler et al., 2015), only few studies are known to have assessed the effects on more complex higher-level skill learning such as mathematical and science concept learning as well as inductive category learning and making complex judgments (e.g., Kang & Pashler, 2012; Kornell & Bjork, 2008; Rohrer, 2009; Rohrer & Taylor, 2006, 2007; Vlach et al., 2008) or even the effects of perceptual and coordinated motor tasks (Baddeley & Longman, 1978; Carpenter et al., 2012; Cepeda et al., 2006; Dempster 1996). Yet, no applied study is known to have assessed the impacts, if any, spaced learning could have on acquiring procedural knowledge. Therefore, it is recommended, that future research carefully investigates if and how the effect appears with different knowledge categories (in this case: factual/conceptual and procedural) or occurs at all. If the effect is not widely demonstrable for procedural knowledge, this could be because procedural knowledge is learnt differently by the brain than factual or conceptual knowledge (Anderson, Krathwohl et al., 2001). But this, too, would have to be further investigated in future research. In case the outcomes concur with those detected in the present research and verify the findings, practical recommendations should then be extended on how to best design and run work-based e-learning interventions.

Since the present study did not find any evidence that previously identified factors such as testing and direct feedback, which are supposed to strengthen the spaced learning effect (e.g., Cogliano et al., 2019; Dunlosky et al., 2013; Morehead et al., 2015; Roediger & Karpicke, 2006a/b), actually enhance knowledge retention in real-life work-based learning interventions, it remains to be investigated why this is the case and whether these factors become obsolete with higher complexity of the learning material or if other ways of implementing these lead to an even stronger impact on knowledge retention. The lack of evidence for those factors in the present research appears quite intriguing. Obvious differences between the research at hand and previously conducted research are the complexity of learning materials (simple vocabulary or fact learning versus learning of concepts or new behaviours), the change of media between the independent learning sessions (videos, slides, live session(s)), the incorporation of more than two spaced learning sessions (in the research at hand three or even five repetition sessions) as well as the limited range of tested learning schedules within this experiment (ISI either four or seven days and RI either two weeks or four weeks). Even though these learning schedules have been based on Cepeda's and teams (2008) optimal gap on restudying, it could be that shorter gaps do not produce any benefits with increasing complexity of learning content and changes in learning media. Since the primary goal of any work-based learning intervention should be enabling learners to apply and transfer the knowledge learnt (Connors, 2021; Gallardo, 2021; Göldi, 2011; Ryo & Moon, 2019; Sala & Gobet, 2017), the above should further be investigated and thereby clearly differentiate between the correlations between knowledge types (e.g., factual/conceptual and procedural), knowledge complexity, learning media used, RIs as well as ISIs (shorter ones versus longer ones with relevance to real-life application).

Also, it is recommended to investigate drivers of learner's confidence, self-perception, and satisfaction further as, again, previous findings of mainly laboratory-based spaced learning research could not be confirmed in the field of work-based learning with the given research and at times even opposing findings were observed (e.g., Bjork, 1999; Dunlosky et al., 2013; Morehead et al., 2015; Simon & Bjork, 2001; Vlach et al., 2019; Zechmeister & Shaughnessy, 1980). This in turn is also of great interest, as the body of research in this field is relatively large and the findings do not seem like coincidences. A reasoning for the lack of evidence in the two experiments at hand could be that the single learning sessions were too long and as a result perceived as independent massed learning entities rather than repetitions of the same materials.

Arguments that could speak in favour to this are first, that learning sessions were on average quite long (shortest session 15 minutes and longest session 60 minutes) and second, the change in learning media used, which addressed different independent working-memory channels. As a result, one could assume that rather than forming one strong memory trace, several weaker memory traces were formed on the topics taught, which also led to learning success, yet did not confirm the metacognitive findings of previous research. It is thus recommended to investigate in more depth, how interactions of chosen instructional methods, learning formats and media, and especially different lengths of retention intervals impact learner's confidence, self-perception, and satisfaction. Further, since it has not been part of this research, it is recommended to investigate why learners in work-based learning environments are motivated to learn and how this impacts real learning success as well as learner's metacognition.

This research has nevertheless shown evidence that the spaced learning effect is applicable to work-based e-learning interventions, to the extent of factual and conceptual knowledge. Yet, to increase the validity of the findings of the two field experiments conducted, further field experiments should be considered. In case the resources and time of future research allow it, the two presented experiments could be replicated. Thereby, care should be taken to reach a much larger number of participants to be able to detect small effects. As a suggestion for a future experimental setup, the following is proposed:

- Two separate experimental investigations, one teaching factual/conceptual knowledge and one teaching procedural knowledge, with about the same level of difficulty.
- Within each of the experimental investigations, two separate ISI-RI conditions are differentiated: the first in accordance with the longer ISI-RI conditions of the experiments of the research at hand, i.e., an ISI of seven days and a RI of four weeks. The second should be even longer with an ISI of ten days and a RI of eight weeks.

- For each of the conditions, the following nine groups should be considered: a) a massed control learning group, b) a spaced learning group with three learning sessions, c) a spaced learning group with three learning sessions and changing learning media, d) a spaced learning group with three learning sessions and direct testing and feedback, e) a spaced learning group with three sessions, changing learning media, direct testing and feedback, f) a spaced learning group with five learning sessions, g) a spaced learning group with five learning sessions and changing media, h) a spaced learning group with five learning sessions, direct testing and feedback and also, i) a spaced learning group with five learning sessions, changing media, direct testing and feedback. Thus, resulting in 18 groups per experiment.
- Assuming an ideal effect size of at least $\eta^2 = 0.06$; an alpha level of 0.05 and a required power of 0.8 (see Cepeda et al., 2006), at least 18 participants per group would be required, however larger groups should be considered to reveal small effects.
- Once both experiments with each 18 groups have been conducted, conclusions shall be drawn per experiment, knowledge type taught, ISI-RI condition, number of learning sessions, usage of learning media as well as direct testing and feedback. Further, cross-experimental comparisons are recommended to build upon the findings of this research.

All in all, spaced learning, and all its moderating influences as well as retrieval practice are widely recognised and proven instructional methods to enhance long-term knowledge retention. However, the majority of findings and recommendations drawn for real-life application mostly stem from laboratory experiments that are very similar in structure and content. Not every finding on spaced learning and the distributed practice made in the laboratory is applicable to all learning fields, subjects and learning environments but rather different combinations of the techniques should be considered when it comes to real-life application. As a result, much broader research is needed to investigate how to best implement and take advantage of spaced learning in more applied settings such as work-based learning interventions.

References

Acedo, F. J., Barroso, C. & Galan, J. L. (2006). The resource-based theory: Dissemination and main trends. *Strategic Management Journal*, *27*(7), 621–636. https://doi.org/10.1002/smj.532

Adams, N. E. (2015). Bloom's taxonomy of cognitive learning objectives. *Journal of the Medical Library Association*, *103*(3), 152–153. https://doi.org/http://dx.doi.org/10.3163/1536-5050.103.3.010

Agarwal, P. K. (2019). Retrieval practice & bloom's taxonomy: Do students need fact knowledge before higher order learning? *Journal of Educational Psychology*, *111*(2), 189–209. https://doi.org/10.1037/edu0000282

Agrawal, S., De Smet, A., Lacroix, S. & Reich, A. (2020). To Emerge Stronger from the COVID-19 crisis, companies should start reskilling their workforces now. In McKinsey & Company, *McKinsey Insights Organization Practice and McKinsey Accelerate* (Issue May). McKinsey & Company, Global Editorial Services.

Alheit, P. & Dausien, B. (2018). Bildungsprozesse über die Lebensspanne und lebenslanges Lernen. In R. Tippelt & B. Schmidt-Hertha, *Handbuch Bildungsforschung* (4th ed., pp. 877–903). Springer VS, Wiesbaden.

Andersen, M. S. & Makransky, G. (2020). The validation and further development of a multidimensional cognitive load scale for virtual environments. *Journal of Computer Assisted Learning, 37,* 183–196. https://doi.org/10.1111/jcal.12478

Anderson, L. W., Krathwohl, D. R., Airasian, P. W., Cruikshank, K. A., Mayer, R. E., Pintrich, P. R., Raths, J. & Wittrock, M. C. (2001). *A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives*. Longman, New York.

Anderson, J. R. & Schunn, C. D. (2000). Implications of the ACT-R learning theory: No magic bullets. In R. Glaser (Ed.), *Advances in instructional psychology: Educational design and cognitive science* (vol. 5, pp. 1–33). Lawrence Erlbaum Associates Publishers, New Jersey.

Ariel, R. & Karpicke, J. D. (2018). Improving self-regulated learning with a retrieval practice intervention. *Journal of Experimental Psychology: Applied*, *24*(1), 43–56. https://doi.org/10.1037/xap0000133

Ayres, P. & Paas, F. (2012). Cognitive Load Theory: New Directions and Challenges. *Applied Cognitive Psychology, 26*(6), 827–832. https://doi.org/10.1002/ACP.2882

Baddeley, A. D. (1986). *Working memory*. New York: Oxford University Press.

Baddeley, A. D. (2000). The episodic buffer: a new component of working memory? *Trends in Cognitive Sciences, 4*(11), 417–423. https://doi.org/10.1016/s1364-6613(00)01538-2

Baddeley, A. D. & Longman, D. J. A. (1978). The influence of length and frequency of training session on the rate of learning to type. *Ergonomics*, *21*(8), 627–635. https://doi.org/10.1080/00140137808931764

Bahrick, H. P. (1979). Maintenance of Knowledge: Questions About Memory We Forgot to Ask. *Journal of Experimental Psychology: General*, *108*(3), 296–308. https://doi.org/10.1037/0096-3445.108.3.296

Bahrick, H. P., Bahrick, L. E., Bahrick, A. S. & Bahrick, P. E. (1993). Maintenance of Foreign Language Vocabulary and the Spacing Effect. *Psychological Science, 4*(5), 316–321. https://doi.org/10.1111/j.1467-9280.1993.tb00571.x

Bahrick, H. P. & Phelphs, E. (1987). Retention of Spanish vocabulary over 8 years. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *13*(2), 344–349. https://doi.org/10.1037/0278-7393.13.2.344

Balota, D. A., Duchek, J. M. & Logan, J. M. (2011). Is expanded retrieval practice a superior form of spaced retrieval? A critical review of the extant literature. In J. S. Nairne (Ed.), *The Foundations of Remembering: Essays in Honor Of Henry L. Roediger, III* (pp. 83–105). Psychology Press. https://doi.org/10.4324/9780203837672

Balota, D. A., Duchek, J. M., Sergent-Marshall, S. D. & Roediger, H. L. (2006). Does expanded retrieval produce benefits over equal-interval spacing? Explorations of spacing effects in healthy aging and early stage Alzheimer's disease. *Psychology and Aging*, *21*(1), 19–31. https://doi.org/10.1037/0882-7974.21.1.19

Balota, D. A., Yap, M. J., Cortese, M. J., Hutchison, K. A., Kessler, B., Loftis, B., Neely, J. H., Nelson, D. L., Simpson, G. B. & Treiman, R. (2007). The English Lexicon project. *Behavior Research Methods, 39*(3), 445–459. https://doi.org/10.3758/BF03193014

Beer, M., Finnström, M. & Schrader, D. (2016). *Why Leadership Training Fails—and What to Do About It*. Retrieved from https://hbr.org/2016/10/why-leadership-training-fails-and-what-to-do-about-it. Accessed: May 25, 2022.

Beck, K., Beedle, M., van Bennekum, A., Cockburn, A., Cunningham, W., Fowle, M., Grenning, J., Highsmith, J., Hunt, A., Jeffries, R., Kern, J., Marick, B., Martin, R. C., Mellor, S., Schwaber, K., Sutherland, J. & Thomas, D. (2001). *Manifesto for Agile software development*. Retrieved from http://agilemanifesto.org/. Accessed: February 2, 2022.

Becker, G. S. (1962). Investment in Human Capital: A Theoretical Analysis. *Journal of Political Economy, 70*(5), 9–49. http://dx.doi.org/10.1086/258724

Beier, M. E. (2021). Life-Span Learning and Development and Its Implications for Workplace Training. *Current Directions in Psychological Science, 31*(1), 56–61. https://doi.org/10.1177/09637214211003891

Benjamin, A. S. & Craik, F. I. (2001). Parallel effects of aging and time pressure on memory for source: evidence from the spacing effect. *Memory & cognition, 29*(5), 691–697. https://doi.org/10.3758/bf03200471

Benjamin, A. S. & Tullis, J. (2010). What makes distributed practice effective? *Cognitive Psychology*, *61*(3), 228–247. https://doi.org/10.1016/j.cogpsych.2010.05.004

Bersin, J. (2018). *New Research Shows "Heavy Learners" More Confident, Successful, and Happy at Work*. Retrieved from https://www.linkedin.com/pulse/want-happy-work-spend-time-learning-josh-bersin/. Accessed: May 30, 2022.

Billett, S. (2014). Conceptualising lifelong learning in contemporary times. In T. Halttunen, M. Koivisto & S. Billett (Eds.), *Promoting, Assessing, Recognizing and Certifying Lifelong Learning: International Perspectives and Practices International Perspectives and Practices* (vol. 20, pp. 19–35). Springer, Dordrecht. https://doi.org/10.1007/978-94-017-8694-2_2

Billett, S. (2018). Distinguishing lifelong learning from lifelong education. *Journal of Adult Learning, Knowledge and Innovation*, *2*(1), 1–7. https://doi.org/10.1556/2059.01.2017.3

Billing, F., De Smet, A., Reich, A. & Schaninger, B. (2021). Building workforce skills at scale to thrive during — and after — the COVID-19 crisis. In McKinsey & Company Global Publishing, *McKinsey & Company Organization Practice* (Issue April). McKinsey & Company Global Publishing.

Bird, S. (2010). Effects of distributed practice on the acquisition of second language English syntax. *Applied Psycholinguistics, 31*(4), 635–650. https://doi.org/10.1017/S0142716410000172

Bjork, R. A. (1979). Information-processing analysis of college teaching. *Educational Psychologist, 14*, 15–23. https://doi.org/10.1080/00461527909529203

Bjork, R. A. (1999). Assessing our own competence: Heuristics and illusions. In D. Gopher & A. Koriat (Eds.), *Attention and performance XVII: Cognitive regulation of performance: Interaction of theory and application* (pp. 435–459). The MIT Press, Cambridge Massachusetts.

Bjork, R. A. & Allen, T. W. (1970). The spacing effect: consolidation or differential encoding? *Journal of Verbal Learning and Verbal Behavior, 9*(5), 567–572. https://doi.org/10.1016/S0022-5371(70)80103-7

Bjork, R. A. & Bjork, E. L. (1992). A new theory of disuse and an old theory of stimulus fluctuation. In A. F. Healy, S. M. Kosslyn & R. M. Shiffrin (Eds.), *Essays in honor of William K. Estes*, *Vol. 1. From learning theory to connectionist theory; Vol. 2. From learning processes to cognitive processes* (pp. 35–67). Lawrence Erlbaum Associates Publishers, New Jersey.

Bjursell, C. (2020). The COVID-19 pandemic as disjuncture: Lifelong learning in a context of fear. *International Review of Education*, *66*(5–6), 673–689. https://doi.org/10.1007/s11159-020-09863-w

Black, P. & Wiliam, D. (1998). Assessment and Classroom Learning. *Assessment in Education: Principles, Policy & Practice, 5*(1), 7–74. https://doi.org/10.1080/0969595980050102

Black, P. & Wiliam, D. (2009). Developing the theory of formative assessment. Educational Assessment, *Evaluation and Accountability, 21*(1), 5–31. https://doi.org/10.1007/s11092-008-9068-5

Blanchard, K. (2013). Make your training stick. *People Dynamics, 13*(9), 13.

Blanden, J. & Machin, S. (2010). Education and Inequality. In P. Peterson, E. Baker & B. McGaw, *International Encyclopedia of Education* (3rd ed., pp. 272–281). Elsevier

Science. https://doi.org/10.1016/B978-0-08-044894-7.01218-5

Bloom, B. S. (Ed.), Engelhart, M. D., Furst, E. J., Hill, W. H. & Krathwohl, D. R. (1956). *The taxonomy of educational objectives: The classification of educational goals* (Handbook 1: Cognitive domain). David McKay Company, New York.

Blume, B. D., Ford, J. K., Baldwin, T. T. & Huang, J. L. (2010). Transfer of training: A meta-analytic review. *Journal of Management*, *36*(4), 1065–1105. https://doi.org/10.1177/0149206309352880

Bodem-Schrötgens, J., Reich, A., Schaninger, B. & Sharma, K. (2021). *Three keys to building a more skilled postpandemic workforce*. Retrieved from: https://www.mckinsey.com/business-functions/people-and-organizational-performance/our-insights/three-keys-to-building-a-more-skilled-postpandemic-workforce. Accessed: July 20, 2022.

Brandi, U. & Iannone, R. L. (2017). Learning strategies for competence development in enterprises. *Industrial and Commercial Training*, *49*(1), 1–5. https://doi.org/10.1108/ICT-08-2016-0052

Braun, K. & Rubin, D. C. (1998). The Spacing Effect Depends on an Encoding Deficit, Retrieval, and Time in Working Memory: Evidence from Once-presented Words. *Memory*, *6*(1), 37–65. https://doi.org/10.1080/741941599

Broadbent, J., Sharman, S., Panadero, E. & Fuller-Tyszkiewicz, M. (2021). How does selfregulated learning influence formative assessment and summative grade? Comparing online and blended learners. *Internet and Higher Education, 50*, 100805. https://doi.org/10.1016/j.iheduc.2021.100805

Brünken, R., Plass, J. L. & Leutner, D. (2003). Direct measurement of cognitive load in multimedia learning. *Educational Psychologist, 38*(1), 53–61.

https://doi.org/10.1207/S15326985EP3801_7

Bughin, J., Hazan, E., Lund, S., Dahlström, P., Wiesinger, A. & Subramaniam, A. (2018). Skill shift: Automation and the future of the workforce. In McKinsey & Company, *McKinsey Global Institute: Skill shift: Automation and the future of the workforce*. Discussion Paper (Issue May). McKinsey & Company, Global Editorial Services.

Butler, A. C., Godbole, N. & Marsh, E. J. (2013). Explanation feedback is better than correct answer feedback for promoting transfer of learning. *Journal of Educational Psychology*, *105*(2), 290–298. https://doi.org/10.1037/a0031026

Butler, A. C., Karpicke, J. D. & Roediger, H. L. III. (2007). The effect of type and timing of feedback on learning from multiple-choice tests. *Journal of Experimental Psychology: Applied, 13*(4), 273–281. https://doi.org/10.1037/1076-898X.13.4.273

Butler, A. C., Karpicke, J. D. & Roediger, H. L. III. (2008). Correcting a metacognitive error: Feedback increases retention of low-confidence correct responses. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*(4), 918–928. https://doi.org/10.1037/0278-7393.34.4.918

Carneiro, P., Dearden, L. & Vignoles, A. (2010). The economics of vocational education and training. In P. Peterson, E. Baker & B. McGaw, *International Encyclopedia of Education* (3rd ed., pp. 255–261). Elsevier Science. https://doi.org/10.1016/B978-0-08-044894-7.01737-1

Carpenter, S. K., Pashler, H. & Cepeda, N. (2009). Using Tests to Enhance 8th Grade Students' Retention of U.S. History Facts. *Applied Cognitive Psychology*, *23*(6), 760–771. https://doi.org/10.1002/acp.1507

Carpenter, S. K., Cepeda, N. J., Rohrer, D., Kang, S. H. K. & Pashler, H. (2012). Using Spacing to Enhance Diverse Forms of Learning: Review of Recent Research and

Implications for Instruction. *Educational Psychology Review*, *24*(3), 369–378. https://doi.org/10.1007/s10648-012-9205-z

Carpenter, S. K. & DeLosh, E. L. (2005). Application of the Testing and Spacing Effects to Name Learning. *Applied Cognitive Psychology*, *19*(5), 619–636. https://doi.org/10.1002/acp.1101

Caspersen, M. E. & Bennedsen, J. (2007). *Instructional design of a programming course*. *September 2007.* 111–122. https://doi.org/10.1145/1288580.1288595

CEDEFOP (2011). *Learning while working: Success stories on workplace learning in Europe*. Luxembourg: Publications Office of the European Union.

Cepeda, N. J., Coburn, N., Rohrer, D., Wixted, J. T., Mozer, M. C. & Pashler, H. (2009). Optimizing Distributed Practice Theoretical Analysis and Practical Implications. *Experimental Psychology*, *56*(4), 236–246. https://doi.org/10.1027/1618-3169.56.4.236

Cepeda, N. J., Vul, E., Rohrer, D., Wixted, J. T. & Pashler, H. (2008). Spacing effects in learning: a temporal ridgeline of optimal retention. *Psychological science, 19*(11), 1095–1102. https://doi.org/10.1111/j.1467-9280.2008.02209.x

Cepeda, N. J., Pashler, H., Vul, E., Wixted, J. T. & Rohrer, D. (2006). Distributed Practice in Verbal Recall Tasks: A Review and Quantitative Synthesis. *Psychological Bulletin*, *132*(3), 354–380. https://doi.org/10.1037/0033-2909.132.3.354

Chandler, P. & Sweller, J. (1991). Cognitive Load Theory and the Format of Instruction. *Cognition and Instruction*, *8*(4), 293–332. https://doi.org/10.1207/s1532690xci0804_2

Chelovechkov, A. & Spar, B. (2019). *3rd Annual 2019 Workplace Learning Report*. Retrieved from: https://learning.linkedin.com/content/dam/me/business/en-us/amp/learning-solutions/images/workplace-learning-report-2019/pdf/workplace-learning-report-2019.pdf. Accessed: May 30, 2022.

Chi, M., Glaser, R. & Rees, E. (1982). Expertise in problem solving. In R. Steinberg (Ed.), *Advances in the psychology of human intelligence* (pp. 7–75). Lawrence Erlbaum Associates Publishers, New Jersey.

Clariana, R. B., Wagner, D. & Murphy, L. C. R. (2000). Applying a connectionist description of feedback timing. *Educational Technology Research and Development, 48*(3), 5–21. https://doi.org/10.1007/BF02319855

Clark, R. C. & Mayer, R. E. (2016). *e-Learning and the science of instruction* (4th ed.). Hoboken, NJ: Wiley.

ClassMarker (2021). Retrieved from: www.classmarker.com. First accessed: August 12, 2021.

Cogliano, M. C., Kardash, C. A. M. & Bernacki, M. L. (2019). The effects of retrieval practice and prior topic knowledge on test performance and confidence judgments. *Contemporary Educational Psychology*, *56*, 117–129. https://doi.org/10.1016/j.cedpsych.2018.12.001

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, New Jersey: Lawrence Erlbaum Associates

Commission of the European Communities (2000). *A Memorandum on Lifelong Learning*, 1–36. https://uil.unesco.org/i/doc/lifelong-learning/policies/european-communities-a-memorandum-on-lifelong-learning.pdf. Accessed: October 13, 2021.

Conklin, J. (2005). Review Reviewed Work(s): A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives. *Educational Horizons*, *83*(4), 154–159. http://www.jstor.org/sTable/42926529

Connors, C. B. (2021). Summative and Formative Assessments: An Educational Polarity. *Kappa Delta Pi Record*, *57*(2), 70–74. https://doi.org/10.1080/00228958.2021.1890441

Corral, D., Carpenter, S. K. & Clingan-Siverly, S. (2021). The effects of immediate versus delayed feedback on complex concept learning. *Quarterly Journal of Experimental Psychology*, *74*(4), 786–799. https://doi.org/10.1177/1747021820977739

Corral, D., Healy, A. F., Rozbruch, E. V. & Jones, M. (2019). Building a testing-based training paradigm from cognitive psychology principles. *Scholarship of Teaching and Learning in Psychology, 5*(3), 189–208. https://doi.org/10.1037/stl0000146

Crook, T. R., Todd, S. Y., Combs, J. G., Woehr, D. J. & Ketchen, D. J. (2011). Does human capital matter? A meta-analysis of the relationship between human capital and firm performance. *Journal of Applied Psychology*, *96*(3), 443–456. https://doi.org/10.1037/a0022147

Cull, W. L. (2000). Untangling the benefits of multiple study opportunities and repeated testing for cued recall. *Applied Cognitive Psychology, 14*(3), 215–235. https://doi.org/10.1002/(SICI)1099-0720(200005/06)14:3<215::AID-ACP640>3.0.CO;2-1

Cull, W. L., Shaughnessy, J. J. & Zechmeister, E. B. (1996). Expanding Understanding of the Expanding-Pattern-of-Retrieval Mnemonic: Toward Confidence in Applicability. *Journal of Experimental Psychology: Applied*, *2*(4), 365–378. https://doi.org/10.1037/1076-898X.2.4.365

Dabbagh, N. & Bannan-Ritland, B. (2004). *Online learning: Concepts, strategies, and application*. Pearson Education, Inc. New Jersey.

Deakin Crick, R., Haigney, D., Huang, S., Coburn, T. & Goldspink, C. (2013). Learning power in the workplace: The effective lifelong learning inventory and its reliability and validity and implications for learning and development. *International Journal of Human Resource Management*, *24*(11), 2255–2272. https://doi.org/10.1080/09585192.2012.725075

Deisig, N., Sandoz, J. C., Giurfa, M. & Lachnit, H. (2007). The trial-spacing effect in olfactory patterning discriminations in honeybees. *Behavioural Brain Research, 176*(2), 314–322. https://doi.org/10.1016/j.bbr.2006.10.019

Delaney, P. F., Spirgel, A. S. & Toppino, T. C. (2012). A deeper analysis of the spacing effect after "deep" encoding. *Memory and Cognition*, *40*(7), 1003–1015. https://doi.org/10.3758/s13421-012-0207-3

Delaney, P. F., Verkoeijen, P. P. J. L. & Spirgel, A. (2010). Spacing and testing effects: A deeply critical, lengthy, and at times discursive review of the literature. In B. H. Ross (Ed.), *The psychology of learning and motivation: Advances in research and theory* (pp. 63–147). Elsevier Academic Press, Cambridge Massachusetts. https://doi.org/10.1016/S0079-7421(10)53003-2

Dempster, F. N. (1988). The spacing effect: A case study in the failure to apply the results of psychological research. *American Psychologist, 43*(8), 627–634. https://doi.org/10.1037/0003-066X.43.8.627

Dempster, F. N. (1989). Spacing effects and their implications for theory and practice. *Educational Psychology Review, 1*(4), 309–330. https://doi.org/10.1007/BF01320097

Dempster, F. N. (1996). Distributing and managing the conditions of encoding and practice. In E. L. Bjork & R. A. Bjork (Eds.), *Memory* (pp. 317–344). Academic Press. https://doi.org/10.1016/B978-012102570-0/50011-2

de Groot, A. (1965). *Thought and choice in chess* (2nd ed.). The Gruyter Mouton, The Hague.

Diamond, R. M. (1989). *Designing and improving courses and curricula in higher education: A systematic approach.* Jossey-Bass, Inc., San Francisco.

Dirksen, J. (2016). *Design for How People Learn* (2nd ed). New Riders, San Francisco.

Dixson, D. D. & Worrell, F. C. (2016). Formative and Summative Assessment in the Classroom. *Theory into Practice*, *55*(2), 153–159. https://doi.org/10.1080/00405841.2016.1148989

Dondi, M., Klier, J., Panier, F. & Schubert, J. (2021). *Defining the skills citizens will need in the future world of work*. Retrieved from: https://www.mckinsey.com/industries/public-and-social-sector/our-insights/defining-the-skills-citizens-will-need-in-the-future-world-of-work. Accessed: July 20, 2022.

Donovan, J. J. & Radosevich, D. J. (1999). A meta-analytic review of the distribution of practice effect: now you see it, now you don't. *Journal of Applied Psychology, 84*(5), 795–805. https://doi.org/10.1037/0021- 9010.84.5.795

Dunlosky, J. A. & Metcalfe, J. (2009). *Metacognition*. Sage Publications, Inc., Thousand Oaks, California.

Dunlosky, J. A, Rawson, K. A., Marsh, E. J., Nathan, M. J. & Willingham, D. T. (2013). Improving Students' Learning With Effective Learning Techniques: Promising Directions From Cognitive and Educational Psychology. *Psychological Science in the Public Interest*, *14*(1), 4–58. https://doi.org/10.1177/1529100612453266

Dunlosky, J. A. & Rawson, K. A. (2015). Practice tests, spaced practice, and successive relearning: Tips for classroom use and for guiding students' learning. *Scholarship of Teaching and Learning in Psychology*, *1*(1), 72–78. https://doi.org/10.1037/stl0000024

Ebbinghaus, H. (1885/1962). *Memory: A contribution to experimental psychology*. Dover Publications, New York.

Elfert, M. (2015). UNESCO, the Faure Report, the Delors Report, and the Political Utopia of Lifelong Learning. *European Journal of Education*, *50*(1), 88–100. https://doi.org/10.1111/ejed.12104

Ellinger, A. D. (2004). The concept of self-directed learning and its implications for human resource development. *Advances in Developing Human Resources, 6*(2), 158–177. https://doi.org/10.1177/1523422304263327

Ellis, P. F. & Kuznia, K. D. (2014). Corporate Elearning Impact On Employees. *Global Journal of Business Research, 8*(4), 1–15.

English, L. M. & Mayo, P. (2021). *Lifelong Learning, Global Social Justice, and Sustainability*. Palgrave Pivot, California.

Ertmer, P. A. & Newby, T. J. (2013). Behaviorism, Cognitivism, Constructivism: Comparing Critical Features From an Instructional Design Perspective. *Performance Improvement Quarterly*, *26*(2), 43–71. https://doi.org/10.1111/j.1937-8327.1993.tb00605.x

European Commission, Directorate-General for Employment, Social Affairs and Inclusion (2018). *Promoting adult learning in the workplace : final report of the ET 2020 working group 2016 – 2018 on adult learning*. Publications Office. https://data.europa.eu/doi/10.2767/865359

European Commission/Joint Research Centre, Sala, A., Punie, Y., Garkov, V. (2020). *LifeComp : the European Framework for personal, social and learning to learn key competence*. Publications Office. https://data.europa.eu/doi/10.2760/922681

European Parliament (2008). *Adult learning: it is never too late to learn European Parliament resolution of 16 January 2008 on adult learning: it is never too late to learn* (2007/2114(INI)) (OJ C, C/41, 19.02.2009, p. 46 https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52008IP0013)

Finn, B. & Tauber, S. K. (2015). When Confidence Is Not a Signal of Knowing: How Students' Experiences and Beliefs About Processing Fluency Can Lead to Miscalibrated Confidence. *Educational Psychology Review*, *27*(4), 567–586.

https://doi.org/10.1007/s10648-015-9313-7

Foot-Seymour, V., Foot, J. & Wiseheart, M. (2019). Judging credibility: Can spaced lessons help students think more critically online? *Applied Cognitive Psychology*, *33*(6), 1032–1043. https://doi.org/10.1002/acp.3539

Fritz, C. O., Morris, P. E., Nolan, D. & Singleton, J. (2007). Expanding retrieval practice: an effective aid to preschool children'slearning. *Quarterly Journal of Experimental Psychology, 60*(7), 991–1004. https://doi. org/10.1080/17470210600823595

Gagné, R. M., Wager, W. W., Golas, K. C. & Keller, J. M. (2005). *Principles of instructional design* (5th ed.). Cengage Learning, Inc., Boston, Massachussetts.

Gallardo, K. (2021). The Importance of Assessment Literacy: Formative and Summative Assessment Instruments and Techniques. In R. Babo, N. Dey & A. S. Ashour, *Workgroups eAssessment: Planning, Implementing and Analysing Frameworks* (pp. 3–25). Springer, Singapore.

Gartner (2018). *Employee expectations of work changes resulting from digitalization*. Retrieved from: https://www.gartner.com/en/human-resources/role/learning-development. Accessed: May 22, 2022.

Gazzaniga, M. S., Ivry, R. B. & Mangun, G. R. (2014). *Cognitive Neuroscience The Biology of the Mind* (4th ed.). W. W. Norton & Company, Inc., New York.

Geary, D. C. (2002). Principles of evolutionary educational psychology. *Learning and Individual Differences, 12*(4), 317–345. https://doi.org/10.1016/S1041-6080(02)00046-8

Geary, D. C. (2008). An evolutionarily informed education science. *Educational Psychologist, 43*(4), 179–195. https://doi.org/10.1080/00461520802392133

Geary, D. C. (2011). Evolutionary Educational Psychology. In K. Harris, S. Graham & T. Urdan (Eds.), *APA educational psychology handbook* (vol. 1, pp. 597–621). American

Psychological Association, Washington DC.

Geary, D. & Berch, D. (2016). Evolution and children's cognitive and academic development. In D. Geary & D. Berch (Eds.), *Evolutionary perspectives on child development and education* (pp. 217–249). Springer Nature, Switzerland.

Gerber, A. S. (2011). Field experiments in political science. In Druckman, J. N., Green, D. P., Kulinski, J. H. & Lupia, A. (Eds.), *Cambridge handbook of experimental political science* (pp. 206–251). Cambridge University Press, New York.

Gerbier, E. & Toppino, T. C. (2015). The effect of distributed practice: Neuroscience, cognition, and education. *Trends in Neuroscience and Education*, *4*, 49–59. https://doi.org/10.1016/j.tine.2015.01.001

Glaveski, S. (2019). *Where Companies Go Wrong with Learning and Development*. Retrieved from: https://hbr.org/2019/10/where-companies-go-wrong-with-learning-and-development. Accessed: May 30, 2022.

Glenberg, A. M. (1976). Monotonic and Nonmonotonic Lag Effects in Paired-Associate and Recognition Memory Paradigms. *Journal of Verbal Learning and Verbal Behavior*, *15*(1), 1–16. https://doi.org/10.1016/S0022-5371(76)90002-5

Glenberg, A. M. (1979). Component-levels theory of the effects of spacing of repetitions on recall and recognition. *Memory & Cognition*, *7*(2), 95–112. https://doi.org/10.3758/BF03197590

Göldi, S. (2011). *Von der bloomschen Taxonomy zu aktuellen Bildungsstandards : zur Entstehungs- und Rezeptionsgeschichte eines pädagogischen Bestsellers* (1st ed.). hep Verlag, Switzerland.

Goldfinch, J. & Hughes, M. (2007). Skills, learning styles and success of first-year undergraduates. *Active Learning in Higher Education, 8*(3), 259–273.

https://doi.org/10.1177/1469787407081881

Greene, R. L. (1989). Spacing Effects in Memory: Evidence for a Two-Process Account. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*(3), 371–377. https://doi.org/10.1037/0278-7393.15.3.371

Greene, R. L. (2008). Spacing and repetition effects. In H. L. Roediger III (Ed.), *Learning and memory: A comprehensive reference. Cognitive psychology of memory* (vol. 2. pp. 65–78). Elsevier, Amsterdam. https://doi.org/10.1016/B978-012370509-9.00168-6

Grote, M. G. (1995). The Effect of Massed Versus Spaced Practice on Retention and Problem-Solving in High School Physics. *The Ohio Journal of Science*, *95*(3), 243–247. http://hdl.handle.net/1811/23655

Gryger, L., Saar, T. & Schaar, P. (2010). *Building organizational capabilities: McKinsey Global Survey results*. Retrieved from: https://www.mckinsey.com/business-functions/people-and-organizational-performance/our-insights/building-organizational-capabilities-mckinsey-global-survey-results. Accessed: May 22, 2022.

Gustafson, K. L. & Branch, R. M. (1997). *Survey of instructional development models* (3rd ed.). ERIC Clearing-house on Information and Technology, Syracuse, New York.

Gustafson, K. L. & Branch, R. M. (2002). *Survey of instructional development models* (4th ed.). ERIC Clearing-house on Information and Technology, Syracuse, New York.

Halawi, L. A., McCarthy, R. V. & Pires, S. (2009). An Evaluation of E-Learning on the Basis of Bloom’s Taxonomy: An Exploratory Study. *Journal of Education for Business*, *84*(6), 374–380. https://doi.org/10.3200/JOEB.84.6.374-380

Hanushek, E. A. & Wößmann, L. (2010). Education and Economic Growth. In P. Peterson, E. Baker, B. McGaw, *International Encyclopedia of Education* (3rd ed., pp. 245–252). Elsevier Science. https://doi.org/10.1016/B978-0-08-044894-7.01227-6

Hanushek, E. & Wößmann, L. (2020). *The economic impacts of learning losses, OECD Education Working Papers*, No. 225, OECD Publishing, Paris. https://doi.org/10.1787/21908d74-en

Harris, R. & Clayton, B. (2019). The current emphasis on learning outcomes. *International Journal of Training Research*, *17*(2), 93–97. https://doi.org/10.1080/14480220.2019.1644777

Hattie, J. & Timperley, H. (2007). The Power of Feedback. *Review of Educational Research, 77*(1), 81–112. https://doi.org/10.3102/003465430298487

Hemmerich, W. (2018). *StatistikGuru: Stichprobengröße für die einfaktorielle ANOVA berechnen*. Retrieved from: https://statistikguru.de/rechner/stichprobengroesse-einfaktorielle-anova.html. Accessed: August 1, 2021.

Hintzman, D. L. (1974). Theoretical implications of the spacing effect. In R. L. Solso (Ed.), *Theories in cognitive psychology: The Loyola Symposium*. Lawrence Erlbaum Associates Publishers, New Jersey.

Hintzman, D. L. & Block, R. A. (1973). Memory for the spacing of repetitions. *Journal of Experimental Psychology, 99*(1), 70–74. https://doi.org/10.1037/h0034761

Houston, D. & Thompson, J. N. (2017). Blending formative and summative assessment in a capstone subject: 'it's not your tools, it's how you use them.' *Journal of University Teaching and Learning Practice*, *14*(3), 2–13. https://doi.org/10.53761/1.14.3.2

Huppelsberg, J. & Walter, K. (2009). Kurzlehrbuch Physiologie (3rd ed.). Georg Thieme Verlag KG, Stuttgart.

International Labour Organization (ILO) (2021). *Skilling, upskilling and reskilling of employees, apprentices & interns during the COVID-19 pandemic*. ILO, Geneva.

Issa, N., Mayer, R. E., Schuller, M., Wang, E., Shapiro, M. B. & Darosa, D. A. (2013). Teaching for understanding in medical classrooms using multimedia design principles. *Medical Education*, *47*(4), 388–396. https://doi.org/10.1111/medu.12127

Issa, N., Schuller, M., Santacaterina, S., Shapiro, M., Wang, E., Mayer, R. E. & Darosa, D. A. (2011). Applying multimedia design principles enhances learning in medical education. *Medical Education*, *45*(8), 818–826. https://doi.org/10.1111/j.1365-2923.2011.03988.x

James, N. & Thériault, V. (2020). Adult education in times of the COVID-19 pandemic: Inequalities, changes, and resilience. *Studies in the Education of Adults*, *52*(2), 129–133. https://doi.org/10.1080/02660830.2020.1811474

Jastrzembski, T. S. & Gluck, K. A. (2009). A formal comparison of model variants for performance prediction. In *Proceedings of the 9th International Conference of Cognitive Modeling* (July 24-26, 2009; paper 258), Manchester, United Kingdom.

Joseph, S. V. (2019). *The Secrets To Keeping Employees Engaged And Productive*. Retrieved from: https://www.forbes.com/sites/shelcyvjoseph/2019/08/15/the-secrets-to-keeping-employees-engaged-and-productive/?sh=46484c9c4310. Accessed: May 25. 2022.

Kalenberg, K. (2017). Spaced and expanded practice: an investigation ofmethods to enhance retention. J*ournal of Undergraduate Research at Minnesota State University, Mankato, 17*, 18. https://cornerstone.lib.mnsu.edu/jur/vol17/iss1/6

Kane, G. C., Palmer, D., Phillips, A. N., Kiron, D. & Buckley, N. (2018). Coming of Age Digitally. In *MIT Sloan Management Review and Deloitte Insights* (Issue June). MIT Sloan Management Review.

Kang, S. H. K. (2016). Spaced Repetition Promotes Efficient and Effective Learning: Policy Implications for Instruction. *Policy Insights from the Behavioral and Brain Sciences*, *3*(1), 12–19. https://doi.org/10.1177/2372732215624708

Kang, S. H. K. & Pashler, H. (2012). Learning painting styles: Spacing is advantageous when it promotes discriminative contrast. *Applied Cognitive Psychology, 26*(1), 97–103. https://doi.org/10.1002/acp.1801

Kapler, I. V., Weston, T. & Wiseheart, M. (2015). Spacing in a simulated undergraduate classroom: Long-term benefits for factual and higher-level learning. *Learning and Instruction*, *36*, 38–45. https://doi.org/10.1016/j.learninstruc.2014.11.001

Karpicke, J. D. & Bauernschmidt, A. (2011). Spaced retrieval: Absolute spacing enhances learning regardless of relative spacing. Journal of Experimental Psychology: Learning, *Memory, and Cognition, 37*(5), 1250–1257. https://doi.org/10.1037/a0023436

Karpicke, J. D., Blunt, J. R. & Smith, M. A. (2016). Retrieval-based learning: Positive effects of retrieval practice in elementary school children. *Frontiers in Psychology, 7*, Article 350. https://doi.org/10.3389/fpsyg.2016.00350

Karpicke, J. D., Butler, A. C. & Roediger, H. L., III (2009). Metacognitive strategies in student learning: Do students practise retrieval when they study on their own? *Memory, 17*(4), 471–479. https://doi.org/10.1080/09658210802647009.

Karpicke, J. D. & Roediger, H. L. (2007a). Expanding Retrieval Practice Promotes Short-Term Retention, but Equally Spaced Retrieval Enhances Long-Term Retention. *Journal of Experimental Psychology: Learning Memory and Cognition*, *33*(4), 704–719. https://doi.org/10.1037/0278-7393.33.4.704

Karpicke, J. D. & Roediger, H. L. (2007b). Repeated retrieval during learning is the key to long-term retention. *Journal of Memory and Language*, *57*(2), 151–162. https://doi.org/10.1016/j.jml.2006.09.004

Karpicke, J. D. & Roediger, H. L. (2010). Is expanding retrieval a superior method for learning text materials? *Memory & Cognition, 38*(1), 116–124. https://doi.org/10.3758/MC.38.1.116

Khalil, M. K. & Elkhider, I. A. (2016). Applying learning theories and instructional design models for effective instruction. *Advances in Physiology Education*, *40*(2), 147–156. https://doi.org/10.1152/advan.00138.2015

Kirchherr, J., Klier, J., Meyer-Guckel, V. & Winde, M. (2020). *Die Zukunft der Qualifizierung in Unternehmen nach Corona.* Stifterverband für die Deutsche Wissenschaft e.V., Essen. Available at: https://www.stifterverband.org/medien/die-zukunft-der-qualifizierung-in-unternehmen-nach-corona

Kirkpatrick, D. L. (1967). Evaluation of training. In R. L. Craig & L. R. Bittel (Eds.), *Training and development handbook: A guide to human resource development* (pp. 87–112). New York: McGraw-Hill.

Kirschner, F., Kester, L. & Corbalan, G. (2011). Cognitive load theory and multimedia learning, task characteristics and learning engagement: The Current State of the Art. *Computers in Human Behavior*, *27*(1), 1–4. https://doi.org/10.1016/j.chb.2010.05.003

Klees, S. J. (2016). Human Capital and Rates of Return: Brilliant Ideas or Ideological Dead Ends? *Comparative Education Review 60*(4), 644-672. https://doi.org/10.1086/688063

Knight, P. T. (2002). Summative assessment in higher education: Practices in disarray. *Studies in Higher Education*, *27*(3), 275–286. https://doi.org/10.1080/03075070220000662

Kondratjew, H. & Kahrens, M. (2018). Leveraging experiential learning training through spaced learning. *Journal of Work-Applied Management, 11*(2), 30–52. https://doi.org/10.1108/JWAM-05-2018-0011

Kontoghiorghes, C., Awbre, S. M. & Feurig, P. L. (2005). Examining the relationship between learning organisation characteristics and change adaptation, innovation, and organizational performance. *Human Resource Development Quarterly, 16*(2), 185–212. https://doi.org/10.1002/hrdq.1133

Kornell, N. & Bjork, R. A. (2007). The promise and perils of self-regulated study. *Psychonomic Bulletin & Review, 14*(2), 219–224. https://doi.org/10.3758/BF03194055

Kornell, N. & Bjork, R. A. (2008). Learning Concepts and Categories Is Spacing the "Enemy of Induction"? *Psychological Science*, *19*(6), 585–592. https://doi.org/10.1111/j.1467-9280.2008.02127.x

Kornell, N., Castel, A. D., Eich, T. S. & Bjork, R. A. (2010). Spacing as the friend of both memory and induction in young and older adults. *Psychology and Aging, 25*(2), 498–503. https://doi.org/10.1037/a0017807

Krathwohl, D. R. (2002). A Revision of Bloom's Taxonomy: An Overview. *Theory into Practice*, *41*(4), 212–218. https://doi.org/10.1207/s15430421tip4104_2

Kubbe, I. (2018). Experimente und experimentelle Forschungsdesigns. In: Wagemann, C., Goerres, A., Siewert, M. (Eds.) *Handbuch Methoden der Politikwissenschaft*. Springer Reference Sozialwissenschaften. Springer VS, Wiesbaden. https://doi.org/10.1007/978-3-658-16937-4_7-1

Küpper-Tetzel, C. E. & Erdfelder, E. (2012). Encoding, maintenance, and retrieval processes in the lag effect: A multinomial processing tree analysis. *Memory*, *20*(1), 37–47. https://doi.org/10.1080/09658211.2011.631550

Küpper-Tetzel, C. E., Kapler, I. V. & Wiseheart, M. (2014). Contracting, equal, and expanding learning schedules: the optimal distribution of learning sessions depends on retention interval. *Memory and Cognition, 42*(5), 729–741. https://doi.org/10.3758/s13421-014-0394-1.

Kugler, F., Piopiunik, M. & Wößmann, L. (2017). *Bildung hat Zukunft - Bildungsstudie 2017*. Union Investment, Frankfurt. Available at: https://www.ifo.de/publikationen/2017/monographie-autorenschaft/bildung-hat-zukunft-bildungsstudie-2017

Kulik, J. A. & Kulik, C. L. C. (1988). Timing of Feedback and Verbal Learning. *Review of Educational Research, 58*(1), 79–97. https://doi.org/10.3102/00346543058001079

Landauer, T. K. (1969). Reinforcement as consolidation. *Psychological Review, 76*(1), 82–96. https://doi.org/10.1037/h0026746

Landauer, T. K. & Bjork, R. A. (1978). Optimum rehearsal patterns and name learning. In M. M. Gruneberg, P. E. Morris & R. N. Sykes (Eds.), *Practical aspects of memory* (pp. 625–632). Academic Press, London.

Lapkin, S., Hart, D. & Harley, B. (1998). Case study of compact core French models: Attitudes and achievement. In S. Lapkin (Ed.), *French second language education in Canada: Empirical studies* (pp. 3–30). University of Toronto Press, Toronto.

Larsen, D. P. (2018). Planning education for long-term retention: the cognitive science and implementation of retrieval practice. *Seminars in Neurology, 38*(4), 449–456. https://doi.org/10.1055/s-0038-1666983

Latimier, A., Peyre, H. & Ramus, F. (2021). A Meta-Analytic Review of the Benefit of Spacing out Retrieval Practice Episodes on Retention. *Educational Psychology Review*, *33*(3), 959–987. https://doi.org/10.1007/s10648-020-09572-8

Lengrand, P. (1970). *An Introduction to Lifelong Education*. UNESCO, Paris.

Lightbown, P. M. & Spada, N. (1994). An Innovative Program for Primary ESL Students in Quebec. *TESOL Quarterly, 28*(3), 563–579. https://doi.org/10.2307/3587308

Lindsey, R., Mozer, M. C., Cepeda, N. J. & Pashler, H. (2009). Optimizing memory retention with cognitive models. In A. Howes, D. Peebles & R. Cooper (Eds.), *Proceedings of the 9th International Conference of Cognitive Modeling* (July 24-26, 2009; paper 258), Manchester, United Kingdom.

Logan, J. M. & Balota, D. A. (2008). Expanded vs. equal interval spaced retrieval practice: Exploring different schedules of spacing and retention interval in younger and older adults. *Aging, Neuropsychology, and Cognition*, *15*(3), 257–280. https://doi.org/10.1080/13825580701322171

Maddox, G. B. (2016). Understanding the underlying mechanism of the spacing effect in verbal learning: a case for encoding variability and study- phase retrieval. *Journal of Cognitive Psychology, 28*(6), 684–706. https://doi.org/10.1080/20445911.2016.1181637

Manyika, J., Lund, S., Chui, M., Bughin, J., Woetzel, J., Batra, P., Ko, R. & Sanghvi, S. (2017). Jobs Lost, Jobs Gained: Workforce Transitions in a Time of Automation Executive Summary. In Mckinsey Global Institute, *Jobs Lost, Jobs Gained: Workforce Transitions in a Time of Automation* (Issue December). McKinsey Global Insitute.

Marquez-Ramos, L. & Mourelle, E. (2019). Education and economic growth: an empirical analysis of nonlinearities. *Applied Economic Analysis*, *27*(79), 21–45. https://doi.org/10.1108/AEA-06-2019-0005

Mayer, R. E. (2019). Thirty years of research on online learning. *Applied Cognitive Psychology*, *33*(2), 152–159. https://doi.org/10.1002/acp.3482

Mayer, R. E., Fennell, S., Farmer, L. & Campbell, J. (2004). A personalization effect in multimedia learning: Students learn better when words are in conversational style rather than formal style. *Journal of Educational Psychology*, *96*(2), 389–395. https://doi.org/10.1037/0022-0663.96.2.389

Mayer, R. E. & Moreno, R. (1998). A split-attention effect in multimedia learning: Evidence for dual processing systems in working memory. *Journal of Educational Psychology, 90*(2), 312–320. https://doi.org/10.1037/0022-0663.90.2.312

Mayer, R. E. & Moreno, R. (2002). Aids to computer-based multimedia learning. *Learning and Instruction*, *12*(1), 107–119. https://doi.org/10.1016/S0959-4752(01)00018-4

Mayer, R. E. & Moreno, R. (2003). Nine ways to reduce cognitive load in multimedia learning. *Educational Psychologist, 38*(1), 43–52. https://doi.org/10.1207/S15326985EP3801_6

McDaniel, M. A., Agarwal, P. K., Huelser, B. J., McDermott, K. B. & Roediger, H. L. (2011). Test-Enhanced Learning in a Middle School Science Classroom: The Effects of Quiz Frequency and Placement. *Journal of Educational Psychology*, *103*(2), 399–414. https://doi.org/10.1037/a0021782

Melton, A. W. (1970). The Situation with Respect to the Spacing of Repetitions and Memory. *Journal of Verbal Learning and Verbal Behavior*, *9*(5), 596–606. https://doi.org/10.1016/S0022-5371(70)80107-4

Merrill, M. D. (2002). A pebble-in-the-pond model for instructional design. *Performance Improvement, 41*(7), 41–46. https://doi.org/10.1002/pfi.4140410709

Mettler, E., Massey, C. M. & Kellman, P. J. (2016). A comparison of adaptive and fixed schedules of practice. *Journal of Experimental Psychology. General, 145*(7), 897–917. https://doi.org/10.1037/xge0000170

Milana, M. (2012). Political globalization and the shift from adult education to lifelong learning. *European Journal for Research on the Education and Learning of Adults*, *3*(2), 103–117. https://doi.org/10.3384/rela.2000-7426.rela0070

Miller, G. A. (1956). The magical number, plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 63*(2), 81–97.

https://doi.org/10.1037/h0043158

Milligan, C., Littlejohn, A. & Margaryan, A. (2014). Workplace learning in informal networks. *Reusing Open Resources: Learning in Open Networks for Work, Life and Education*, *March*, 93–103. https://doi.org/10.5334/2014-06

Moore, D. A. & Healy, P. J. (2008). The trouble with overconfidence. *Psychological Review, 115*(2), 502–517. https://doi.org/10.1037/0033-295X.115.2.502

Morehead, K., Rhodes, M. G. & DeLozier, S. (2015). Instructor and student knowledge of study strategies. *Memory*, *24*(2), 257–271. https://doi.org/10.1080/09658211.2014.1001992

Moritz, B. E., Ellis, K., Kayser, H., Chao, R., Ryan, T., Oldfield, R., Hyde, S., Cobourn, B., Davies, M., Steffen, I. & Mallon, H. (2019). *22nd Annual Global CEO Survey - CEOs' curbed confidence spells caution*. Retrieved from: https://www.pwc.de/de/ceosurvey2019/pwc-22nd-annual-global-ceo-survey-report.pdf. Accessed: May 25, 2022

Mozer, M. C., Pashler, H., Cepeda, N., Lindsey, R. & Vul, E. (2009). Predicting the optimal spacing of study: A multiscale context model of memory. In Y. Bengio, D. Schuurmans, J. Lafferty, C.K.I. Williams & A. Culotta (Eds.), *Advances in Neural Information Processing Systems 22* (pp. 1321–1329). NIPS Foundation, La Jolla, California

Mullet, H. G., Butler, A. C., Verdin, B., von Borries, R. & Marsh, E. J. (2014). Delaying feedback promotes transfer of knowledge despite student preferences to receive feedback immediately. *Journal of Applied Research in Memory and Cognition, 3*(3), 222–229. https://doi.org/10.1016/j.jarmac.2014.05.001

Mulligan, N. W. & Peterson, D. J. (2014). The spacing effect and metacognitive control. *Journal of Experimental Psychology: Learning Memory and Cognition*, *40*(1), 306–311.

https://doi.org/10.1037/a0033866

Murphy, E. J. (2007). Prior learning assessment: A review of bloom's taxonomy and kolb's theory of experiential learning: Practical uses for prior learning assessment. *Journal of Continuing Higher Education*, *55*(3), 64–66. https://doi.org/10.1080/07377366.2007.10400135

Newby, T. J., Stepich, D., Lehman, J. & Russell, J. D. (1996). *Instructional technology for teaching and learning: Designing, integrating computers, and using media*. Upper Saddle River, NJ: Pearson Education, Inc.

Ogunleye, J. (2013). Workplace learning as a linchpin of Europe's lifelong learning policy: An examination of national policies with particular reference to people with long-term mental illness. *Higher Education, Skills and Work-Based Learning*, *3*(3), 180–195. https://doi.org/10.1108/HESWBL-04-2011-0025

Organisation for Economic Co-operation and Development (OECD) (2001). *Economics and Finance of Lifelong Learning.* OECD Publishing, Paris. https://doi.org/10.1787/9789264195868-en

Organisation for Economic Co-Operation and Development (OECD) (2004). Lifelong Learning (Policy Briefs series). Paris: OECD Secretary-General.

Organisation for Economic Co-operation and Development (OECD) (2007). *Bildung auf einen Blick 2007: OECD-Indikatoren*. W. Bertelsmann Verlag, Bielefeld. https://doi.org/10.1787/eag-2007-de

Organisation for Economic Co-operation and Development (OECD) (2019). *OECD Employment Outlook 2019: The Future of Work.* OECD Publishing, Paris. https://doi.org/10.1787/9ee00155-en

Organisation for Economic Co-operation and Development (OECD) (2021). *OECD Skills*

Outlook 2021: Learning for Life. OECD Publishing, Paris. https://doi.org/10.1787/0ae365b4-en

Paas, F. & Ayres, P. (2014). Cognitive Load Theory: A Broader View on the Role of Memory in Learning and Education. *Educational Psychology Review*, *26*(2), 191–195. https://doi.org/10.1007/s10648-014-9263-5

Paas, F., Gog, T. Van & Sweller, J. (2010). Cognitive Load Theory : New Conceptualizations, Specifications, and Integrated Research Perspectives. *Educational Psycology Review, 22*(2), 115–121. https://doi.org/10.1007/s10648-010-9133-8

Paas, F., Renkl, A. & Sweller, J. (2003). Cognitive load theory and instructional design: Recent developments. *Educational Psychologist*, *38*(1), 1–4. https://doi.org/10.1207/S15326985EP3801_1

Paas, F., Renkl, A. & Sweller, J. (2004). *Cognitive Load Theory : Instructional Implications of the Interaction between Information Structures and Cognitive Architecture*, *32*, 1–8. https://doi.org/10.1023/B:TRUC.0000021806.17516.d0

Paas, F. & Sweller, J. (2012). An Evolutionary Upgrade of Cognitive Load Theory: Using the Human Motor System and Collaboration to Support the Learning of Complex Cognitive Tasks. *Educational Psychology Review*, *24*(1), 27–45. https://doi.org/10.1007/s10648-011-9179-2

Pashler, H., Rohrer, D., Cepeda, N. J. & Carpenter, S. K. (2007). Enhancing learning and retarding forgetting: Choices and consequences. *Psychonomic Bulletin and Review*, *14*(2), 187–193. https://doi.org/10.3758/BF03194050

Pavlik, P. I. & Anderson, J. R. (2005). Practice and forgetting effects on vocabulary memory: An activation-based model of the spacing effect. *Cognitive Science*, *29*(4), 559–586. https://doi.org/10.1207/s15516709cog0000_14

Perera-Diltz, D. & Moe, J. (2014). Formative and Summative Assessment in Online Education. *Journal of Research in Innovative Teaching*, *7*(1), 130–142. https://digitalcommons.odu.edu/chs_pubs/37

Peterson, L. R. & Peterson, M. J. (1959). Short-term retention of individual verbal items. *Journal of Experimental Psychology, 58*, 193–198. https://doi.org/10.1037/h0049234

Pfeffer, J. & Veiga, J. F. (1999). Putting people first for organizational success. *Academy of Management Perspectives, 13*(2), 37–48. https://www.jstor.org/stable/4165538

Pimsleur, P. (1967). A Memory Schedule. *The Modern Language Journal*, *51*(2), 73–75. https://doi.org/10.1111/j.1540-4781.1967.tb06700.x

Poquet, O. & de Laat, M. (2021). Developing capabilities: Lifelong learning in the age of AI. *British Journal of Educational Technology*, *52*(4), 1695–1708. https://doi.org/10.1111/bjet.13123

Porth, S. J., McCall, J. & Bausch, T. A. (1999). Spiritual themes of the 'learning organization'. *Journal of Organizational Change Management, 12*(3), 211–220. https://doi.org/10.1108/09534819910273883

Putnam, A. L., Sungkhasettee, V. W. & Roediger, H. L. (2016). Optimizing Learning in College: Tips From Cognitive Psychology. *Perspectives on Psychological Science*, *11*(5), 652–660. https://doi.org/10.1177/1745691616645770

Raaijmakers, J. G. W. (2003). Spacing and repetition effects in human memory: Application of the SAM model. *Cognitive Science*, *27*(3), 431–452. https://doi.org/10.1016/S0364-0213(03)00007-7

Radmehr, F. & Drake, M. (2019). Revised Bloom's taxonomy and major theories and frameworks that influence the teaching, learning, and assessment of mathematics: a comparison. *International Journal of Mathematical Education in Science and*

Technology, *50*(6), 895–920. https://doi.org/10.1080/0020739X.2018.1549336

Rea, C. P. & Modigliani, V. (1985). The effect of expanded versus massed practice on the retention of multiplication facts and spelling lists. *Human Learning: Journal of Practical Research and Applications, 4*(1), 11–18.

Rea, C. P. & Modigliani, V. (1987). The spacing effect in 4- to 9-year-old children. *Memory & Cognition, 15*(5), 436–443. https://doi.org/10.3758/BF03197733

Rees, G. (2010). The political economy of adult education. In P. Peterson, E. Baker & B. McGaw, *International Encyclopedia of Education* (3rd ed., pp. 272–281). Elsevier Science. https://doi.org/10.1016/B978-0-08-044894-7.00046-4

Reynolds, J.H. & Glaser, R.E. (1964). Effects of repetition and spaced review upon retention of a complex learning task. *Journal of Educational Psychology, 55,* 297-308. https://doi.org/10.1037/h0040734

Richards, D. & Taylor, M. (2015). A comparison of learning gains when using a 2D simulation tool versus a 3D virtual world: An experiment to find the right representation involving the Marginal Value Theorem. *Computers & Education, 86,* 157–171. https://doi.org/10.1016/j.compedu.2015.03.009

Robinson, E. S. (1921). The Relative Efficiencies of Distributed and Concentrated Study in Memorizing. *Journal of Experimental Psychology, 4*(5), 327–343. https://doi.org/10.1037/h0073490

Rodriguez, F., Kataoka, S., Janet Rivas, M., Kadandale, P., Nili, A. & Warschauer, M. (2021). Do spacing and self-testing predict learning outcomes? *Active Learning in Higher Education*, *22*(1), 77–91. https://doi.org/10.1177/1469787418774185

Roediger, H. L. & Karpicke, J. D. (2006a). Test-Enhanced Learning: Taking Memory Tests Imporves Long-Term Retention. *Psychological Science*, *17*(3), 249–255.

https://doi.org/10.1111/j.1467-9280.2006.01693.x

Roediger, H. L. & Karpicke, J. D. (2006b). The Power of Testing Memory Basic Research and Implications for Educational Practice. *Perspectives on Psychological Science*, *1*(3), 181–210. https://doi.org/10.1111/j.1745-6916.2006.00012.x

Rohrer, D. (2009). The effects of spacing and mixing practice problems. *Journal for Research in Mathematics Education, 40*, 4–17. https://doi.org/10.5951/jresematheduc.40.1.0004

Rohrer, D. & Pashler, H. (2007). Increasing Retention Without Increasing Study Time. *Current Directions in Psychological Science, 16*(4), 183–186. https://doi.org/10.1111/j.1467-8721.2007.00500.x

Rohrer, D. & Taylor, K. (2006). The Effects of Overlearning and Distributed Practise on the Retention of Mathematics Knowledge. *Applied Cognitive Psychology*, *20*(9), 1209–1224. https://doi.org/10.1002/acp.1266

Rohrer, D. & Taylor, K. (2007). The shuffling of mathematics problems improves learning. *Instructional Science*, *35*(6), 481–498. https://doi.org/10.1007/s11251-007-9015-8

Rose, R. C., Kumar, N. & Pak, O. G. (2009). The effect of organizational learning on organizational commitment, job satisfaction and work performance. *Journal of Applied Business Research, 25*(6), 55–66. https://doi.org/10.19030/jabr.v25i6.995

Roth, G. (2011). *Bildung braucht Persönlichkeit. Wie Lernen gelingt* (4th ed.). Klett-Cotta, Stuttgart.

Rubenson, K. (2019). Assessing the status of lifelong learning: Issues with composite indexes and surveys on participation. *International Review of Education*, *65*(2), 295–317. https://doi.org/10.1007/s11159-019-09768-3

Ryu, G. & Moon, S. G. (2019). The effect of actual workplace learning on job satisfaction and organizational commitment. *Journal of Workplace Learning, 31*(8), 481–497.

https://doi.org/10.1108/JWL-05-2019-0061

Sala, G. & Gobet, F. (2017). Does far transfer exist? Negative evidence from chess, music, and working memory training. *Current Directions in Psychological Science*, *26*(6), 515–520. https://doi.org/10.1177/0963721417712760

Saunders, M., Lewis, P. and Thornhill, A. (2016) Research Methods for Business Students. 7th Edition, Pearson, Harlow.

Schacter, D. L. (1996). *Searching for memory: The brain, the mind, and the past*. Basic Books, New York.

Schellekens, L. H., Bok, H. G. J., de Jong, L. H., van der Schaaf, M. F., Kremer, W. D. J. & van der Vleuten, C. P. M. (2021). A scoping review on the notions of Assessment as Learning (AaL), Assessment for Learning (AfL), and Assessment of Learning (AoL). *Studies in Educational Evaluation*, *71*, 1–15. https://doi.org/10.1016/j.stueduc.2021.101094

Schmidt, R., Lee, T., Winstein, C., Wulf, G. & Zelaznik, H. (2019). *Motor Control and Learning* (6th ed). Human Kinetics.

Schmidt-Hertha, B. (2018). Bildung im Erwachsenenalter. In R. Tippelt & B. Schmidt-Hertha, *Handbuch Bildungsforschung* (4 th ed.). Springer VS, Wiesbaden.

Schnotz, W. & Kürschner, C. (2007). A consideration of cognitive load theory. *Educational Psychology Review*, *19*(4), 469–508. https://doi.org/10.1007/s10648-007-9053-4

Schönherr, K. W. & Tiberius, V. (2014). *Lebenslanges Lernen Wissen und Können als Wohlstandsfaktoren*. Springer VS, Wiesbaden.

Schunk, D. H. (2012). *Learning Theories, an Educational Perspective* (6th ed.). Pearson Education Inc., Boston Massachussetts.

Scriven, M. (1967). The Methodology of Evaluation. In R. Tyler, R. Gagne & M. Scriven *Perspectives on Curriculum Evaluation, AERA Monograph Series – Curriculum Evaluation*, Chicago: Rand McNally & Co.

Seabrook, R., Brown, G. D. A. & Solity, J. E. (2005). Distributed and Massed Practice: From Laboratory to Classroom. *Applied Cognitive Psychology*, *19*(1), 107–122. https://doi.org/10.1002/acp.1066

Serrat, O. (2017). *Knowledge Solutions Tools, Methods, and Approaches to Drive Organizational Performance.* Springer Open.

Seyda, S. & Placke, B. (2020). Weiterbildungserhebung 2020 - Weiterbildung auf Wachstumskurs.*IW-Trends, 47*(4), 3–21. Institut der deutschen Wirtschaft Köln. Availabe at: https://www.iwkoeln.de/studien/susanne-seyda-beate-placke-weiterbildung-auf-wachstumskurs-493427.html

Shavelson, R. J. & Towne, L. (2002). *Scientific research in education*. National Academy Press, Washington, DC.

Shipstead, Z., Lindsey, D. R. B., Marshall, R. L. & Engle, R. W. (2014). The mechanisms of working memory capacity: Primary memory, secondary memory, and attention control. *Journal of Memory and Language*, *72*(1), 116–141. https://doi.org/10.1016/j.jml.2014.01.004

Simon, D. A. & Bjork, R. A. (2001). Metacognition in Motor Learning. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *27*(4), 907–912. https://doi.org/10.1037/0278-7393.27.4.907

Sitte, W. (2001). Taxonomie. In Sitte, W. & Wohlschlägl, H. (Eds.), *Beiträge zur Didaktik des „Geographie und Wirtschaftskunde"-Unterrichts* (16th ed., pp. 473–475). Institut für Geographie und Regionalforschung der Universität Wien.

Skinner, B. F. (1968). *The technology of teaching*. Prentice Hall, New York.

Smith, P. L. & Ragan, T. J. (2005). *Instructional Design* (3rd ed). John Wiley & Sons, Inc., New Jersey.

Smith, S. M. & Rothkopf, E. Z. (1984). Contextual enrichment and distribution of practice in the classroom. *Cognition and Instruction, 1*(3), 341–358. https://doi.org/10.1207/s1532690xci0103_4

Smith, C. D. & Scarf, D. (2017). Spacing Repetitions Over Long Timescales: A Review and a Reconsolidation Explanation. *Frontiers in psychology*, *8*, 962. https://doi.org/10.3389/fpsyg.2017.00962

Smolen, P., Zhang, Y. & Byrne, J. H. (2016). The right time to learn: mechanisms and optimization of spaced learning. *Nature reviews. Neuroscience, 17*(2), 77–88. https://doi.org/10.1038/nrn.2015.18

Sobel, H. S., Cepeda, N. J. & Kapler, I. V. (2011). Spacing Effects in Real-World Classroom Vocabulary Learning. *Applied Cognitive Psychology*, *25*(5), 763–767. https://doi.org/10.1002/acp.1747

Son, L. K. (2010). Metacognitive Control and the Spacing Effect. *Journal of Experimental Psychology: Learning Memory and Cognition*, *36*(1), 255–262. https://doi.org/10.1037/a0017892

Sorden, S. D. (2005). A cognitive approach to instructional design for multimedia learning. *Informing Science*, *8*, 263–279. https://doi.org/10.28945/498

Spicer, D. P. & Sadler-Smith, E. (2006). Organizational learning in smaller manufacturing firms. *International Small Business Journal: Researching Entrepreneurship, 24*(2), 133–158. https://doi.org/10.1177/0266242606061836

Spielman, R., Dumper, K., Jenkins, W., Lacombe, A., Lovett, M. & Perlmutter, M. (2018). Memory. In R. E. West, *Foundations of Learning and Instructional Design Technology: The Past, Present, and Future of Learning and Instructional Design* (pp. 71–79). EdTech Books. Retrieved from: https://edtechbooks.org/lidtfoundations/memory

Spitzer, H. F. (1939). Studies in retention. *Journal of Educational Psychology, 30*(9), 641–656. https://doi.org/10.1037/h0063404

Squire, L. R. (1987). *Memory and brain*. Oxford University Press, New York.

Stanny, C. J. (2016). Reevaluating bloom's taxonomy: What measurable verbs can and cannot say about student learning. *Education Sciences*, *6*(4), 37. https://doi.org/10.3390/educsci6040037

Statistisches Bundesamt (2021). 3*26 Milliarden Euro für Bildung, Forschung und Wissenschaft*. Retrieved from: https://www.forschung-und-lehre.de/politik/326-milliarden-euro-fuer-bildung-forschung-und-wissenschaft-3615/. Accessed: October 1, 2021

Strack, R., Antebi, P., Kovács-Ondrejkovic, O., Kavanagh, K., Baier, J. & Lóepz Gobernado, A. (2021). *Decoding Global Reskilling and Career Paths*. Boston Consulting Group.

Sweller, J. (1988). Cognitive Load During Problem Solving: Effects on Learning - Sweller - 2010 - Cognitive Science - Wiley Online Library. *Cognitive Science*, *12*(2), 257–285. https://doi.org/10.1016/0364-0213(88)90023-7

Sweller, J. (2010). Element interactivity and intrinsic, extraneous, and germane cognitive load. *Educational Psychology Review, 22*(2), 123–138. https://doi.org/10.1007/s10648-010-9128-5

Sweller, J. (2011). Cognitive load theory. In J. P. Mestre & B. H. Ross (Eds.), *The psychology of learning and motivation: Cognition in education* (pp. 37–76). Elsevier Academic

Press, Cambridge Massasuchetts. https://doi.org/10.1016/B978-0-12-387691-1.00002-8

Sweller, J. (2016). Working Memory, Long-term Memory, and Instructional Design. *Journal of Applied Research in Memory and Cognition*, *5*(4), 360–367. https://doi.org/10.1016/j.jarmac.2015.12.002

Sweller, J., Ayres, P. & Kalyuga, S. (2011). *Cognitive load theory*. Springer, New York.

Sweller, J., Chandler, P., Tierney, P. & Cooper, M. (1990). Cognitive Load as a Factor in the Structuring of Technical Material. *Journal of Experimental Psychology: General*, *119*(2), 176–192. https://doi.org/10.1037/0096-3445.119.2.176

Sweller, J. & Paas, F. (2017). Should self-regulated learning be integrated with cognitive load theory? A commentary. *Learning and Instruction*, *51*, 85–89. https://doi.org/10.1016/j.learninstruc.2017.05.005

Sweller, J. & Sweller, S. (2006). Natural information processing systems. *Evolutionary Psychology, 4*, 434–458. https://doi.org/10.1177/147470490600400135

Sweller, J., van Merriënboer, J. J. G. & Paas, F. (1998). Cognitive Architecture and Instructional Design. *Educational Psychology Review*, *10*(3), 251–296. https://doi.org/10.1023/A:1022193728205

Sweller, J., van Merriënboer, J. J. G. & Paas, F. (2019). Cognitive Architecture and Instructional Design: 20 Years Later. *Educational Psychology Review*, *31*(2), 261–292. https://doi.org/10.1007/s10648-019-09465-5

Tannenbaum, S. I. (1997). Enhancing continuous learning: Diagnostic findings from multiple companes. *Human Resource Management*, *36*(4), 437–452. https://doi.org/10.1002/(sici)1099-050x(199724)36:4<437::aid-hrm7>3.0.co;2-w

Taras, M. (2005). Assessment - Summative and formative - Some theoretical reflections. *British Journal of Educational Studies*, *53*(4), 466–478. https://doi.org/10.1111/j.1467-8527.2005.00307.x

Tauber, T., Lazzareschi, B, Welna, N., Danzl, S., Johnson, D., Mott, A., Foster, G., Olmstead, A., Naughton, M. & Sprague, J. (2019). *How the workforce learns*. Retrieved from: https://get.degreed.com/hubfs/How%20Workers%20Learn%20Report_no%20date.pdf. Accessed: May 30, 2022.

The Jamovi project (2021). Jamovi. (Version 2.2.5) [Computer Software]. Retrieved from: https://www.jamovi.org.

The World Bank. (2003). *Lifelong learning in the global knowledge economy: Challenges for developing countries*. The World Bank. https://doi.org/10.1596/978- 0- 8213- 5475- 9

Thios, S. J. & D'Agostino, P. R. (1976). Effects of repetition as a function of study-phase retrieval. *Journal of Verbal Learning and Verbal Behavior, 15*(5), 529–536. https://doi.org/10.1016/0022-5371(76)90047-5

Toppino T. C. (1991). The spacing effect in young children's free recall: support for automatic-process explanations. *Memory & cognition, 19*(2), 159–167. https://doi.org/10.3758/bf03197112

Toppino, T. C. & DiGeorge, W. (1984). The spacing effect in free recall emerges with development. *Memory & Cognition*, *12*(2), 118–122. https://doi.org/10.3758/BF03198425

Toppino, T. C. & Gerbier, E. (2014). About practice: Repetition, spacing, and abstraction. In B. H. Ross (Ed.), *The psychology of learning and motivation* (pp. 113–189). Elsevier Academic Press, Cambridge Massasuchetts.

Toppino, T. C. & Gracen, T. F. (1985). The lag effect and differential organization theory:

Nine failures to replicate. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 11*(1), 185–191. https://doi.org/10.1037/0278-7393.11.1.185

Toppino, T. C., Phelan, H. A. & Gerbier, E. (2018). Level of initial training moderates the effects of distributing practice over multiple days with expanding, contracting, and uniform schedules: evidence for study-phase retrieval. *Memory & Cognition, 46*(6), 969–978. https://doi.org/10.3758/s13421-018-0815-7.

Tricot, A. & Sweller, J. (2014). Domain-Specific Knowledge and Why Teaching Generic Skills Does Not Work. *Educational Psychology Review*, *26*(2), 265–283. https://doi.org/10.1007/s10648-013-9243-1

Tuijnman, A. & Boström, A. (2002). Changing Notions of Lifelong Education and Lifelong Learning. *International Review of Education*, *48*(1/2), 93–110. https://doi.org/10.1023/A:1015601909731

Tullis, J. G., Finley, J. R. & Benjamin, A. S. (2013). Metacognition of the testing effect: Guiding learners to predict the benefits of retrieval. *Memory and Cognition*, *41*(3), 429–442. https://doi.org/10.3758/s13421-012-0274-5

Tully, T., Preat, T., Boynton, S. C. & Del Vecchio, M. (1994). Genetic dissection of consolidated memory in Drosophila. *Cell, 79*(1), 35–47. https://doi.org/10.1016/0092-8674(94)90398-0

Twelker, P. A., Urbach, F. D. & Buck, J. E. (1972). *The systematic development of instruction: An overview and basic guide to the literature*. Stanford, CA: ERIC Clearinghouse on Educational Media and Technology.

Tynjälä, P. (2008). Perspectives into learning at the workplace. *Educational Research Review*, *3*(2), 130–154. https://doi.org/10.1016/j.edurev.2007.12.001

Typeform (2021). Retrieved from: www.typeform.com. First accessed: August 12, 2021.

Udemy (2018). *State of the ROI of Learning Report*. Retrieved from: https://business.udemy.com/resources/state-of-the-roi-of-learning-report/?utm_source=cultureamp&utm_medium=partner&utm_campaign=state-of-roi-of-learning-report. Accessed: May 22, 2022

UNESCO (2015). *Rethinking education: towards a global common good?* UNESCO Publishing. Available at: https://unesdoc.unesco.org/ark:/48223/pf0000232555

University4Industry (2021). Company website. Retrieved from: www.university4industry.com. First accessed: July 1, 2021

van Merriënboer, J. J. G. (1997). *Training complex cognitive skills: A four-component instructional design model for technical training*. Educational Technology Publications, New Jersey.

van Merriënboer, J. J. G. & Sluijsmans, D. M. A. (2009). Toward a synthesis of cognitive load theory, four-component instructional design, and self-directed learning. *Educational Psychology Review*, *21*(1), 55–66. https://doi.org/10.1007/s10648-008-9092-5

van Merriënboer, J. J. G. & Sweller, J. (2005). Cognitive Load Theory and Complex Learning: Recent Developments and Future Directions. *Educational Psychology Review, 17*(2), 147–177. https://doi.org/10.1007/s10648-005-3951-0

Vargas, C. (2017). *Lifelong learning from a social justice perspective. Education Research and Foresight Working Papers Series*, No. 21. UNESCO, Paris. https://en.unesco.org/node/268820

Verkoeijen, P., Rikers, R. & Özsoy, B. (2008). Distributed Rereading can Hurt the Spacing Effect in Text Memory. *Applied Cognitive Psychology*, *22*(5), 685–695. https://doi.org/10.1002/acp.1388

Vlach, H. A., Bredemann, C. A. & Kraft, C. (2019). To mass or space? Young children do not possess adults' incorrect biases about spaced learning. *Journal of Experimental Child Psychology*, *183*, 115–133. https://doi.org/10.1016/j.jecp.2019.02.003

Vlach, H. A., Sandhofer, C. M. & Kornell, N. (2008). The spacing effect in children's memory and category induction. *Cognition, 109*(1), 163–167. https://doi.org/10.1016/j.cognition.2008.07.013

Waller, R., Hodge, S., Holford, J., Milana, M. & Webb, S. (2020). Lifelong education, social inequality and the COVID-19 health pandemic. *International Journal of Lifelong Education*, *39*(3), 243–246. https://doi.org/10.1080/02601370.2020.1790267

Walsh, M. M., Gluck, K. A., Gunzelmann, G., Jastrzembski, T. & Krusmark, M. (2018a). Evaluating the Theoretic Adequacy and Applied Potential of Computational Models of the Spacing Effect. *Cognitive Science*, *42*, 644–691. https://doi.org/10.1111/cogs.12602

Walsh, M. M., Gluck, K. A., Gunzelmann, G., Jastrzembski, T., Krusmark, M., Myung, J. I., Pitt, M. A. & Zhou, R. (2018b). Mechanisms underlying the spacing effect in learning: A comparison of three computational models. *Journal of Experimental Psychology: General*, *147*(9), 1325–1348. https://doi.org/10.1037/XGE0000416

Weinstein, Y., Madan, C. R. & Sumeracki, M. A. (2018). Teaching the science of learning. *Cognitive Research: Principles and Implications*, *3*(1), 2. https://doi.org/10.1186/s41235-017-0087-y

Wickelgren, W. A. (1972). Trace resistance and the decay of long-term memory. *Journal of Mathematical Psychology, 9*(4), 418–455. https://doi.org/10.1016/0022-2496(72)90015-6

Wiliam, D. (2011). What is assessment for learning? *Studies in Educational Evaluation*, *37*(1), 3–14. https://doi.org/10.1016/j.stueduc.2011.03.001

Wilson, B. A. & Evans, J. J. (1996). Error-free learning in the rehabilitation of people with memory impairments. *The Journal of Head Trauma Rehabilitation, 11*(2), 54–64. https://doi.org/10.1097/00001199-199604000-00006

Wise, K. (2020). Measuring memory. In Ven den Bulck, J. (Ed.), *The International Encyclopedia of Media Psychology*, pp. 1–4. John Wiley & Sons, Inc., New York. https://doi.org/10.1002/9781119011071.iemp0201

Wiseheart, M., Küpper-Tetzel, C. E. , Weston, T., Kim, A. S.N., Kapler, I. V. & Foot-Seymour, V. (2019). Enhancing the quality of student learning using distributed practice. In: Dunlosky, J. and Rawson, K. A. (eds.) *The Cambridge Handbook of Cognition and Education. Series: Cambridge handbooks in psychology* (pp. 550–584). Cambridge University Press, Cambridge, United Kingdom. https://doi.org/10.1017/9781108235631.023

Yam, S. & Rossini, P. (2013). Formative assessment in property education: A comparison between blended and online learning. *Pacific Rim Property Research Journal*, *19*(1), 101–127. https://doi.org/10.1080/14445921.2013.11104376

Yan, V. X., Bjork, E. L. & Bjork, R. A. (2016). On the difficulty of mending metacognitive illusions: A priori theories, fluency effects, and misattributions of the interleaving benefit. *Journal of Experimental Psychology: General, 145*(7), 918–933. https://doi.org/10.1037/xge0000177

Youssef-Shalala, A., Ayres, P., Schubert, C. & Sweller, J. (2014). Using a general problem-solving strategy to promote transfer. *Journal of Experimental Psychology: Applied, 20*(3), 215–231. https://doi.org/10.1037/xap0000021

Zahidi, S. (2020). *We need a global reskilling revolution – here's why*. Retrieved from: https://www.weforum.org/agenda/2020/01/reskilling-revolution-jobs-future-skills/. Accessed: October 9, 2021

Zahidi, S., Ratcheva, V., Hingel, G. & Brown, S. (2020). *The Future of Jobs Report 2020*. Retrieved from: https://www.weforum.org/reports/the-future-of-jobs-report-2020/. Accessed: July 22, 2022.

Zechmeister, E. B. & Shaughnessy, J. J. (1980). When you know that you know and when you think that you know but you don't. *Bulletin of the Psychonomic Society*, *15*(1), 41–44. https://doi.org/10.3758/BF03329756

Zuber, U. (2014). Der Arbeitsplatz der Zukunft Entwicklungspfad für eine lern- und wandlungsfähige Institution. In K. W. Schönherr & V. Tiberius, *Lebenslanges Lernen Wissen und Können als Wohlstandsfaktoren* (pp. 171–202). Springer VS, Wiesbaden.

Appendices

Appendix A – Typeform assessment for the e-learning on "platform business models"

Typeform assessment for platform business models

Question no.	Question	Correct answer
1	Was ist eine Alternative zum klassischen Pipeline-Geschäftsmodell?	Plattform Geschäftsmodell
2	Wie funktioniert das Qualitätsmanagement in einem Plattform-Geschäftsmodell?	Die Anbieter und Kunden auf der Plattform bewerten sich gegenseitig. Schlechte Anbieter/Kunden finden keine weiteren Kunden/Anbieter.
3	Erläutern Sie einen konkreten Vorteil des Plattform-Geschäftsmodells gegenüber dem Pipeline-Geschäftsmodell.	Die Plattform hat keine Gatekeeper (Anbieter und Kunden bestimmen selbst, wer auf der Plattform erfolgreich ist); ist nicht begrenzt durch Ressourcenknappheit (z.B. Finanzmittel) und kann besser Daten nutzen (z.B. in Form von Kundenbewertungen).
4	Was ist keine Plattformgeschäftsstrategie? • *Follow-the-Rabbit-Strategie* • *Piggyback-Strategie* • *Buy-and-build-Strategie* • *Micromarket-Strategie*	Buy-and-build-Strategie
5	Bitte erklären Sie das Konzept einer Plattformstrategie.	*Beispiel*: Follow-the-rabbit Strategie: Man öffnet sein bestehendes Geschäft für andere Partner (andere Strategien denkbar).
6	Bitte erläutern Sie den Unterschied zwischen zwei Plattformstrategien Ihrer Wahl.	*Beispiel*: In der follow-the-rabbit Strategie öffnet man sein bestehendes erfolgreiches Geschäftsmodell für andere Partner. In der Marquee-Strategie überzeugt man dagegen Branchenführer auf die eigene Plattform zu kommen, ohne dass sie notwendigerweise bereits erfolgreich ist (andere Antworten denkbar).
7	Wie funktioniert das Konzept der erfolgsbasierten Gebühren bei einem Plattformgeschäft?	Es wird bei jeder erfolgten Transaktion eine Gebühr erhoben (z.B. x% des Transaktionsvolumens oder eine fixe Gebühr).
8	Was ist keine Preisstrategie für ein Plattformgeschäft? • *Erfolgsbasierte Gebühren* • *Gebühren für zusätzliche Services* • *Gebühren für Produktentwicklung*	Gebühren für Produktentwicklung

Typeform assessment for platform business models (continued)

Question no.	Question	Correct answer
9	Erklären Sie den Unterschied zwischen erfolgsbasierten Gebühren und Gebühren für zusätzliche Qualitätssicherung.	Bei der erfolgsbasierten Gebühr wird nur bei abgeschlossenen Transaktionen gecharged. Beim anderen Konzept werden dagegen für die Teilnahme an der Plattform Gebühren fällig, da eine zusätzliche Qualitätskontrolle erfolgt.
10	Nennen Sie bitte eine Art von Netzwerkeffekten.	Same-side effects/cross-side effects/reinforcing economies of scale/supply-side effect/demand-side effects
11	Welche der folgenden Aussagen ist richtig (Mehrfachnennung möglich)? • *Seitenübergreifende Effekte gelten sowohl für die Nachfrage als auch für das Angebot* • *Effekte auf einer Seite gelten sowohl für die Nachfrage als auch für das Angebot* • *Seitenübergreifende Effekte gelten nur für die Nachfrage* • *Effekte auf einer Seite gelten nur für das Angebot*	Seitenübergreifende Effekte gelten sowohl für die Nachfrage als auch für das Angebot Effekte auf einer Seite gelten sowohl für die Nachfrage als auch für das Angebot
12	Wie funktionieren Skaleneffekte auf Plattformen?	Der Nutzen einer Plattform steigt mit jedem weiteren Nutzer an, da der Markt auf der Plattform wächst, die Plattform attraktiver macht und weitere Nutzer anzieht.
13	Bitte nennen Sie den Namen einer Phase und eine für diese Phase geeignete Erfolgskennzahl für Plattform-Geschäftsmodelle.	*Beispiel*: Skalierungsphase: Anzahl neuer aktiver Nutzer (andere Antworten denkbar).
14	Warum brauchen Sie unterschiedliche Leistungskennzahlen entlang der verschiedenen Lebenszyklusphasen eines Plattformgeschäfts?	Die Herausforderungen einer Plattform sind in jeder Phase unterschiedlich. Daher sieht auch Erfolg in jeder Phase anders aus, sodass es unterschiedlicher KPIs bedarf.
15	Nennen Sie die 4 Phasen im Lebenszyklus eines Plattformunternehmens. • *Einführung* • *Start* • *Entwicklung* • *Skalierung* • *Ausweitung* • *Reifegrad* • *Wachstum* • *Rückgang* • *Weiterentwicklung*	Einführung Ausweitung Reifegrad Weiterentwicklung

Typeform assessment for platform business models (continued)

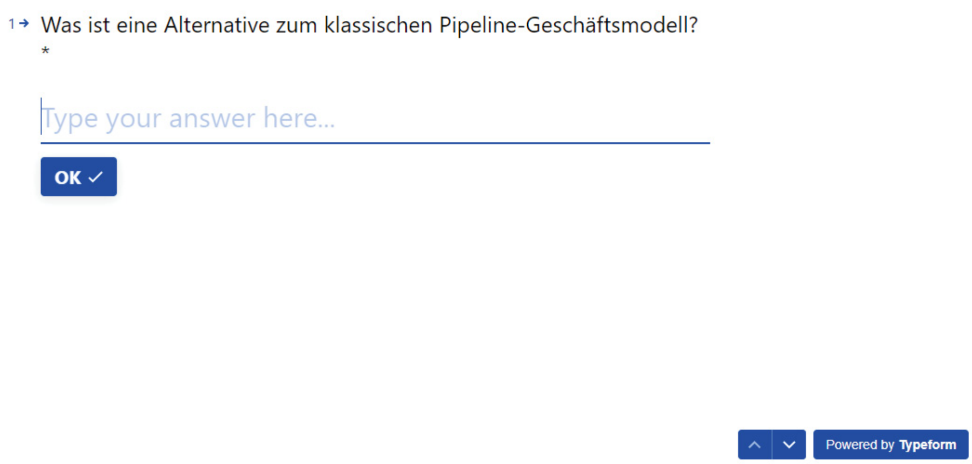
Question no.	Question
16	Was schätzen Sie, wie viele der zuletzt gestellten 15 Testfragen haben Sie richtig beantwortet?
17	Haben Sie sich zwischen dem letzten Training und diesem Test nochmals mit dem Lernmaterial befasst?
18	Glauben Sie, Ihr spezifischer Trainingsablauf hatte einen positiven Effekt auf Ihren persönlichen Lernerfolg?
19	In Ihrem jetzigen beruflichen Umfeld, wie würden Sie Ihre Trainings beschreiben? • *Mehrere kürzere Wiederholungseinheiten* • *Eine längere Lerneinheit* • *Interaktiv* • *Nicht-interaktiv* • *Selbstgesteuert* • *Geführt*
20	Welche der folgenden Arten von Training wären für Ihr persönliches berufliches Umfeld am besten geeignet? • *Mehrere kürzere Wiederholungseinheiten* • *Eine längere Lerneinheit* • *Interaktiv* • *Nicht-interaktiv* • *Selbstgesteuert* • *Geführt*
21	Glauben Sie, es wäre effektiver für Ihren persönlichen Lernerfolg gewesen, wenn Sie nur eine Lerneinheit gehabt hätten?
22	Glauben Sie, dass sich Ihr Verständnis zu den gelehrten Konzepten zu Plattform Geschäftsmodellen signifikant, in geringem Maße oder gar nicht verändert hat?

Typeform assessment for platform business models (continued)

Question no.	Question
16	Was schätzen Sie, wie viele der zuletzt gestellten 15 Testfragen haben Sie richtig beantwortet?
17	Haben Sie sich zwischen dem letzten Training und diesem Test nochmals mit dem Lernmaterial befasst?
18	Glauben Sie, Ihr spezifischer Trainingsablauf hatte einen positiven Effekt auf Ihren persönlichen Lernerfolg?
19	In Ihrem jetzigen beruflichen Umfeld, wie würden Sie Ihre Trainings beschreiben? • *Mehrere kürzere Wiederholungseinheiten* • *Eine längere Lerneinheit* • *Interaktiv* • *Nicht-interaktiv* • *Selbstgesteuert* • *Geführt*
20	Welche der folgenden Arten von Training wären für Ihr persönliches berufliches Umfeld am besten geeignet? • *Mehrere kürzere Wiederholungseinheiten* • *Eine längere Lerneinheit* • *Interaktiv* • *Nicht-interaktiv* • *Selbstgesteuert* • *Geführt*
21	Glauben Sie, es wäre effektiver für Ihren persönlichen Lernerfolg gewesen, wenn Sie nur eine Lerneinheit gehabt hätten?
22	Glauben Sie, dass sich Ihr Verständnis zu den gelehrten Konzepten zu Plattform Geschäftsmodellen signifikant, in geringem Maße oder gar nicht verändert hat?
23	Welchem Lernformat haben Sie am meisten Aufmerksamkeit geschenkt? • *Folien* • *Videos* • *Live Session*
24	Wie bewerten Sie in Summe die Effektivität des Trainingsprogrammes? Auf einer Skala von 1-5, würden Sie zustimmen, dass mehrere Wiederholungseinheiten einen größeren Lerneffekt haben als eine einzige längere?
25	Wie bewerten Sie in Summe die Effektivität des Trainingsprogrammes? Auf einer Skala von 1-5, würden Sie zustimmen, dass eine einzige längere Lerneinheit Sie sicherer macht bei der tatsächlichen Anwendung der Lerninhalte als mehrere kürzere Wiederholungseinheiten?
26	Wie bewerten Sie in Summe die Effektivität des Trainingsprogrammes? Auf einer Skala von 1-5, würden Sie zustimmen, dass eine selbst-gesteuerte Lerneinheit effektiver für den Lernerfolg ist als eine geführte Lerneinheit (Live Session)?

Typeform assessment for platform business models (continued)	
Question no.	**Question**
27	Wie bewerten Sie in Summe die Effektivität des Trainingsprogrammes? Auf einer Skala von 1-5, würden Sie zustimmen, dass eine nicht-interaktive Lerneinheit effektiver für den Lernerfolg ist als eine interaktive Lerneinheit?
28	Was mochten Sie besonders an Ihrer Lernerfahrung?
29	Was mochten Sie am Wenigsten? Was hätten wir besser machen können?
30	Gibt es noch weitere Themen, die Sie uns mitteilen möchten?

Preview mode on Typeform



Appendix B – ClassMarker multiple choice test for experimental Group 4 for the e-learning on “platform business models”

Classmarker multiple choice test for Group 4 on platform business models

Question no.	Question	Correct answer
1	Was ist ein wesentlicher Unterschied bei einem Plattform Geschäftsmodell im Vergleich zu einem Pipeline Geschäftsmodell? • *Auf einer Plattform ermöglicht Technologie die Integration mehrerer Unternehmen und Kunden aus den unterschiedlichen Phasen der Wertschöpfungskette.* • *Auf einer Plattform gibt es in der Regel eine größere Anzahl von homogenen Kunden als in einem Pipeline Geschäftsmodell.* • *Auf einer Plattform gibt es in der Regel eine größere Anzahl von homogenen Unternehmen als in einem Pipeline Geschäftsmodell.* • *Auf einer Plattform wird Technologie dazu verwendet, um eine höhere Kundenzufriedenheit zu erreichen verglichen zu einem klassischen Pipeline Geschäftsmodell.*	Auf einer Plattform ermöglicht Technologie die Integration mehrerer Unternehmen und Kunden aus den unterschiedlichen Phasen der Wert-schöpfungskette.
2	Was ist kein typischer Vorteil eines Plattform Geschäftsmodells? • *Plattformen können die Arbeitsweise von Unternehmen signifikant verbessern, da es sich voll auf die eigene Strategie und Wertschöpfung fokussieren kann.* • *Plattformen können ihr Geschäft effizienter ausweiten, weil sie die ‚Gatekeeper‘-Funktion ausschalten.* • *Plattformen könne neue Quellen für potenzielle Angebote an Kunden erschließen.* • *Plattformen können ihr Geschäft besser optimieren, da ihnen oft signifikant mehr Daten zur Verfügung stehen, die sie nutzen können.*	Plattformen können die Arbeitsweise von Unternehmen signifikant verbessern, da es sich voll auf die eigene Strategie und Wertschöpfung fokussieren kann.
3	Was versteht man unter einer ‚Follow-the-rabbit‘ Strategie? • *Öffnung des eigenen bestehenden Pipeline-Geschäfts für Partner.* • *Verwendung von Content anderer Unternehmen oder Plattformen für die eigene Plattform.* • *Gezielte Überzeugungsarbeit beim wichtigsten Akteur in der eigenen Industrie, damit er Ihre Plattformen nutzt und andere zum Nachmachen animiert.* • *Start mit einem Service für die Angebots- oder Nachfrageseite Ihrer künftigen Plattform und Öffnung der Plattform, sobald auf der jeweiligen Seite genügend Zugkraft und Dynamik vorhanden ist.*	Öffnung des eigenen bestehenden Pipeline-Geschäfts für Partner.

Classmarker multiple choice test for Group 4 on platform business models (continued)

Question no.	Question	
4	Bei welcher Pricing Strategie trifft die folgende Überlegung zu: Die Eintrittshürden für Nutzer sind niedrig, da sie nur im Falle einer Transaktion zahlen müssen? • *Erfolgsbasierte Gebühren* • *Gebühren für zusätzliche Services* • *Gebühren für Qualitätssicherung* • *Gebühren für Anmeldung*	Erfolgsbasierte Gebühren
5	Welche Art von Netzwerkeffekten beschreibt das folgende Beispiel: Wenn AirBnB mehr Nutzer hat, steigt das Angebot (d.h. die Anzahl von Apartments) ebenfalls, da die Eigentümer von mehr Buchungen ausgehen. Das gestiegene Angebot zieht weitere Nutzer an, die eine höhere Chance haben, die für sie passende Unterkunft zu finden. Dies führt erneut zu einem größeren Angebot. Das größere Angebot lockt wiederum mehr Nutzer an. • *Sich gegenseitig verstärkende Skaleneffekte* • *Skaleneffekte auf der Nachfrageseite* • *Skaleneffekte auf der Angebotsseite* • *Economies of scale*	Sich gegenseitig verstärkende Skaleneffekte

Preview mode on ClassMarker

H0560B1_Plattform Geschäftsmodelle

Alle Fragen

Frage 1 von 5

Was versteht man unter einer "Follow-the-rabbit" Strategie?

- **A.** Start mit einem Service für die Angebots- oder Nachfrageseite Ihrer künftigen Plattform und Öffnung der Plattform, sobald auf der jeweiligen Sei-te genügend Zugkraft und Dynamik vorhanden ist.
- **B.** Öffnung des eigenen bestehenden Pipeline-Geschäfts für Partner.
- **C.** Gezielte Überzeugungsarbeit beim wichtigsten Akteur in der eigenen Industrie, damit er Ihre Plattformen nutzt und andere zum Nachmachen animiert.
- **D.** Verwendung von Content anderer Unternehmen oder Plattformen für die eigene Plattform.

Weiter

Appendix C – Design plan experiment 1

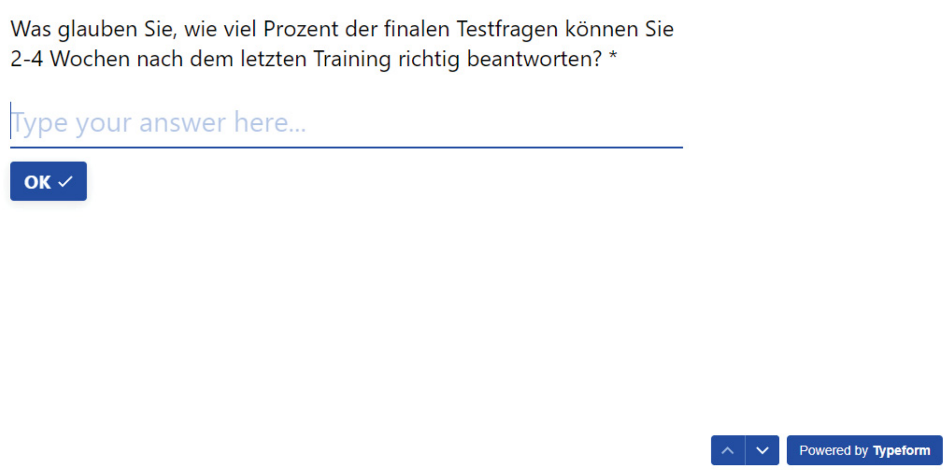
Training set-up and schedule – platform business models
(V: video, S: slides, L: live session)

Group	**Session 1** Date Content	**Session 2** Date Content	**Session 3** Date Content	**Session 4** Date Content	**Session 5** Date Content	**Session 6** Date Content
Group 1	27.09. V1-5 S1-5 L1-5	11.10. Final test (*2 weeks RI group*)	25.10. Final test (*4 weeks RI groups*)	-	-	-
Group 2	27.09. V1-5	01.10. S1-5	06.10. L1-5	20.10. Final test	-	-
Group 3	27.09. S1-5	01.10. L1-5	06.10. V1-5	20.10. Final test	-	-
Group 4	27.09. V1-5 Test	01.10. S1-5 Test	06.10. L1-5 Test	20.10. Final test	-	-
Group 5	27.09. V1-2	01.10. V3 S1	06.10. S2-3 L1-3	11.10. V4-5	15.10. S4-5 L4-5	29.10. Final test
Group 6	27.09. V1-5	04.10. S1-5	11.10. L1-5	08.11. Final test	-	-
Group 7	27.09. V1-2	04.10. V3 S1	11.10. S2-3 L1-3	18.10. V4-5	25.10. S4-5 L4-5	22.11. Final test

Appendix D – Typeform survey on metacognition for the e-learning on "platform business models"

Question no.	Question
1	Was glauben Sie, wie viel Prozent der finalen Testfragen können Sie 2-4 Wochen nach dem letzten Training richtig beantworten?
2	Welche Lernerfahrung bevorzugen Sie? • *Eine einzige längere* • *Mehrere kürzere*
3	Wenn Sie möchten, können Sie uns Ihre Präferenz erläutern.
4	Wie zufrieden waren Sie mit der heutigen Lernerfahrung? • *Sehr unzufrieden* • *Unzufrieden* • *Weder zufrieden, noch unzufrieden* • *Zufrieden* • *Sehr zufrieden*
5	Wenn Sie möchten, erläutern Sie uns bitte, was zu Ihrer (Un)Zufriedenheit beitrug.

Preview mode on Typeform



Appendix E – Typeform assessment for the e-learning on “time management”

Typeform assessment for time management

Question no.	Question	Correct answer
1	Welche drei Zeittypen bzw. Zeitpersönlichkeiten werden typischerweise differenziert? Bitte erläutern Sie die Charakteristik von einer davon.	Jongleure, Improvisierer, Feuerwehrleute
2	Für welchen Zeittyp ist die Pomodoro-Technik geeignet?	Feuerwehr
3	Warum hilft die Eisenhower-Matrix beim Zeitmanagement?	Fokussierung auf die wichtigen und dringenden Dinge
4	Wie oft sollte man die Pomodoro Technik in einer Sitzung wieder holen?	4 mal wiederholen (3 oder 5 zählt auch als Antwort)
5	Was muss man tun, wenn man die Anker-Methode richtig anwendet?	Zeit einer vergleichbaren Aufgabe in der Vergangenheit nehmen und extrapolieren
6	Wozu dient die Anker-Methode?	Bessere Abschätzung von Zeiten, dadurch bessere Planung
7	Warum funktioniert die Pomodoro-Technik?	Konzentration und Fokus auf nur 1 Aufgabe; Pausen helfen
8	Welche Methode sollten Sie zur besseren Priorisierung verwenden?	Eisenhower-Matrix
9	Was sind gute Ansatzpunkte zur Verbesserung der Planung? Die Auswahl von drei Antworten ist genügend: • *Start- und Endzeiten notieren* • *Pufferzeiten aus Planung entfernen* • *Dauer von Aufgaben abschätzen* • *Ähnliche Aktivitäten über den Tag verteilen/balancieren* • *Schwierigkeiten der Aufgaben an Biorhythmus anpassen* • *Aufgaben mit Stoppuhr messen*	Start- und Endzeiten notieren Dauer von Aufgaben abschätzen Aufgaben mit Stoppuhr messen

Typeform assessment for time management (continued)

Question no.	Question	Correct answer
10	Nennen Sie zwei der besprochenen Zeitfallen: • *Technologie* • *Geld und Wohlstand* • *Unterschätzte Zeit* • *Beschäftigtsein als Status* • *Ablehnung von Leerlauf* • *Planungsfehlentscheidung*	*Technologie* *Geld und Wohlstand* *Unterschätzte Zeit* *Beschäftigtsein als Status* *Ablehnung von Leerlauf* *Planungsfehlentscheidung*
11	Nach welchen Dimensionen unterscheiden Sie Aufgaben in der Eisenhower-Matrix? • *Nutzen* • *Mehrwert* • *Dringlichkeit* • *Zeitaufwand* • *Wichtigkeit* • *Delegation* • *Zeitmanagement* • *Priorisierung*	Dringlichkeit Wichtigkeit
12	Was sind die drei ‚P' im Zeitmanagement? • *Präsentation* • *Planung* • *Priorisierung* • *Produktion* • *Planungsfehlentscheidung* • *Profit* • *Produktivität* • *Prämisse* • *Portfolio* • *Pomodoro* • *Power* • *Psychologie*	Planung Priorisierung Produktivität
13	Was schätzen Sie, wie viele der zuletzt gestellten 12 Testfragen haben Sie richtig beantwortet?	
14	Haben Sie sich zwischen dem letzten Training und diesem Test nochmals mit dem Lernmaterial befasst?	
15	Glauben Sie, Ihr spezifischer Trainingsablauf hatte einen positiven Effekt auf Ihren persönlichen Lernerfolg?	

Typeform assessment for time management (continued)

Question no.	Question
16	In Ihrem jetzigen beruflichen Umfeld, wie würden Sie Ihre Trainings beschreiben? • *Mehrere kürzere Wiederholungseinheiten* • *Eine längere Lerneinheit* • *Interaktiv* • *Nicht-interaktiv* • *Selbstgesteuert* • *Geführt*
17	Welche der folgenden Arten von Training wären für Ihr persönliches berufliches Umfeld am besten geeignet? • *Mehrere kürzere Wiederholungseinheiten* • *Eine längere Lerneinheit* • *Interaktiv* • *Nicht-interaktiv* • *Selbstgesteuert* • *Geführt*
18	Glauben Sie, es wäre effektiver für Ihren persönlichen Lernerfolg gewesen, wenn Sie nur eine Lerneinheit gehabt hätten?
19	Glauben Sie, dass sich Ihr Verständnis zu den gelehrten Konzepten zum Zeitmanagement signifikant, in geringem Maße oder gar nicht verändert hat?
20	Welchem Lernformat haben Sie am meisten Aufmerksamkeit geschenkt? • *Folien* • *Videos* • *Live Session*
21	Wie bewerten Sie in Summe die Effektivität des Trainingsprogrammes? Auf einer Skala von 1-5, würden Sie zustimmen, dass mehrere Wiederholungseinheiten einen größeren Lerneffekt haben als eine einzige längere?
22	Wie bewerten Sie in Summe die Effektivität des Trainingsprogrammes? Auf einer Skala von 1-5, würden Sie zustimmen, dass eine einzige längere Lerneinheit Sie sicherer macht bei der tatsächlichen Anwendung der Lerninhalte als mehrere kürzere Wiederholungseinheiten?
23	Wie bewerten Sie in Summe die Effektivität des Trainingsprogrammes? Auf einer Skala von 1-5, würden Sie zustimmen, dass eine selbst-gesteuerte Lerneinheit effektiver für den Lernerfolg ist als eine geführte Lerneinheit (Live Session)?
24	Wie bewerten Sie in Summe die Effektivität des Trainingsprogrammes? Auf einer Skala von 1-5, würden Sie zustimmen, dass eine nicht-interaktive Lerneinheit effektiver für den Lernerfolg ist als eine interaktive Lerneinheit?
25	Was mochten Sie besonders an Ihrer Lernerfahrung?
26	Was mochten Sie am Wenigsten? Was hätten wir besser machen können?
27	Gibt es noch weitere Themen, die Sie uns mitteilen möchten?

Preview mode on Typeform

1→ Welche drei Zeittypen bzw. Zeitpersönlichkeiten werden typischerweise differenziert? Bitte erläutern Sie die Charakteristik von einer davon. *

Type your answer here...

Shift ⇧ + Enter ↵ to make a line break

OK ✓

Powered by Typeform

Appendix F – ClassMarker multiple choice test for experimental Group 4 for the e-learning on “time management”

Classmarker multiple choice test for Group 4 on time management

Question no.	Question	Correct answer
1	Nennen Sie bitte zwei Aspekte, die typischerweise auf Improvisierer:innen zutreffen. • *Wollen die Dinge richtig und mit hoher Qualität erledigen.* • *Sagen lieber ‚Ja‘ als ‚Nein‘* • *Können spontan Dinge angehen und umsetzen* • *Arbeiten sehr gut unter hohem Druck und in Krisensituationen* • *Finden schnell kreative Lösungen*	Können spontan Dinge angehen und umsetzen Finden schnell kreative Lösungen
2	Welche der folgenden Aussagen ist richtig: • *Wenige Aufgaben gleichzeitig zu managen benötigt viel Gehirnkapazität.* • *Viele Aufgaben gleichzeitig zu managen benötigt viel Gehirnkapazität.* • *Wenige Aufgaben gleichzeitig zu managen benötigt wenig Gehirnkapazität.* • *Viele Aufgaben gleichzeitig zu managen benötigt wenig Gehirnkapazität.*	Viele Aufgaben gleichzeitig zu managen benötigt viel Gehirnkapazität.
3	Geben Sie zwei Antworten an, warum eine gute Planung hilfreich ist: • *Zeiten müssen oft geschätzt werden* • *Geringes Risiko in Stress zu geraten* • *Erleichtert die Planung* • *Aufgaben in der Vergangenheit müssen identifiziert und angepasst werden, um den Biorhythmus zu ändern* • *Um wichtige von dringenden Themen zu unterscheiden* • *Durch das Abhaken von Aufgaben wird das Belohnungssystem im Gehirn aktiviert und Dopamin ausgeschüttet*	Zeiten müssen oft geschätzt werden Geringes Risiko in Stress zu geraten
4	Welche Aufgabe der folgenden eignet sich am besten für die Pomodoro-Methode: • *Buch lesen* • *Seminararbeit schreiben* • *Mittagsschlaf* • *Exceltabellen formatieren* • *Vorlesungsstunde im Internet ansehen*	Exceltabellen formatieren

Preview mode on ClassMarker

H0616B1_Zeitmanagement

See all questions

Question 1 of 4

Welche Aufgabe der folgenden eignet sich am besten für die Pomodoro-Methode:

- **A.** Seminararbeit schreiben
- **B.** Vorlesungsstunde im Internet ansehen
- **C.** Buch lesen
- **D.** Mittagsschlaf
- **E.** Exceltabellen formatieren

Next

Appendix G – Design plan experiment 2

Training set-up and schedule – time management
(V: video, S: slides, L: live session)

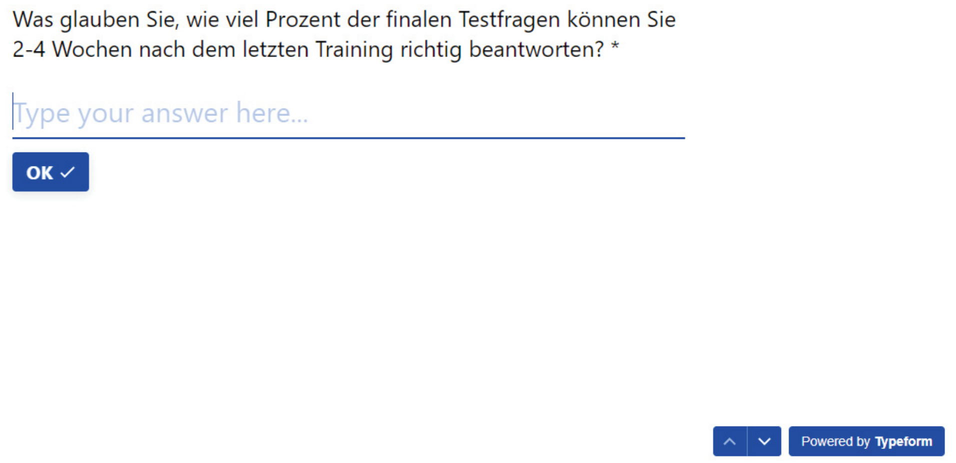
Group	**Session 1** Date Content	**Session 2** Date Content	**Session 3** Date Content	**Session 4** Date Content	**Session 5** Date Content	
Group 1	02.05. V1-4 S1-4 L1-4	16.05. Final test (*2 weeks RI group*)	30.05. Final test (*4 weeks RI groups*)	-	-	-
Group 2	02.05. V1-4	06.05. S1-4	11.05. L1-4	25.05. Final test	-	-
Group 3	02.05. S1-4	06.05. L1-4	11.05. V1-4	25.05. Final test	-	-
Group 4	02.05.. V1-4 Test	06.05. S1-4 Test	11.05. L1-4 Test	25.05. Final test	-	-
Group 5	02.05. V1-2	06.05. V3 S1	11.05. V4 S2 L1-2	16.05. S3-4	20.05. L3-4	03.06. Final test
Group 6	02.05. V1-4	09.05. S1-4	16.05. L1-4	13.06. Final test	-	-
Group 7	02.05. V1-2	09.05. V3 S1	16.05. V4 S2 L1-2	23.05. S3-4	30.05. L3-4	27.06. Final test

Appendix H – Typeform survey on metacognition for the e-learning on "time management"

Typeform survey on metacognition for time management

Question no.	Question
1	Was glauben Sie, wie viel Prozent der finalen Testfragen können Sie 2-4 Wochen nach dem letzten Training richtig beantworten?
2	Welche Lernerfahrung bevorzugen Sie? • *Eine einzige längere* • *Mehrere kürzere*
3	Wenn Sie möchten, können Sie uns Ihre Präferenz erläutern.
4	Wie zufrieden waren Sie mit der heutigen Lernerfahrung? • *Sehr unzufrieden* • *Unzufrieden* • *Weder zufrieden, noch unzufrieden* • *Zufrieden* • *Sehr zufrieden*
5	Wenn Sie möchten, erläutern Sie uns bitte, was zu Ihrer (Un)Zufriedenheit beitrug.
6	Wie motiviert waren Sie für die heutige Lerneinheit? • *Sehr motiviert* • *Motiviert* • *Weder motiviert, noch unmotiviert* • *Unmotiviert* • *Sehr unmotiviert*

Preview mode on Typeform



Appendix I – Data protection policies of Heinrich Heine University, Düsseldorf

HEINRICH HEINE
UNIVERSITÄT DÜSSELDORF

Forschungsdaten-Richtlinie der Heinrich-Heine-Universität Düsseldorf

I. Präambel

Nach § 3 HG NRW ist Aufgabe der Heinrich-Heine-Universität Düsseldorf die Gewinnung wissenschaftlicher Erkenntnisse sowie die Pflege und Entwicklung der Wissenschaften im Wege der Forschung. Die Verfügbarkeit von Forschungsdaten ist ein Aspekt guter wissenschaftlicher Praxis, der in der Ordnung über die Grundsätze zur Sicherung guter wissenschaftlicher Praxis an der Heinrich-Heine-Universität Düsseldorf vom 19. Februar 2014 bereits festgehalten ist. Die Planung, Erhebung, Verarbeitung, Aufbewahrung und nachhaltige Bereitstellung von Forschungsdaten muss anerkannten Standards und hohen Anforderungen genügen, damit Forschungsergebnisse nachvollziehbar und überprüfbar sind und die Daten langfristig genutzt werden können. Daher und basierend auf den Empfehlungen der 16. Hochschulrektorenkonferenz vom 13. Mai 2014 soll mit dieser Richtlinie des Rektorats der Heinrich-Heine-Universität Düsseldorf ein zentraler Rahmen für die Arbeit mit Forschungsdaten geschaffen werden, dessen konkrete fachspezifische Ausgestaltung in Eigenverantwortung der unterschiedlichen Wissenschaftsbereiche geleistet werden muss.

II. Anwendungsbereich

1. **Forschungsdaten** sind alle Daten, die im Zuge von Forschungsprozessen gesammelt, beobachtet, simuliert, abgeleitet oder generiert werden. Dies gilt unabhängig von der Fachdisziplin, dem Format oder der angewandten Methode. Erfasst sind insbesondere Primärdaten, Sekundäranalysen, Visualisierungen, Modelle, Analysewerkzeuge, Objektsammlungen oder Produkte, die während des wissenschaftlichen Arbeitsprozesses erzeugt und benutzt werden.

2. **Forschungsprimärdaten** sind Daten, die im Verlauf von Quellenforschungen, Experimenten, Messungen, Erhebungen oder Umfragen entstanden sind. Sie stellen die Grundlagen für die wissenschaftlichen Publikationen dar.

3. **Forschungsdatenmanagement** umfasst

   a. Planung und Erfassung

   b. Verarbeitung und Speicherung

   c. Aufbewahrung, Zugriff und Nutzung

4. **Vorgaben von Drittmittelgebern** sind vorrangig gegenüber dieser Richtlinie zu berücksichtigen.

III. Grundsätze

1. Verantwortlichkeit

a. Die eigenverantwortlich forschenden Mitglieder der Heinrich-Heine-Universität sind für das Forschungsdatenmanagement innerhalb ihrer Vorhaben verantwortlich. Die Verantwortlichkeit beginnt mit der Erzeugung der Daten und endet mit ihrer endgültigen Löschung.

b. Im Rahmen von Forschungskollaborationen gelten diese Grundsätze, soweit die anderen Beteiligten keine gleichwertigen oder strengeren Vorgaben treffen.

2. Einhaltung von rechtlichen Rahmenbedingungen und Standards

Im Rahmen des Forschungsdatenmanagements sind gesetzliche Vorgaben, anerkannte Standards guter wissenschaftlicher Praxis sowie etwaige fachbezogene Grundsätze einzuhalten. Dabei sind insbesondere das Datenschutz- und Urheberrecht, der Geheimnisschutz und bei der Drittmittelforschung vertragliche Vorgaben zu beachten. Persönliche Daten von durch die Datenerhebung betroffenen Personen sind, soweit nach den einschlägigen Forschungsstandards möglich, zu anonymisieren, hilfsweise zu pseudonymisieren.

3. Aufstellung eines Datenmanagementplans; fachspezifische Richtlinien

a. Die Verantwortlichen sollen für Forschungsprojekte mit Forschungsdaten einen **Datenmanagementplan** aufstellen. Dieser muss insbesondere Vorgaben für die Authentizität, Integrität, Vollständigkeit, Vertraulichkeit und Veröffentlichung von Daten unter Berücksichtigung der fachspezifischen Besonderheiten enthalten. Es ist dabei festzulegen, welche Daten wie lange aufbewahrt werden müssen.

b. Die Fächer und Fakultäten können **fachspezifische Richtlinien** für typische datenintensive Forschungsvorhaben erstellen.

c. Die Heinrich-Heine-Universität Düsseldorf unterstützt die Verantwortlichen durch geeignete Informationen bei der Erstellung der Datenmanagementpläne.

4. Pflicht zur Dokumentation und Datenaufbewahrung

a. Die für ein Forschungsprojekt Verantwortlichen dokumentieren den gesamten Forschungszyklus sowie die verwendeten Werkzeuge und Verfahren.

b. Die für ein Forschungsprojekt Verantwortlichen stellen sicher, dass die einer Veröffentlichung zugrundeliegenden Forschungsprimärdaten aufbewahrt werden und im Zweifelsfall zugreifbar sind.

5. Quellenangabe, Inhaberschaft und Berechtigung

a. Daten sind persönlich zu kennzeichnen und unter dem Namen der Verantwortlichen abzulegen.

b. Etwaige Urheber- oder Leistungsschutzrechte an Daten, insbesondere das Datenbankrecht (§ 87a UrhG) verbleiben im Zweifel bei den Verantwortlichen. Dies umfasst insbesondere das Recht, die Daten weitergehend zu nutzen oder zu veröffentlichen. Für Daten, die Grundlage von schutzfähigem, geistigem Eigentum sind, gilt grundsätzlich die Verpflichtung zur Einreichung einer Erfindungsmeldung gemäß §§ 5, 42 Nr. 2 Arbeitnehmererfindungsgesetz.

c. Unberührt bleiben abweichende vertragliche Vereinbarungen, insbesondere im Rahmen von Drittmittelprojekten.

6. Vorgaben für die Speicherung

a. Die Speicherung der Forschungsdaten erfolgt im Rahmen anerkannter, hilfsweise im Datenmanagementplan definierter Standards.

b. Die Speicherung und Archivierung digitaler Forschungsdaten erfolgt in einem vom Zentrum für Informations- und Medientechnologie der Heinrich-Heine-Universität Düsseldorf bereitgestellten System oder in anerkannten externen oder internen Fachrepositorien. Soweit Daten in externen Repositorien gespeichert werden, soll dies beim Zentrum für Informations- und Medientechnologie angezeigt werden.

c. Die Daten werden durch geeignete und möglichst im Datenmanagementplan spezifizierte Metadaten beschrieben und durch einen Zeitstempel sowie eine qualifizierte elektronische Signatur vor Veränderungen geschützt.

7. Aufbewahrungsdauer, Archivierung

a. Forschungsdaten, die die Grundlage einer Publikation bilden, sollen langfristig in einem geeigneten vertrauenswürdigen Datenarchiv bzw. Repositorium archiviert und/oder veröffentlicht werden. Sie zählen zur wissenschaftlichen Leistung der Forschenden.

b. Forschungsprimärdaten sind entsprechend der „Vorschläge zur Sicherung guter wissenschaftlicher Praxis" der DFG von 1998 i.d.F. von 2013 auf haltbaren und gesicherten Datenträgern zehn Jahre nach Abschluss des Vorhabens zu sichern. Weitergehende Aufbewahrungspflichten aufgrund gesetzlicher Bestimmungen sowie Schutzmaßnahmen (z.B. Zugriffskontrollen bei personenbezogenen Daten, Identifikation von Quellen durch digitale Wasserzeichen zur Diebstahls- oder Plagiatprävention) bleiben unberührt.

8. Zugriff und Verbreitung

a. Die Verantwortlichen bestimmen, zu welchem Zeitpunkt und zu welchen rechtlichen Bedingungen Forschungsdaten zugänglich gemacht werden.

b. Die Heinrich-Heine-Universität Düsseldorf empfiehlt, Forschungsdaten ebenso wie die wissenschaftliche Publikation gemäß der Open-Access-Resolution der Heinrich-Heine Universität Düsseldorf öffentlich zugänglich zu machen, soweit keine entgegenstehenden rechtlichen Verpflichtungen bestehen (z.B. Verträge mit Verlagen, Datenschutz).

IV. Finanzierung

1. Die Heinrich-Heine Universität Düsseldorf stellt einen zentralen Speicherdienst für Forschungsdaten im Zentrum für Informations- und Medientechnologie zur Verfügung. Bei besonderen Anforderungen ist eine vorherige Einzelfallregelung zu treffen.

2. Eine Datenspeicherung bei externen Anbietern oder aufgrund von Vorgaben durch Drittmittelgeber bleibt unberührt.

V. Forschungsdatenmanagement als Teil der guten wissenschaftlichen Praxis

Zur nachhaltigen Verankerung und Entwicklung hochwertigen Forschungsdatenmanagements müssen die Prinzipien guter wissenschaftlicher Datenverarbeitung im Rahmen der Unterweisungen in guter wissenschaftlicher Praxis thematisiert werden.

VI. Überprüfung, Aktualisierung

1. Diese Forschungsdatenrichtlinie wird laufend auf ihre Vereinbarkeit mit den jeweiligen wissenschaftlichen Standards und der Praxis überprüft. Sie ist spätestens drei Jahre nach Inkrafttreten an die geltenden Maßstäbe anzupassen.

2. Für die Einhaltung und Anpassung dieser Ordnung ist das Rektorat zuständig.

VII. Inkrafttreten

Diese Richtlinie tritt am 26.11.2015 in Kraft.

Ausgefertigt aufgrund des Rektoratsbeschlusses vom 26.11.2015.

Die Rektorin der

Heinrich-Heine-Universität Düsseldorf

Steinbeck

Prof. Dr. Anja Steinbeck

Appendix J – Additional analysis on learner's learning success (H5, experiment 1)

All calculations shown in this section were adapted from Jamovi, 2022.

Post hoc comparisons – Instructional method

Comparison						
Instructional method	Instructional method	Mean difference	SE	df	t	p_{tukey}
Massed	vs. Spaced	1.37	0.441	239	3.12	0.002

Note: Comparisons are based on estimated marginal means

Post hoc comparisons – RI

Comparison						
RI	RI	Mean difference	SE	df	t	p_{tukey}
2 w	vs. 4 w	-2.72	0.441	239	-6.18	< 0.001

Note: Comparisons are based on estimated marginal means

Post hoc comparisons – Instructional method * RI

Comparison								
Instruc-tional method	RI	Instruc-tional method	RI	Mean difference	SE	df	t	p_{tukey}
Massed	2 w	vs. Massed	4 w	-5.190	0.733	239	-7.083	< 0.001
Massed	2 w	vs. Spaced	2 w	-1.094	0.577	239	-1.896	0.233
Massed	2 w	vs. Spaced	4 w	-1.350	0.660	239	-2.046	0.174
Massed	4 w	vs. Spaced	2 w	4.096	0.584	239	7.011	< 0.001
Massed	4 w	vs. Spaced	4 w	3.840	0.666	239	5.765	< 0.001
Spaced	2 w	vs. Spaced	4 w	-0.256	0.490	239	-0.523	0.953

Note: Comparisons are based on estimated marginal means

Repeated measures ANOVA learner's self-perception (within subjects effects)

	Sum of squares	df	Mean square	F	*p*
Self-perception	22.5	1	22.50	5.16	0.024
Self-perception * Instructional method	42.3	1	42.34	9.71	0.002
Self-perception * RI	166.6	1	166.59	38.19	< 0.001
Self-perception * Instructional method * RI	136.7	1	136.72	31.34	< 0.001

Note: Type 3 sums of squares

Repeated measures ANOVA learner's self-perception (between subjects effects)

	Sum of squares	df	Mean square	F	*p*
Instructional method	461.1	1	461.1	29.486	< 0.001
RI	138.0	1	138.0	8.822	0.003
Instructional method * RI	12.4	1	12.4	0.796	0.373

Note: Type 3 sums of squares

Post hoc comparisons – Learning success * Instructional method * RI

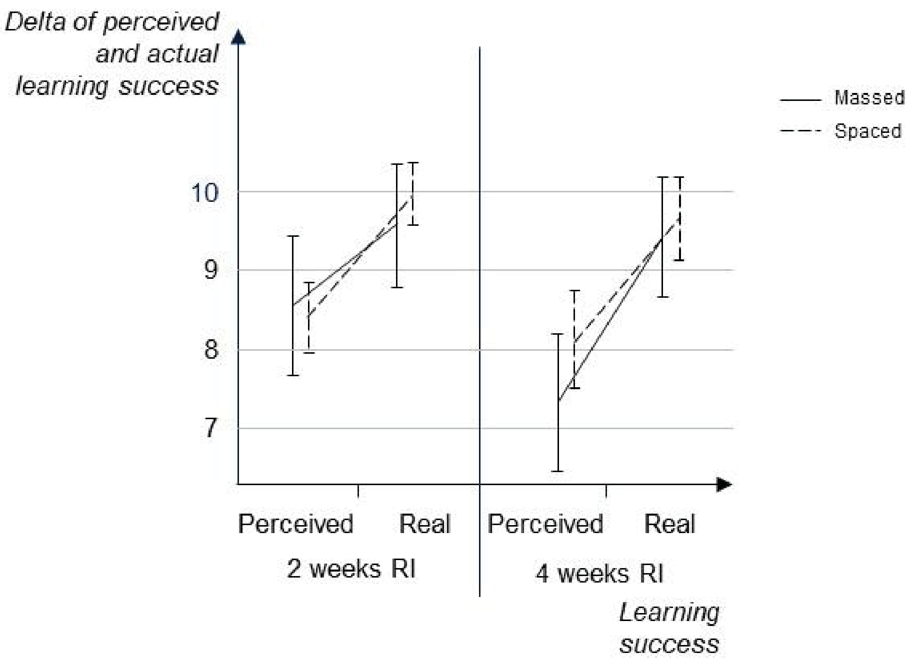
Comparison											
Lear-ning success	Instruc-tional method	RI		Lear-ning success	Instruc-tional method	RI	Mean difference	SE	df	t	p_{tukey}
Perceived	Massed	2 w	vs.	Perceived	Massed	4 w	-0.9839	0.789	239	-1.2475	0.917
Perceived	Massed	2 w	vs.	Perceived	Spaced	2 w	-2.4403	0.621	239	-3.9288	0.003
Perceived	Massed	2 w	vs.	Perceived	Spaced	4 w	-1.7014	0.710	239	-2.3957	0.248
Perceived	Massed	2 w	vs.	Real	Massed	2 w	-2.4091	0.514	239	-4.6854	< 0.001
Perceived	Massed	2 w	vs.	Real	Massed	4 w	1.7973	0.785	239	2.2910	0.303
Perceived	Massed	2 w	vs.	Real	Spaced	2 w	-3.7553	0.620	239	-6.0589	< 0.001
Perceived	Massed	2 w	vs.	Real	Spaced	4 w	-2.7602	0.707	239	-3.9026	0.003
Perceived	Massed	4 w	vs.	Perceived	Spaced	2 w	-1.4564	0.629	239	-2.3162	0.289
Perceived	Massed	4 w	vs.	Perceived	Spaced	4 w	-0.7175	0.717	239	-1.0008	0.974
Perceived	Massed	4 w	vs.	Real	Massed	2 w	-1.4252	0.785	239	-1.8164	0.610

Post hoc comparisons – Learning success * Instructional method * RI (continued)

Comparison												
Lear-ning success	Instruc-tional method	RI		Lear-ning success	Instruc-tional method	RI	Mean difference	SE	df	t		p_{tukey}
Perceived	Massed	4 w	vs.	Real	Massed	4 w	2.7813	0.522	239	5.3267	<	0.001
Perceived	Massed	4 w	vs.	Real	Spaced	2 w	-2.7714	0.627	239	-4.4167	<	0.001
Perceived	Massed	4 w	vs.	Real	Spaced	4 w	-1.7763	0.714	239	-2.4878		0.206
Perceived	Spaced	2 w	vs.	Perceived	Spaced	4 w	0.7389	0.527	239	1.4021		0.856
Perceived	Spaced	2 w	vs.	Real	Massed	2 w	0.0313	0.616	239	0.0507		1.000
Perceived	Spaced	2 w	vs.	Real	Massed	4 w	4.2377	0.624	239	6.7963	<	0.001
Perceived	Spaced	2 w	vs.	Real	Spaced	2 w	-1.3150	0.262	239	-5.0171	<	0.001
Perceived	Spaced	2 w	vs.	Real	Spaced	4 w	-0.3199	0.523	239	-0.6116		0.999
Perceived	Spaced	4 w	vs.	Real	Massed	2 w	-0.7077	0.706	239	-1.0028		0.974
Perceived	Spaced	4 w	vs.	Real	Massed	4 w	3.4988	0.712	239	4.9119	<	0.001
Perceived	Spaced	4 w	vs.	Real	Spaced	2 w	-2.0539	0.525	239	-3.9090		0.003
Perceived	Spaced	4 w	vs.	Real	Spaced	4 w	-1.0588	0.414	239	-2.5601		0.176
Real	Massed	2 w	vs.	Real	Massed	4 w	4.2064	0.780	239	5.3900	<	0.001
Real	Massed	2 w	vs.	Real	Spaced	2 w	-1.3462	0.615	239	-2.1903		0.361
Real	Massed	2 w	vs.	Real	Spaced	4 w	-0.3512	0.703	239	-0.4997		1.000
Real	Massed	4 w	vs.	Real	Spaced	2 w	-5.5527	0.622	239	-8.9243	<	0.001
Real	Massed	4 w	vs.	Real	Spaced	4 w	-4.5576	0.709	239	-6.4247	<	0.001
Real	Spaced	2 w	vs.	Real	Spaced	4 w	0.9951	0.521	239	1.9082		0.547

Appendix K – Additional analysis on learner's learning success (H5, experiment 2)

All calculations shown in this section were adapted from Jamovi, 2022.



Post hoc comparisons – Learning success * Instructional method * RI

Comparison											
Learning success	**Instructional method**	**RI**		**Learning success**	**Instructional method**	**RI**	**Mean difference**	**SE**	**df**	**t**	p_{tukey}
Perceived	Massed	2 w	vs.	Perceived	Massed	4 w	12.107	0.625	215	1.937	0.527
Perceived	Massed	2 w	vs.	Perceived	Spaced	2 w	0.1554	0.496	215	0.313	1.000
Perceived	Massed	2 w	vs.	Perceived	Spaced	4 w	0.4286	0.541	215	0.792	0.993
Perceived	Massed	2 w	vs.	Real	Massed	2 w	-10.179	0.338	215	-3.012	0.057
Perceived	Massed	2 w	vs.	Real	Massed	4 w	-0.8750	0.587	215	-1.491	0.811
Perceived	Massed	2 w	vs.	Real	Spaced	2 w	-14.184	0.484	215	-2.931	0.072
Perceived	Massed	2 w	vs.	Real	Spaced	4 w	-11.071	0.519	215	-2.132	0.398
Perceived	Massed	4 w	vs.	Perceived	Spaced	2 w	-10.553	0.496	215	-2.126	0.402
Perceived	Massed	4 w	vs.	Perceived	Spaced	4 w	-0.7821	0.541	215	-1.445	0.835
Perceived	Massed	4 w	vs.	Real	Massed	2 w	-22.286	0.587	215	-3.798	0.005

Post hoc comparisons – Learning success * Instructional method * RI (continued)

Comparison											
Learning success	**Instructional method**	**RI**		**Learning success**	**Instructional method**	**RI**	**Mean difference**	**SE**	**df**	**t**	p_{tukey}
Perceived	Massed	4 w	vs.	Real	Massed	4 w	-20.857	0.338	215	-6.172	< 0.001
Perceived	Massed	4 w	vs.	Real	Spaced	2 w	-26.291	0.484	215	-5.432	< 0.001
Perceived	Massed	4 w	vs.	Real	Spaced	4 w	-23.179	0.519	215	-4.463	< 0.001
Perceived	Spaced	2 w	vs.	Perceived	Spaced	4 w	0.2731	0.386	215	0.708	0.997
Perceived	Spaced	2 w	vs.	Real	Massed	2 w	-11.733	0.447	215	-2.623	0.154
Perceived	Spaced	2 w	vs.	Real	Massed	4 w	-10.304	0.447	215	-2.303	0.297
Perceived	Spaced	2 w	vs.	Real	Spaced	2 w	-15.738	0.173	215	-9.105	< 0.001
Perceived	Spaced	2 w	vs.	Real	Spaced	4 w	-12.626	0.354	215	-3.562	0.011
Perceived	Spaced	4 w	vs.	Real	Massed	2 w	-14.464	0.497	215	-2.912	0.075
Perceived	Spaced	4 w	vs.	Real	Massed	4 w	-13.036	0.497	215	-2.625	0.153
Perceived	Spaced	4 w	vs.	Real	Spaced	2 w	-18.470	0.370	215	-4.997	< 0.001
Perceived	Spaced	4 w	vs.	Real	Spaced	4 w	-15.357	0.239	215	-6.427	< 0.001
Real	Massed	2 w	vs.	Real	Massed	4 w	0.1429	0.546	215	0.262	1.000
Real	Massed	2 w	vs.	Real	Spaced	2 w	-0.4005	0.434	215	-0.924	0.983
Real	Massed	2 w	vs.	Real	Spaced	4 w	-0.0893	0.473	215	-0.189	1.000
Real	Massed	4 w	vs.	Real	Spaced	2 w	-0.5434	0.434	215	-1.253	0.915
Real	Massed	4 w	vs.	Real	Spaced	4 w	-0.2321	0.473	215	-0.491	1.000
Real	Spaced	2 w	vs.	Real	Spaced	4 w	0.3112	0.337	215	0.924	0.983

Appendix L – Additional analysis on learner's satisfaction (H6, experiment 1)

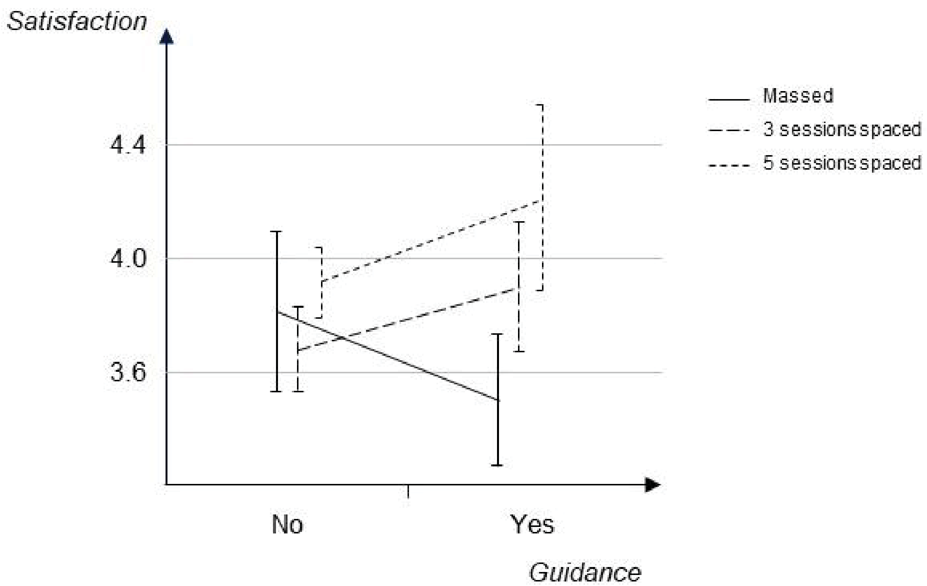
All calculations shown in this section were adapted from Jamovi, 2022.

ANOVA – Satisfaction

	Sum of squares	df	Mean square	F	p	η^2
Guidance	0.0164	1	0.0164	0.0151	0.902	0.000

ANOVA – Satisfaction

	Sum of squares	df	Mean square	F	p	η^2
Instructional method	11.509	2	5.754	5.405	0.005	0.016
Guidance	0.469	1	0.469	0.441	0.507	0.001
Instructional method * Guidance	7.405	2	3.703	3.478	0.031	0.010



Post hoc comparisons – Instructional method * Interactivity

Comparison									
Instructional method	Inter-activity		Instructional method	Inter-activity	Mean difference	SE	df	t	p_{tukey}
Massed	No	vs.	Massed	Yes	0.3257	0.1911	676	1.705	0.529
Massed	No	vs.	Spaced 3 sessions	No	0.1427	0.1645	676	0.868	0.954
Massed	No	vs.	Spaced 3 sessions	Yes	-0.0811	0.1879	676	-0.431	0.998
Massed	No	vs.	Spaced 5 sessions	No	-0.1032	0.1591	676	-0.648	0.987
Massed	No	vs.	Spaced 5 sessions	Yes	-0.4032	0.2238	676	-1.802	0.465
Massed	Yes	vs.	Spaced 3 sessions	No	-0.1830	0.1448	676	-1.264	0.805
Massed	Yes	vs.	Spaced 3 sessions	Yes	-0.4068	0.1709	676	-2.380	0.165
Massed	Yes	vs.	Spaced 5 sessions	No	-0.4289	0.1387	676	-3.092	0.025
Massed	Yes	vs.	Spaced 5 sessions	Yes	-0.7290	0.2097	676	-3.476	0.007
Spaced 3 sessions	No	vs.	Spaced 3 sessions	Yes	-0.2238	0.1406	676	-1.592	0.604
Spaced 3 sessions	No	vs.	Spaced 5 sessions	No	-0.2459	0.0989	676	-2.485	0.130
Spaced 3 sessions	No	vs.	Spaced,5 sessions	Yes	-0.5459	0.1858	676	-2.938	0.040
Spaced 3 sessions	Yes	vs.	Spaced 5 sessions	No	-0.0221	0.1343	676	-0.165	1.000
Spaced 3 sessions	Yes	vs.	Spaced 5 sessions	Yes	-0.3222	0.2068	676	-1.558	0.627
Spaced 5 sessions	No	vs.	Spaced 5 sessions	Yes	-0.3001	0.1811	676	-1.657	0.561

Note: Comparisons are based on estimated marginal means

Post hoc comparisons – Instructional method * Guidance

Comparison									
Instructional method	**Guidance**		**Instructional method**	**Guidance**	**Mean difference**	**SE**	**df**	**t**	**p_{tukey}**
Massed	No	vs.	Massed	Yes	0.3257	0.1911	676	1.705	0.529
Massed	No	vs.	Spaced 3 sessions	No	0.1427	0.1645	676	0.868	0.954
Massed	No	vs.	Spaced 3 sessions	Yes	-0.0811	0.1879	676	-0.431	0.998
Massed	No	vs.	Spaced 5 sessions	No	-0.1032	0.1591	676	-0.648	0.987
Massed	No	vs.	Spaced 5 sessions	Yes	-0.4032	0.2238	676	-1.802	0.465
Massed	Yes	vs.	Spaced 3 sessions	No	-0.1830	0.1448	676	-1.264	0.805
Massed	Yes	vs.	Spaced 3 sessions	Yes	-0.4068	0.1709	676	-2.380	0.165
Massed	Yes	vs.	Spaced 5 sessions	No	-0.4289	0.1387	676	-3.092	0.025
Massed	Yes	vs.	Spaced 5 sessions	Yes	-0.7290	0.2097	676	-3.476	0.007
Spaced 3 sessions	No	vs.	Spaced 3 sessions	Yes	-0.2238	0.1406	676	-1.592	0.604
Spaced 3 sessions	No	vs.	Spaced 5 sessions	No	-0.2459	0.0989	676	-2.485	0.130
Spaced 3 sessions	No	vs.	Spaced 5 sessions	Yes	-0.5459	0.1858	676	-2.938	0.040
Spaced 3 sessions	Yes	vs.	Spaced 5 sessions	No	-0.0221	0.1343	676	-0.165	1.000
Spaced 3 sessions	Yes	vs.	Spaced 5 sessions	Yes	-0.3222	0.2068	676	-1.558	0.627
Spaced 5 sessions	No	vs.	Spaced 5 sessions	Yes	-0.3001	0.1811	676	-1.657	0.561

Note: Comparisons are based on estimated marginal means

Appendix M – Additional analysis on learner's satisfaction (H6, experiment 2)

All calculations shown in this section were adapted from Jamovi, 2022.

ANOVA – Satisfaction

	Sum of squares	df	Mean square	F	*p*	η^2
Guidance	0.181	1	0.181	0.162	0.687	0.000

ANOVA – Satisfaction

	Sum of squares	df	Mean square	F	*p*	η^2
Instructional method	2.08	2	1.04	0.933	0.394	0.003
Guidance	1.10	1	1.10	0.982	0.322	0.002
Instructional method * Guidance	2.25	2	1.13	1.010	0.365	0.003

Appendix N – Additional information and analyses on comparison of experiment 1 and experiment 2

All calculations shown in this section were adapted from Jamovi, 2022.

Post hoc comparisons – Instructional method

Comparison							
Instructional method		Instructional method	Mean difference	SE	df	t	p_{tukey}
Massed	vs.	Spaced	-0.109	0.0214	452	-5.11	< 0.001

Note: Comparisons are based on estimated marginal means

Post hoc comparisons – RI

Comparison							
RI		RI	Mean difference	SE	df	t	p_{tukey}
2 w	vs.	4 w	0.0910	0.0214	452	4.25	< 0.001

Note: Comparisons are based on estimated marginal means

Post hoc comparisons – Knowledge type taught

Comparison							
Knowledge type taught		Knowledge type taught	Mean difference	SE	df	t	p_{tukey}
Procedural	vs.	Factual	0.224	0.0214	452	10.5	< 0.001

Note: Comparisons are based on estimated marginal means

Post hoc comparisons – Instructional method * RI

Comparison									
Instructional method	RI		Instructional method	RI	Mean difference	SE	df	t	p_{tukey}
Massed	2 w	vs.	Massed	4 w	0.1412	0.0361	452	3.907	< 0.001
Massed	2 w	vs.	Spaced	2 w	-0.0592	0.0282	452	-2.097	0.156
Massed	2 w	vs.	Spaced	4 w	-0.0183	0.0315	452	-0.580	0.938
Massed	4 w	vs.	Spaced	2 w	-0.2004	0.0289	452	-6.923	< 0.001
Massed	4 w	vs.	Spaced	4 w	-0.1595	0.0322	452	-4.954	< 0.001
Spaced	2 w	vs.	Spaced	4 w	0.0409	0.0230	452	1.782	0.283

Note: Comparisons are based on estimated marginal means

Post hoc comparisons – Instructional method * Knowledge type taught

Comparison									
Instructional method	Knowledge type taught		Instructional method	Knowledge type taught	Mean difference	SE	df	t	
Massed	Procedural	vs.	Massed	Factual	0.3110	0.0361	452	8.606	< 0.001
Massed	Procedural	vs.	Spaced	Procedural	-0.0227	0.0314	452	-0.722	0.888
Massed	Procedural	vs.	Spaced	Factual	0.1150	0.0315	452	3.651	0.002
Massed	Factual	vs.	Spaced	Procedural	-0.3336	0.0290	452	-11.504	< 0.001
Massed	Factual	vs.	Spaced	Factual	-0.1960	0.0291	452	-6.735	< 0.001
Spaced	Procedural	vs.	Massed	Factual	0.1376	0.0230	452	5.995	< 0.001

Note: Comparisons are based on estimated marginal means

Post hoc comparisons – RI * Knowledge type taught

RI	Knowledge type taught		RI	Knowledge type taught	Mean difference	SE	df	t	p_{tukey}
2 w	Procedural	vs.	2 w	Factual	0.1485	0.0282	452	5.260	< 0.001
2 w	Procedural	vs.	4 w	Procedural	0.0152	0.0314	452	0.485	0.962
2 w	Procedural	vs.	4 w	Factual	0.3154	0.0302	452	10.434	< 0.001
2 w	Factual	vs.	4 w	Procedural	-0.1333	0.0303	452	-4.395	< 0.001
2 w	Factual	vs.	4 w	Factual	0.1669	0.0291	452	5.734	< 0.001
4 w	Procedural	vs.	4 w	Factual	0.3001	0.0322	452	9.324	< 0.001

Note: Comparisons are based on estimated marginal means

Post hoc comparisons – Instructional method * RI * Knowledge type taught

Instructional method	RI	Knowledge type taught		Instructional method	RI	Knowledge type taught	Mean difference	SE	df	t	p_{tukey}
Massed	2 w	Procedural	vs.	Massed	2 w	Factual	0.17742	0.0503	452	3.527	0.011
Massed	2 w	Procedural	vs.	Massed	4 w	Procedural	0.00762	0.0539	452	0.141	1.000
Massed	2 w	Procedural	vs.	Massed	4 w	Factual	0.45216	0.0503	452	8.990	< 0.001
Massed	2 w	Procedural	vs.	Spaced	2 w	Procedural	-0.03026	0.0416	452	-0.728	0.996
Massed	2 w	Procedural	vs.	Spaced	2 w	Factual	0.08929	0.0408	452	2.186	0.362
Massed	2 w	Procedural	vs.	Spaced	4 w	Procedural	-0.00744	0.0453	452	-0.164	1.000
Massed	2 w	Procedural	vs.	Spaced	4 w	Factual	0.14829	0.0462	452	3.209	0.031
Massed	2 w	Factual	vs.	Massed	4 w	Procedural	-0.16980	0.0519	452	-3.271	0.025
Massed	2 w	Factual	vs.	Massed	4 w	Factual	0.27475	0.0482	452	5.701	< 0.001
Massed	2 w	Factual	vs.	Spaced	2 w	Procedural	-0.20768	0.0390	452	-5.328	< 0.001

Note: Comparisons are based on estimated marginal means

Post hoc comparisons – Instructional method * RI * Knowledge type taught (continued)

Comparison											
Instructional method	**RI**	**Knowledge type taught**		**Instructional method**	**RI**	**Knowledge type taught**	**Mean difference**	**SE**	**df**	**t**	p_{tukey}
Massed	2 w	Factual	vs.	Spaced	2 w	Factual	-0.08813	0.0382	452	-2.306	0.292
Massed	2 w	Factual	vs.	Spaced	4 w	Procedural	-0.18486	0.0430	452	-4.303	< 0.001
Massed	2 w	Factual	vs.	Spaced	4 w	Factual	-0.02913	0.0439	452	-0.664	0.998
Massed	4 w	Procedural	vs.	Massed	4 w	Factual	0.44455	0.0519	452	8.565	< 0.001
Massed	4 w	Procedural	vs.	Spaced	2 w	Procedural	-0.03788	0.0435	452	-0.871	0.988
Massed	4 w	Procedural	vs.	Spaced	2 w	Factual	0.08167	0.0428	452	1.908	0.546
Massed	4 w	Procedural	vs.	Spaced	4 w	Procedural	-0.01506	0.0471	452	-0.320	1.000
Massed	4 w	Procedural	vs.	Spaced	4 w	Factual	0.14067	0.0479	452	2.934	0.068
Massed	4 w	Factual	vs.	Spaced	2 w	Procedural	-0.48243	0.0390	452	-12.377	< 0.001
Massed	4 w	Factual	vs.	Spaced	2 w	Factual	-0.36288	0.0382	452	-9.495	< 0.001
Massed	4 w	Factual	vs.	Spaced	4 w	Procedural	-0.45960	0.0430	452	-10.699	< 0.001
Massed	4 w	Factual	vs.	Spaced	4 w	Factual	-0.30388	0.0439	452	-6.922	< 0.001
Spaced	2 w	Procedural	vs.	Spaced	2 w	Factual	0.11955	0.0256	452	4.662	< 0.001
Spaced	2 w	Procedural	vs.	Spaced	4 w	Procedural	0.02282	0.0323	452	0.707	0.997
Spaced	2 w	Procedural	vs.	Spaced	4 w	Factual	0.17855	0.0335	452	5.325	< 0.001
Spaced	2 w	Factual	vs.	Spaced	4 w	Procedural	-0.09673	0.0314	452	-3.084	0.045
Spaced	2 w	Factual	vs.	Spaced	4 w	Factual	0.05900	0.0326	452	1.807	0.615
Spaced	4 w	Procedural	vs.	Spaced	4 w	Factual	0.15573	0.0381	452	4.089	0.001

Note: Comparisons are based on estimated marginal means

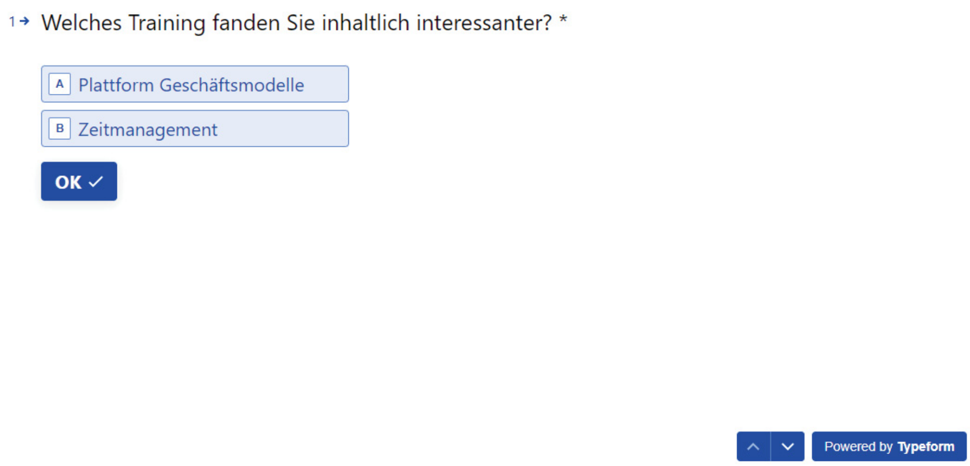
Typeform survey to gather information of learners' metacognitive comparison of experiment 1 and experiment 2

Typeform survey of learners' metacognitive comparison of experiments 1 & 2

Question no.	Question
1	Welches Training fanden Sie inhaltlich interessanter? • *Plattform Geschäftsmodelle* • *Zeitmanagement*
2	Warum fanden Sie genau dieses eben genannte Training interessanter?
3	Welches Training fanden Sie vom Trainingsaufbau effektiver? • *Plattform Geschäftsmodelle* • *Zeitmanagement*
4	Warum fanden Sie genau dieses eben genannte Training effektiver?
5	Welcher Aussage stimmen Sie eher zu? • *Der interessantere Inhalt hat mehr zu meinem persönlichen Lernerfolg beigetragen als der unterschiedliche Aufbau.* • *Der effektivere Trainingsaufbau hat mehr zu meinem persönlichen Lernerfolg beigetragen als der unterschiedliche Inhalt.*
6	Nach welchem Training haben Sie gefühlt mehr über das gelehrte Thema gewusst? • *Plattform Geschäftsmodelle* • *Zeitmanagement*
7	Was hat am meisten zu Ihrem gefühlten Wissenszuwachs beigetragen? Mehrfachnennung möglich • *Interesse* • *Anwendbarkeit* • *Ablauf des Trainings* • *Aufbereitung der Medien (Folien, Videos)* • *Ablauf der Live Session(s)* • *Andere*
8	Auf einer Skala von 1 (stimme überhaupt nicht) bis 5 (stimme komplett zu) wie sehr stimmen Sie der folgenden Aussage (nicht) zu: Das Thema ‚Zeitmanagement' war relevanter für meinen Arbeitsalltag als das Thema ‚Plattform Geschäftsmodelle'.
9	Auf einer Skala von 1 (stimme überhaupt nicht) bis 5 (stimme komplett zu) wie sehr stimmen Sie der folgenden Aussage (nicht) zu: Die Inhalte des Themas ‚Zeitmanagement' konnte ich mir besser merken, als die des Themas ‚Plattform Geschäftsmodelle'.
10	Auf einer Skala von 1 (stimme überhaupt nicht) bis 5 (stimme komplett zu) wie sehr stimmen Sie der folgenden Aussage (nicht) zu: Allgemein merke ich mir Themen, die alltagsrelevanter sind, eher als solche, für die ich nur ein intrinsisches Interesse habe.

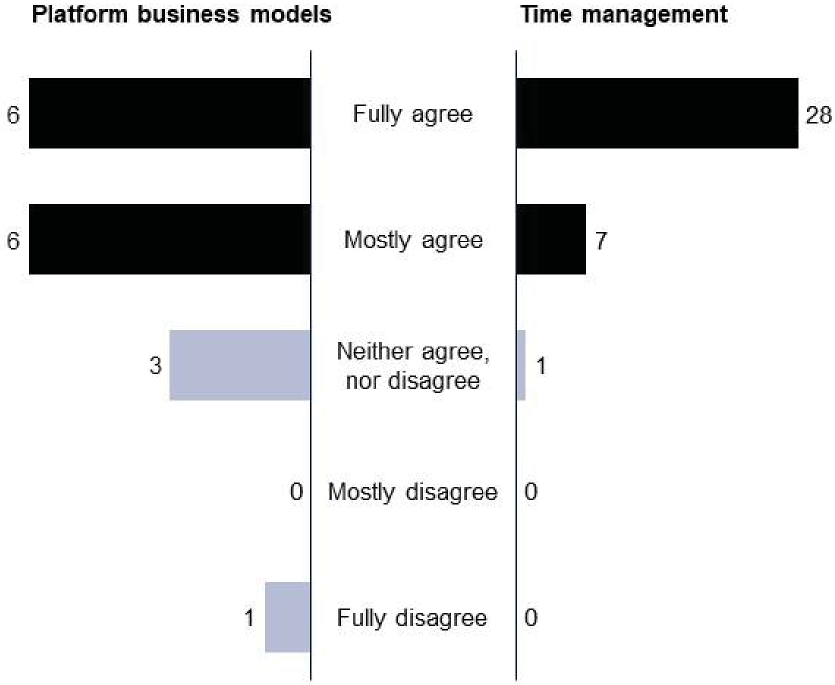
Typeform survey of learners' metacognitive comparison of experiments 1 & 2 (cont'd)	
Question no.	**Question**
11	Auf einer Skala von 1 (stimme überhaupt nicht) bis 5 (stimme komplett zu), was glauben Sie, dass mehrere Wiederholungseinheiten besonders stark bei Fakten basiertem Wissen, wie am Beispiel des Plattform Geschäftsmodells helfen?
12	Auf einer Skala von 1 (stimme überhaupt nicht) bis 5 (stimme komplett zu), was glauben Sie, dass mehrere Wiederholungseinheiten besonders stark bei prozessualem Wissen, wie am Beispiel des Zeitmanagements helfen?
13	Haben Sie noch weitere Anmerkungen zu den Inhalten, Aufbau der Trainings, die Sie mir gerne mitteilen möchten?

Preview mode on Typeform



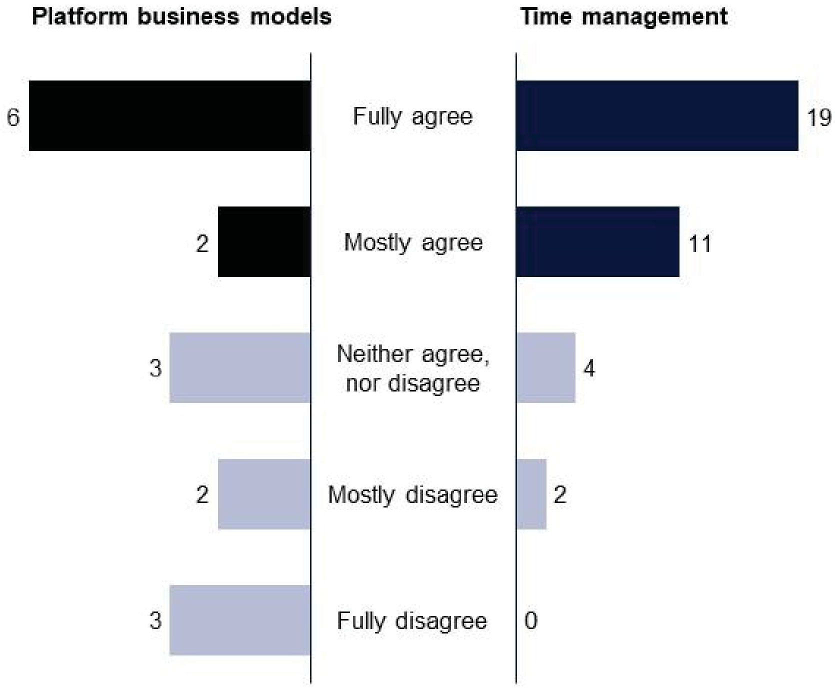
Survey question:
"The topic of time management was more relevant for my daily work than the topic of platform business models."

Survey answers for participants who found below training more interesting:



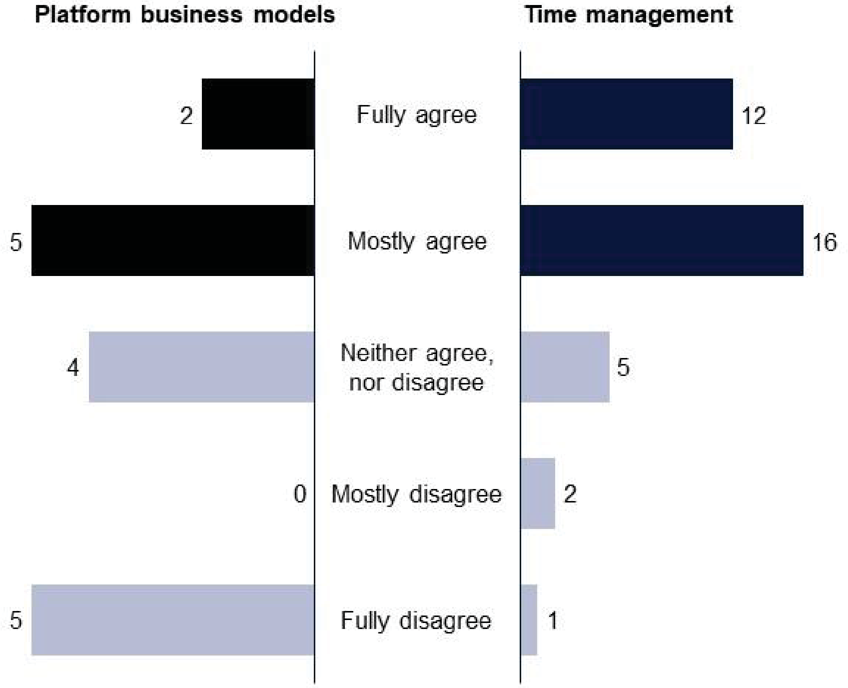
Survey question:
"I could remember the contents for the topic of time management better than these of platform business models."

Survey answers for participants who found below training more interesting:



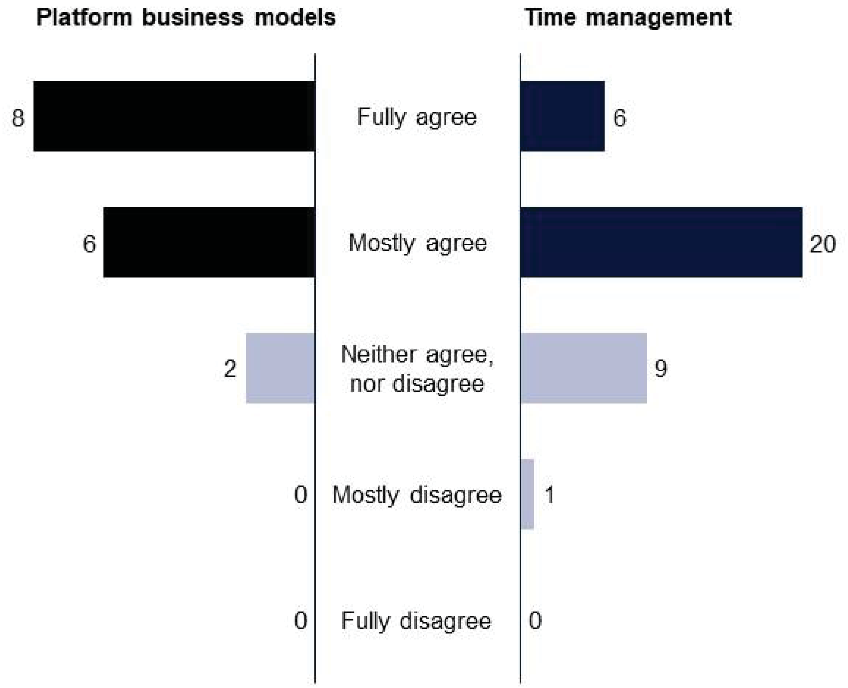
Survey question:
"Generally I remember topics that are relevant for daily work better than those that I just have an intrinsic interest in."

Survey answers for participants who found below training more interesting:



Survey question:
"Multiple repetitions help particularly well for factual knowledge as with the example of platform business models."

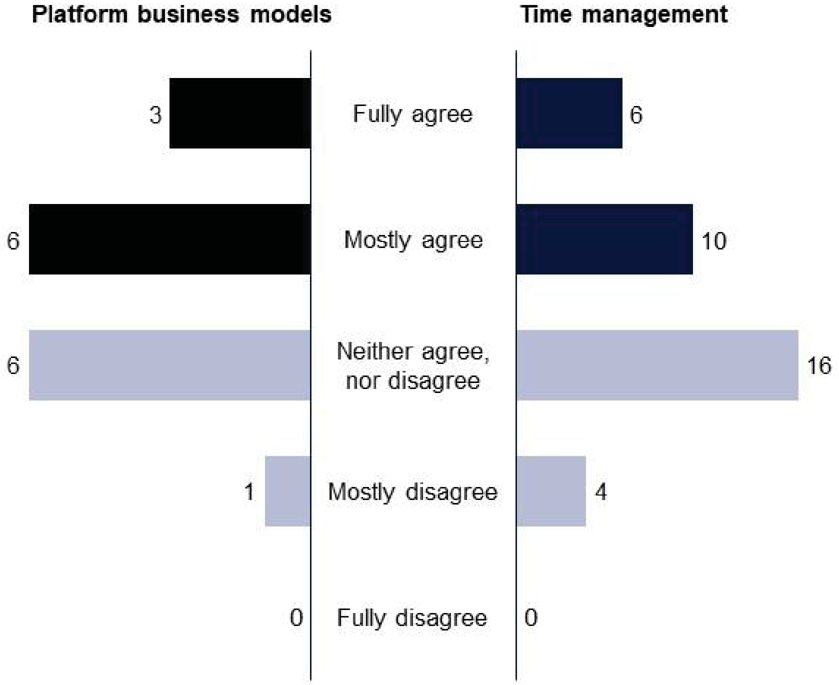
Survey answers for participants who found below training more interesting:



Survey question:

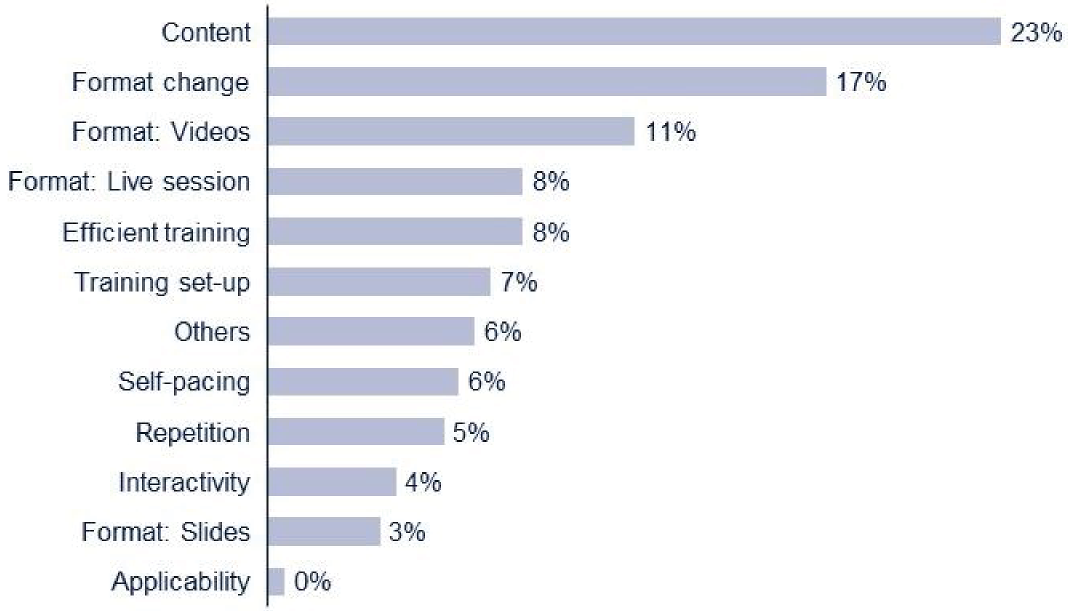
"Multiple repetitions help particularly well for procedural knowledge as with the example of time management."

Survey answers for participants who found below training more interesting:



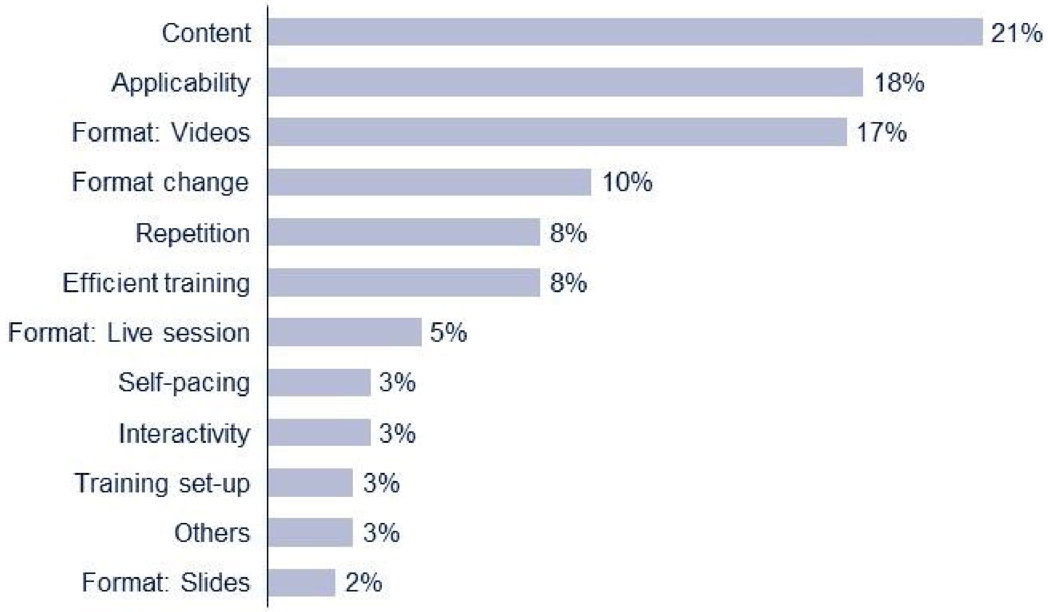
Appendix O – Thematic coding analysis of surveys

Survey answers – Platform business models (n = 202)



Note: Percentages may not add up to 100% due to rounding

Survey answers – Time management (n = 197)



Note: Percentages may not add up to 100% due to rounding

Appendix P – E-Mail traffic with Benjamin Aaron

From: Benjamin, Aaron S <asbenjam@illinois.edu>
Sent: Tuesday, 6 July 2021 20:13
To: Hanan Kondratjew
Subject: Re: Introduction and request

I'm afraid I don't have time to do a thorough review, in part because that would require me to know more about the materials and the relevant work in your own field. I can say: this looks like a reasonable design, if a little complex; it will require an extremely large number of subjects. I haven't computed power but to detect some of the smaller differences with any confidence (say, between expanded and regular RP) I would guess you will need 200+ subjects per condition. This will depend on individual differences in the learning of your materials, of course.

--aaron

____
Aaron S. Benjamin
Professor and Assistant Head, Department of Psychology, University of Illinois at Urbana-Champaign
Editor, *Journal of Experimental Psychology: Learning, Memory, and Cognition*

http://labs.psychology.illinois.edu/~asbenjam/

From: Hanan Kondratjew <hanan.kondratjew@hhu.de>
Sent: Monday, July 5, 2021 7:02 AM
To: Benjamin, Aaron S <asbenjam@illinois.edu>
Subject: RE: Introduction and request

Dear Dr Benjamin,

I am so grateful for you getting back to me. Unfortunately, I fell sick; hence, I couldn't get back to you earlier than now.
Many thanks for the recommendations you sent – I will investigate these.

What would be great for me to discuss is the experimental set up of my research. My research is an applied one: I am working closely with a company which offers online learning on digital topics. I am aiming at running five experimental groups in parallel; trying to find the best schedule for long term memory retention for this training.

For me it would be great to understand, if my set up makes sense at all from a spaced learning research perspective; as I am part of the business economics and management faculty nobody really knows anything about this topic and I would love to be in line with previous research, if possible.

Attached, I am sending you the first draft of what I handed in as proposed experimental design and hypothesis.

Again, many many thanks for getting back to me.

Warm regards,
Hanan

Hanan Kondratjew
Lehrstuhl für BWL, insb. Management | Heinrich-Heine-Universität Düsseldorf
Universitätsstraße 1 | 40225 Düsseldorf