The Testing Effect in the Lecture Hall: How to Adapt it to Individual Characteristics and Situational Circumstances?

Inaugural-Dissertation

zur Erlangung der Doktorwürde der Naturwissenschaft (Dr. rer. nat.) der

Fakultät für Humanwissenschaften

der

Julius-Maximilians-Universität Würzburg

Vorgelegt von

Julia Glaser

aus Alzenau-Albstadt

Würzburg, Januar 2024

Betreuer: Professor Dr. Tobias Richter

Erstgutachter: Professor Dr. Wolfgang Lenhard

Tag der Disputation: 22.03.2024

Summary

Learning accompanies us throughout our lives, from early childhood education through school, training and university to learning at work. However, much of what we learn is quickly forgotten. The use of practice tests is a learning strategy that contributes to the acquisition of sustainable knowledge, i.e. knowledge that is permanently available and can be retrieved when it is needed. This dissertation first presents findings from previous research on testing in real educational contexts and discusses theoretically why certain learner or situational characteristics might influence the effectiveness of the testing effect. Furthermore, a cycle of three experiments is presented, which were used to investigate whether the positive effect of practice tests on retention (testing effect) depends on personal or situational characteristics and also promotes the retention of lecture content that was not directly tested (transfer) in the context of regular psychology lectures in teacher training courses. In an additional chapter, feedback from students on the implementation of the study in the classroom context is examined in more detail. Finally, the results of the three studies are discussed and placed in relation to the theories presented. The central conclusion from the studies presented is that the testing effect appears to be a very effective learning strategy that can be used effectively in university teaching and leads to better learning outcomes regardless of learner characteristics. However, the practice tests should cover the entire range of relevant content, as transfer effects to non-tested content are not to be expected.

Zusammenfassung

Lernen begleitet uns unser Leben lang, von der frühkindlichen Bildung über Schule, Ausbildung und Universität bis hin zum Lernen im Beruf. Vieles von dem, was wir lernen, wird aber rasch wieder vergessen. Die Nutzung von Übungstests ist eine Lernstrategie, die zum Erwerb *nachhaltigen Wissens* beiträgt, also Wissen, das dauerhaft zur Verfügung steht und abgerufen werden kann, wenn es benötigt wird. In dieser Dissertation werden zunächst Befunde aus der bisherigen Forschung zum Testen in realen Bildungskontexten vorgestellt und theoretisch erörtert, warum bestimmte Lerner- oder Situationsmerkmale die Effektivität des Testungseffekts beeinflussen könnten. Weiterhin wird ein Zyklus aus drei Experimenten vorgestellt, anhand derer im Rahmen regulärer Psychologievorlesungen im Lehramtsstudium erforscht wurde, ob der positive Effekt von Übungstests auf das Behalten (Testungseffekt) von Personen- oder Situationsmerkmalen abhängt und auch das Behalten von Vorlesungsinhalten fördert, die nicht direkt getestet wurden (Transfer). In einem Zusatzkapitel wird Feedback von Studierenden zur Umsetzung der Studie im Klassenraumkontext näher beleuchtet. Die Ergebnisse der drei Studien werden abschließend diskutiert und in Bezug zu vorgestellten Theorien gesetzt. Die zentrale Schlussfolgerung aus den vorgestellten Studien ist die Erkenntnis, dass der Testungseffekt eine sehr wirksame Lernstrategie zu sein scheint, die sich sinnvoll in der Hochschullehre einsetzen lässt und unabhängig von Lernermerkmalen zu besseren Lernergebnissen führt. Die Übungstests sollten aber die gesamte Bandbreite relevanter Inhalte abdecken, da Transfereffekte zu nicht getesteten Inhalten nicht zu erwarten sind.

Table of Contents

Introduction

How can learning be optimized? This is the question that countless scientists around the globe are addressing. Especially in times like these, when digital learning is being pushed forward by the pandemic, but also by the wealth of new possibilities, the question arises as to the optimal use of learning strategies and technologies, adapted to the users. The difficulty in this is that there is no teacher to address the individual needs of learners (Wong et al., 2019). So how can the right strategies, adapted to the individual, be offered from the onset to learn optimally?

A promising learning strategy when it comes to broad effectiveness is self-testing: learners acquire the knowledge and can consolidate it by means of practice questions and check their knowledge level at the same time (Roediger & Butler, 2011). This strategy has been found by Dunlosky and colleagues, among others, to be more effective for sustainable retrieval than other common learning strategies, such as memorization, mind mapping, or script writing (Dunlosky et al., 2013). The positive effect of this strategy is called the testing effect and has been demonstrated in many contexts with a medium to large effect size (e.g. $g = 0.51$, Adesope et al., 2017; $g = 0.50$, Rowland, 2014; $g = 0.33$, Yang et al., 2021; $d = 0.73$, Schwieren et al., 2017). It can be divided into the direct effects of testing - the increase in knowledge from answering the questions - and the indirect effects of testing - the detection of knowledge gaps (Arnold & McDermott, 2013).

Why does practice testing in particular have such a lasting effect on memory? Testing is one of the so-called desirable difficulties, which means that the learning process is intentionally made a little more difficult to increase the mental effort in order to achieve greater learning success later on (Bjork & Bjork, 2011).

Jonsson and colleagues (2021) entitled testing as a "learning method for all" due to its benefits appearing relatively independent of student characteristics (Jonsson et al., 2021). And not only for various groups of people, but also for many different contexts the positive effect

of testing could be demonstrated: using different materials, periods of learning and laboratory as well as real educational context (for a overview see Adesope et al., 2017 or Rowland et al., 2014). Confirmed moderators are for example retrievability, which plays an important role dependent on the presence of feedback, or the kind of feedback given (Greving et al., 2022; Greving & Richter, 2018; Kornell & Vaughn, 2016; Kornell et al., 2015). But despite these many already confirmed moderators, there are still several unexplained potential individual moderators on the testing effect.

Therefore, the main goal of this dissertation is to investigate how practice testing as a learning strategy of self-directed learning can be adapted to individual characteristics or situational circumstances to generate best possible learning progress in a real-educational online setting. For this purpose, this dissertation aims to clarify four questions.

The first question addresses the effectiveness of testing in a digital, self-regulated, real-educational learning environment. In the past, learning mostly took part in the classroom in presence of a teaching person. But since there is currently a shift towards more independent work, which increasingly has to manage without the support of a teacher, the question arises as to whether the strategies tried and tested in the classroom also work well in an online setting without much guidance.

The second question to be answered is whether there is any influence of individual learner characteristics on the effectiveness of testing. In contrast to other moderators, learner characteristics have received little attention to date, although there are so many of them like cognitive capacity, motivational disposition, or anxiety. But it can be assumed that they can have an influence on the effectiveness of testing as well as other moderators (Kubik et al., 2021), especially in self-regulated learning contexts when learners are given a lot of flexibility.

The third question to be examined is whether characteristics of the learning situation influence the effectiveness of testing, for example the need to transfer what has been learned

or the metacognitive awareness of the efficacy of testing in the learning situation. Exams are normally not just a simple knowledge test, but usually also involve transfer tasks. There is already research on transfer (see meta-analysis Pan & Rickard, 2018), but it is inconclusive in some respects. Therefore, there is a need for further research into the extent to which testing is also a useful strategy for knowledge transfer. Further, while learning, many students do not think about what they do. In research on metacognitive activation, however, it has been observed that learners who know why the strategy used is effective can achieve better learning outcomes (Schwonke et al., 2013). Therefore, research is still needed in this regard.

The fourth question deals with the feasibility of testing in the classroom, especially from the students' perspective. Students often use less effective strategies in their daily learning, mostly because of a lack of education in learning strategies or false beliefs about effective learning, consolidated over several years (Bjork et al., 2013). But even when knowing good strategies, they often do not execute them spontaneously, concerned that they are expending more resources than the final benefits would be worth (Bannert, 2009; Foerst et al., 2017). When now making students use the strategy and providing them the necessary resources, how do they experience its usefulness for learning? How easily does it fit their daily learning?

To address these four questions, three empirical studies have been conducted. The first study (Chapter II) focuses on the issue of whether individual learner characteristics have an impact on the effectiveness of testing in an online setting. Motivational, cognitive, and emotional variables are examined. The second study (Chapter III), in which the design of the first study was taken up and extended to a longer observation period, investigates whether the testing effect is also evident for transfer knowledge that was learned but not explicitly tested in this way. In addition, further individual characteristics are included as potential moderators. The third study (Chapter IV), as a conceptual replication of the first study, examines whether metacognitive activation of students has an impact on the effectiveness of testing. In the last two studies, student data were collected on the perceived usefulness and handling of testing as

a learning strategy in order to evaluate the implementation of the strategy in everyday teaching or learning (Chapter V).

This dissertation is structured as follows: Chapter I provides an overview of previous research and relevant theories in the field of testing, setting the foundation for the aims of this dissertation. Chapters II to IV present the studies that were conducted to achieve the aims of this dissertation. In Chapter V, a cross-study review of student feedback on the use of testing as a learning strategy is presented and discussed. Chapter VI summarizes all of the studies' findings and discusses their collective implications for theory and practice.

References

Adesope, O. O., Trevisan, D. A., & Sundararajan, N. (2017). Rethinking the use of tests: A meta-analysis of practice testing. *Review of Educational Research, 87*(3), 659–701. https://doi.org/10.3102/0034654316689306

Arnold, K. M., & McDermott, K. B. (2013). Test-potentiated learning: Distinguishing between direct and indirect effects of tests. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39*(3), 940–945. https://doi.org/10.1037/a0029199

Bannert, Maria (2009): Promoting Self-Regulated Learning Through Prompts. *Zeitschrift für Pädagogische Psychologie 23*(2), 139-145. https://doi.org/10.1024/1010-0652.23.2.139

Bjork, E. L., & Bjork, R. A. (2011). Making things hard on yourself, but in a good way: Creating desirable difficulties to enhance learning. In M. A. Gernsbacher, R. W. Pew, L. M. Hough, & J. R. Pomerantz (Eds.), *FABBS Foundation, pschology and the real world: Essays illustrating fundamental contributions to society* (pp. 56–64). Worth Publishers.

Bjork, R. A., Dunlosky, J., & Kornell, N. (2013). Self-regulated learning: Beliefs, techniques, and illusions. *Annual Review of Psychology*, *64*, 417-444. https://doi.org/annurev-psych-113011-143823

Dunlosky, J., 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

Foerst, N. M., Klug, J., Jöstl, G., Spiel, C., & Schober, B. (2017). Knowledge vs. Action: Discrepancies in University Students' Knowledge about and Self-Reported Use of Self-Regulated Learning Strategies. *Frontiers in Psychology, 8,* 1288. https://doi.org/10.3389/fpsyg.2017.01288

Greving, S., Lenhard, W., & Richter, T. (2023). The Testing Effect in University Teaching: Using Multiple-Choice Testing to Promote Retention of Highly Retrievable Information. *Teaching of Psychology, 50*(4), 332-341. https://doi.org/10.1177/00986283211061204

Greving, S., & Richter, T. (2018). Examining the testing effect in university teaching: Retrievability and question format matter. *Frontiers in Psychology, 9*, 2412. https://doi.org/10.3389/fpsyg.2018.02412

Jonsson, B., Wiklund-Hörnqvist, C., Stenlund, T., Andersson, M., & Nyberg, L. (2021). A learning method for all: The testing effect is independent of cognitive ability. *Journal of Educational Psychology, 113*(5), 972–985. https://doi.org/10.1037/edu0000627

Kornell, N., Klein, P. J., & Rawson, K. A. (2015). Retrieval attempts enhance learning, but retrieval success (versus failure) does not matter. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 41*(1), 283. https://doi.org/10.1037/a0037850

Introduction

Kornell, N., & Vaughn, K. E. (2016). How retrieval attempts affect learning: A review and synthesis. In B. H. Ross (Ed.), *The psychology of learning and motivation* (pp. 183–215). Elsevier Academic Press.

Kubik, V., Gaschler, R., & Hausman, H. (2021). Enhancing student learning in research and educational practice: The power of retrieval practice and feedback. *Psychology Learning & Teaching, 20*(1), 1–20. https://doi.org/10.1177/1475725720976462

Pan, S. C., & Rickard, T. C. (2018). Transfer of test-enhanced learning: Meta-analytic review and synthesis. *Psychological Bulletin, 144*(7), 710–756. https://doi.org/10.1037/bul0000151

Roediger, H. L.III, & Butler, A. C. (2011). The critical role of retrieval practice in long-term retention. *Trends in Cognitive Sciences, 15*(1), 20–27. https://doi.org/10.1016/j.tics.2010.09.003

Rowland, C. A. (2014). The effect of testing versus restudy on retention: A meta-analytic review of the testing effect. *Psychological Bulletin, 140*(6), 1432–1463. https://doi.org/10.1037/a0037559

Schwieren, J., Barenberg, J., & Dutke, S. (2017). The testing effect in the psychology classroom: A meta-analytic perspective. *Psychology Learning & Teaching, 16*(2), 179–196. https://doi.org/10.1177/1475725717695149

Schwonke, R., Ertelt, A., Otieno, C., Renkl, A., Aleven, V., & Salden, R. J. C. M. (2013). Metacognitive support promotes an effective use of instructional resources in intelligent tutoring. *Learning and Instruction, 23*, 136–150. https://doi.org/10.1016/j.learninstruc.2012.08.003

Wong, J., Baars, M., Davis, D., Van Der Zee, T., Houben, G. J., & Paas, F. (2019). Supporting self-regulated learning in online learning environments and MOOCs: A systematic review. *International Journal of Human–Computer Interaction, 35*(4-5), 356-373. https://doi.org/10.1080/10447318.2018.1543084

Yang, C., Luo, L., Vadillo, M. A., Yu, R., & Shanks, D. R. (2021). Testing (quizzing) boosts classroom learning: A systematic and meta-analytic review. *Psychological Bulletin*, *147*(4), 399–435. https://doi.org/10.1037/bul0000309

Chapter I

The Testing Effect in the Lecture Hall: Evidence and Open Questions

Learning in the Lecture Hall

Lifelong learning is a frequently used term in today's world. From an early age, learning accompanies us throughout our lives. It is generally about acquiring knowledge in a sustainable way so that it is permanently available, can form the basis for new learning and can be called upon in the situations in which it is needed (Richter et al., 2022). Self-regulated learning in particular is becoming increasingly important in today's highly digitalized world. Learners have to choose the right strategies on their own and apply them correctly in order to learn successfully and sustainably (Wong et al., 2019). Unfortunately, students often choose the wrong strategies that ensure the short-term retention of learning materials but disregard the long-term, for example re-reading or memorization (e.g. Karpicke, Butler & Roediger, 2009; Kornell & Bjork, 2007; Taraban, Maki & Rynearson, 1999). But why do they do that? Mostly because they have wrong beliefs about effective learning (Bjork et al., 2013). The most common of these misconceptions is the belief that learning has to be easy – the so-called feeling of fluency (Biwer et al., 2020). Learners think that learning is only effective if they can quickly reproduce the correct answer, and this comes easily to them. However, this is in contrast to the concept of desirable difficulties in learning, which has been repeatedly confirmed by science (Bjork & Bjork, 2011). This says that learning becomes sustainable when the process of learning is deliberately made more difficult, stimulating deeper processing that leads to longer-term recall. It is this deliberate impediment that also prevents many students – but not all (see Kuhbandner & Emmerdinger (2019), who claim that students use multiple learning strategies, also including testing, spontaneous) - from spontaneously using effective strategies (Bannert, 2009). They misjudge the relationship between the effort required and the benefits of the strategy and do not give it a chance, either out of demotivation or due to a lack of resources (Foerst et al., 2017).

There are many strategies that can be assigned to desirable difficulties and can be easily integrated into everyday learning. One of the most effective strategies is practice testing, which means that previously learned knowledge is recalled and thus deepened (Roediger & Butler, 2011). This is typically researched using the design also used by Roediger & Karpicke (2006): After a one-time presentation of learning material, learners are either asked practice questions to repeat the learning content or are given summarizing statements, so-called restudy. This design has been studied by many researchers (e.g. Karpicke et al., 2009) and has repeatedly shown that tests are disadvantageous for short-term retention but clearly advantageous for long-term retention (even over a few days to several months).

The aim of this dissertation is to investigate whether and how practice tests as a strategy of self-directed learning can be adapted to individual characteristics or situational circumstances in order to achieve the best possible learning progress in a real educational online setting. To this end, I will first explain the theoretical mechanisms driving the testing effect. In the following chapter I review existing research on practice testing in an educational context. Further, I also investigate the moderators that have already been researched and the reasons for adopting further moderators, which ultimately leads to the formulation of the specific research questions that are addressed by the studies presented in Chapters II-V.

Theoretical Mechanisms Underlying the Testing Effect

The testing effect is based on two mechanisms that contribute to its effectiveness: the direct testing effect on the one hand and the indirect testing effect on the other. Direct testing effect means that testing itself improves the later retrieval of knowledge from the long-term memory by practicing the retrieval (e.g., Carpenter, 2009). The indirect testing effect, on the other hand, refers to the metacognitive benefits of testing, e.g., that gaps in knowledge are discovered and future learning activities are influenced (e.g., Barenberg & Dutke, 2019;

Soderstorm & Bjork, 2014). Several theoretical approaches to the functioning of these two effects are discussed in the current literature on the testing effect. Mainly, the different approaches exist side by side and are not mutually exclusive. In the following, I will present the most relevant ones to draw conclusions about potential effects in the classroom context.

Mechanisms Driving the Direct Testing Effect

One approach of mechanisms driving the direct testing effect is Carpenter's (2009) elaborative retrieval hypothesis, according to which the positive effect of testing can be explained by the fact that related information is also activated when the searched information is retrieved, which strengthens the integration of what has been learned into long-term memory (Carpenter, 2009).

Another one is the mediator effectiveness hypothesis of Pyc & Rawson (2010), that is quite similar to Carpenter's approach. It proposes that testing creates a link between other information and the response, a so-called mediator, which can be activated by cues (Pyc & Rawson, 2010).

The idea behind transfer-oriented processing is that the practice and test contexts are more similar when exercising with practice questions than with restudy, thereby creating an advantage for the later test situation (Morris et al., 1977).

In summary, it can also be said that the testing effect is one of the so-called desirable difficulties, which works by making the learning process more difficult, but thereby facilitating later recall through deeper processing and better storage (Bjork & Bjork; 2011).

Mechanisms Driving the Indirect Testing Effect

The indirect effect of testing is strongest when students are given feedback on the validity of their answers. This gives them information about their performance and allows

them to realistically assess what they still need to learn. The so-called illusion of knowing, i.e. that test subjects have the feeling that they know the correct answer by reading it, is prevented (Glenberg et al., 1982). Furthermore, the feedback received can also be used to assess the quality of learning and future learning behavior can be adapted accordingly. In this way, metacognitive awareness of effective learning can improve and have a lasting influence on learning (e.g., Barenberg & Dutke, 2019; Soderstrom & Bjork, 2014).

Practical Implications of Theoretical Accounts on the Testing Effect

From the theories on the functioning of the direct testing effect, it can be concluded that testing is not only a highly effective strategy for consolidating but also for activating linked knowledge (see elaborative retrieval hypothesis, Carpenter, 2009; mediator effectiveness hypothesis, Pyc & Rawson, 2010). This linked knowledge is what we need for transfer tasks. But not exclusively the direct effects of testing favor transfer, also the indirect effects do. So, Chan and colleagues (2006) have developed the concept of retrieval-induced facilitation, which states that under certain conditions, testing leads to better remembering of other but the learned information (Chan et al., 2006). It seems that teachers could advise testing also as a strategy for transfer tasks. Unfortunately, research does not agree about the amount of transfer possible, so further studies are needed, like Study 2.

Further, testing should profit from metacognitive activation. If one is metacognitively activated, i.e., if one is aware of the effectiveness of the learning strategy and how it works, one can use it in a much more targeted way. For example, feedback will be used more effectively because one knows that it is useful. Therefore, it is desirable that students are metacognitive activated when learning with the strategy of testing or any other strategy of the desirable difficulties.

Chapter I

The Testing Effect in the Lecture Hall

Testing was demonstrated to be a very robust phenomenon with medium to large effect sizes both in the laboratory ($g = 0.51$, Adesope et al., 2017; $g = 0.50$, Rowland, 2014) and in the classroom context ($g = 0.33$, Yang et al., 2021; $d = 0.73$, Schwieren et al., 2017). Agarwal and colleagues (2021) could show within a meta-analysis that 57% of 49 studies had medium to large effect sizes in educational settings (Agarwal et al., 2021). Thereby, the testing effect was applied to a wide range of materials, learning periods and contexts.

A major challenge in the classroom, however, is the diversity of individuals and demands on the learning outcome, which cannot always be controlled. Especially in self-regulated learning contexts, the characteristics of individual learners could lead to significant differences in the benefits of practice testing. Both cognitive and non-cognitive (like emotional or motivational) variables are likely moderators of testings' effectiveness but have been little researched to date (Kubik et al., 2022).

Further, there is the proverb "We don't learn for school, we learn for life", which alludes to the desired transfer of learning content. There is some research about the effects of testing on transfer: Pan & Rickard (2018) found a medium-sized positive effect of practice testing across different types of transfer, but not all results are conclusive (see meta-analysis by Pan & Rickard, 2018). From a theoretical point of view, transfer is quite possible due to the testing. As shown by Carpenter's elaborative retrieval hypotheses (2009), when a response is retrieved from long-term memory, not only the specific memory but also linked memories are activated (Carpenter, 2009). In the classroom context, that would be beneficial as teachers rarely use the practice questions in the final exam but ask similar concepts.

Moderators of the Testing Effect

Moderators Already Researched

There is a bunch of already researched moderators concerning the effectiveness of the testing effect, but most of them concentrate on the used materials or circumstances of testing. For example, Butler and Roediger (2008) found out, that the provision of feedback is important (Butler & Roediger, 2008). They showed, that, especially for tests with multiple-choice questions, feedback helps students to benefit from testing. Otherwise, a dangerous feeling of knowing could appear and prevent successful learning (Butler & Roediger, 2008; Roediger & Marsh, 2005). It is therefore recommended that students are always given feedback so that the benefits of the indirect testing effect can be exploited.

Another important moderator of testings' effectiveness is the type of questions used in practice test (Greving & Richter, 2018; McDaniel & Little, 2019). It could be shown that short answer questions were more reliable and led to a stronger testing effect than multiple choice questions. One obvious reason for this is the greater cognitive effort required for free (short-answer questions) compared to cued recall (multiple-choice questions).

Further, characteristics of the learning materials matter, for example their complexity (Van Gog & Sweller, 2015). This is a really important factor thus educational materials are often interactive and therefore high in complexity. Van Gog & Sweller (2015) could show, that with increasing complexity of learning materials, the testing effect decreases or even disappears.

Finally, if testing is in a context without feedback, retrievability (for the first trial) seems to be a very important influence factor on the effectiveness on testing (Greving et al., 2022; Greving & Richter, 2018), as it ensures that the correct answer is kept. With feedback it

is less important (Kornell & Vaughn, 2016; Kornell et al., 2015) due to the possibility of correction.

All of these moderators are issues that need to be considered when planning practice testing into lessons. For example, it is always advisable to give feedback and use short-answer questions. Things like the complexity of the learning materials are less easy to change - here it's important to consider whether other learning strategies might cover this better.

Further Assumed Moderators

Less well researched and therefore relatively unnoticed are moderators that incorporate the individual characteristics of the learners. The following is a brief overview - these are examined in more detail in Study 1 (Chapter II).

In line with the elaborative retrieval hypothesis or the mediator effectiveness hypothesis, learners with better cognitive resources, i.e. better abilities and also more prior knowledge, could benefit more from testing than learners with poorer prerequisites (Carpenter, 2009; Pyc & Rawson, 2010). Contrary to this, however, Jonsson and colleagues (2021) have shown that the testing effect appears to be independent of cognitive abilities (Jonsson et al., 2021). There is therefore a need for further research.

As practice testing is a desirable difficulty (Bjork & Bjork, 2011) and learners have to put in more effort than in other strategies, also motivational factors are assumable moderators of the testing effect. Different motivational factors could play a role, for example the general performance motivation, the concept of errors (so-called *error-orientation*; Rybowiak et al., 1999) or the pleasure in thinking (so-called *Need for Cognition*; Cacioppo et al., 1984).

And also, closely linked to motivational factors, emotional dispositions could play a role for the effectiveness of testing. As the character of desirable difficulties is that failure is part of it and therefore it is more aversive than other learning methods, learners with fear of

failure or test anxiety in general could have problems with testing (e.g., Ramsden, 2003; Hinze & Rapp, 2014).

Finally, metacognitive activation of learners could crucially influence the effectiveness of testing. Within the concept of constructive retrieval (Hinze, Wiley & Pellegrino, 2013), the authors claim that intentional, active use could be more effective than passive use of a strategy, because more cognitive resources get activated and so a deeper processing is achieved.

These potential moderators will be investigated in the studies conducted. To this end, the research questions are formulated below.

Research Questions

In the previous chapters, I described how far research on the testing effect in real educational contexts has come so far and where there are still starting points for further research. I have highlighted the theories behind the functioning of the testing effect and the moderators that are already known. The aim of this dissertation is to gain a more thorough understanding of the testing effect by exploring other potential moderating factors or circumstances of testing. To this purpose, a set of research questions is explored.

The main aim of the first study (Chapter II) was to investigate whether practice testing can be effectively applied in a digital, real-world educational context, meaning that knowledge learned via questions can be better recalled than knowledge learned via restudy. It should also be shown whether individual differences in learners influence the effectiveness of testing. Studies on the testing effect almost always reveal individual differences. Not all learners benefit equally from practice tests. It is conceivable and theoretically plausible to assume that learners must have certain prerequisites - of a cognitive, emotional, or motivational nature - in order to make the best possible use of practice tests for their learning.

In this study, students first completed a questionnaire to assess their individual characteristics. They then attended their lecture session, directly answered knowledge questions about it or read repetitive statements about it (manipulated within participants) and answered test questions a week later in order to observe the testing effect. The moderating effect of the individual characteristics on the testing effect could thus be investigated.

Study 2 (Chapter III), in which the general design of Study 1 was adopted, sought to find out what influence testing has on the transfer of knowledge. Positive effects of practice tests on retention have so far only been clearly demonstrated for content that appears in the practice tests themselves. The extent to which the testing effect also generalizes to thematically related content that occurred in the same teaching unit but was not directly tested is still open (see the meta-analysis by Pan & Rickard, 2018). A practice test with questions or summarizing items was administered 2 days after the lecture session attended by the students. This was repeated over 6 weeks for 6 different lecture topics. In the seventh week of the study, there was a knowledge test in which half of the questions were similar to those from the practice test, while the other half asked about content-related transfer knowledge.

The main aim of Study 3 (Chapter IV), another conceptual replication of Study 1, was to investigate the influence of metacognitive activation on the effectiveness of testing. Research on the role of metacognition in learning (Schneider, Tibken & Richter, 2022) suggests that learners might be able to use practice tests for sustained learning even more effectively if they are aware that and how active retrieval from long-term memory promotes long-term retention. An intervention that makes corresponding metacognitive beliefs salient shortly before completing the practice tests could therefore further strengthen the testing effect. In this study, that was investigated by incorporating a short intervention into the practice session immediately following the lecture for half of the test subjects to remind them of the effectiveness of testing and thus activate their metacognitive knowledge. In addition, a

reminder prompt was activated after every fifth item. The effect of the intervention can be compared in the test one week later.

In the following chapters, the research questions of the three empirical studies in this dissertation are elucidated in more detail.

References

Adesope, O. O., Trevisan, D. A., & Sundararajan, N. (2017). Rethinking the use of tests: A meta-analysis of practice testing. *Review of Educational Research, 87*(3), 659-701. https://doi.org/10.3102/0034654316689306

Agarwal, P. K., Nunes, L. D., & Blunt, J. R. (2021). Retrieval practice consistently benefits student learning: A systematic review of applied research in schools and classrooms. *Educational Psychology Review*, *33*(4), 1409-1453. https://doi.org/10.1007/s10648-021-09595-9

Bannert, Maria (2009): Promoting Self-Regulated Learning Through Prompts. *Zeitschrift für Pädagogische Psychologie 23*(2), 139-145. https://doi.org/10.1024/1010-0652.23.2.139

Barenberg, J., & Dutke, S. (2019). Testing and metacognition: Retrieval practise effects on metacognitive monitoring in learning from text. *Memory, 27*(3), 269-279. https://doi.org/10.1080/09658211.2018.1506481

Biwer, F., Egbrink, M. G. A. o., Aalten, P., & de Bruin, A. B. H. (2020). Fostering effective learning strategies in higher education—A mixed-methods study.*Journal of Applied Research in Memory and Cognition, 9*(2), 186–203. https://doi.org/10.1016/j.jarmac.2020.03.004

Bjork, E. L., & Bjork, R. A. (2011). Making things hard on yourself, but in a good way: Creating desirable difficulties to enhance learning. In M. A. Gernsbacher, R. W. Pew, L. M. Hough, & J. R. Pomerantz (Eds.), FABBS Foundation, psychology and the real

world: Essays illustrating fundamental contributions to society (pp. 56–64). Worth
Publishers.

Bjork, R. A., Dunlosky, J., & Kornell, N. (2013). Self-regulated learning: Beliefs, techniques,
and illusions. *Annual Review of Psychology*, *64*, 417-444. https://doi.org/annurev-
psych-113011-143823

Butler, A. C., & Roediger, H. L. (2008). Feedback enhances the positive effects and reduces
the negative effects of multiple-choice testing. *Memory & Cognition, 36*(3), 604-616.
https://doi.org/10.3758/MC.36.3.604

Carpenter, S. K. (2009). Cue strength as a moderator of the testing effect: The benefits of
elaborative retrieval. *Journal of Experimental Psychology: Learning, Memory, and
Cognition, 35*(6), 1563–1569. https://doi.org/10.1037/a0017021

Chan, J. C. K., McDermott, K. B., & Roediger, H. L. III. (2006). Retrieval-induced
facilitation: Initially nontested material can benefit from prior testing of related
material. *Journal of Experimental Psychology: General, 135*(4), 553–571.
https://doi.org/10.1037/0096-3445.135.4.553

Foerst, N. M., Klug, J., Jöstl, G., Spiel, C., & Schober, B. (2017). Knowledge vs. Action:
Discrepancies in University Students' Knowledge about and Self-Reported Use of
Self-Regulated Learning Strategies. *Frontiers in Psychology, 8,* 1288.
https://doi.org/10.3389/fpsyg.2017.01288

Glenberg, A.M., Wilkinson, A.C. & Epstein, W. (1982). The illusion of knowing: Failure in
the self-assessment of comprehension. *Memory & Cognition 10*(6), 597–602.
https://doi.org/10.3758/BF03202442

Greving, S., Lenhard, W., & Richter, T. (2022). The testing effect in university teaching:
Using multiple-choice testing to promote retention of highly retrievable information.
Teaching of Psychology. Advance online publication.
https://doi.org/10.1177/00986283211061204

Greving, S., & Richter, T. (2018). Examining the testing effect in university teaching: Retrievability and question format matter. *Frontiers in Psychology*, *9,* 2412. https://doi.org/10.3389/fpsyg.2018.02412

Hinze, S. R., & Rapp, D. N. (2014). Retrieval (sometimes) enhances learning: Performance pressure reduces the benefits of retrieval practice. *Applied Cognitive Psychology, 28*(4), 597–606. https://doi.org/10.1002/acp.3032

Hinze, S. R., Wiley, J. & Pellegrino, J. W. (2013). The importance of constructive comprehension processes in learning from tests. *Journal of Memory and Language, 69*(2), 151–164. https://doi.org/10.1016/j.jml.2013.03.002

Jonsson, B., Wiklund-Hörnqvist, C., Stenlund, T., Andersson, M., & Nyberg, L. (2021). A learning method for all: The testing effect is independent of cognitive ability. *Journal of Educational Psychology*, *113*(5), 972–985. https://doi.org/10.1037/edu0000627

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

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., Klein, P. J., & Rawson, K. A. (2015). Retrieval attempts enhance learning, but retrieval success (versus failure) does not matter. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*(1), 283. https://doi.org/10.1037/a0037850

Kornell, N., & Vaughn, K. E. (2016). How retrieval attempts affect learning: A review and synthesis. In B. H. Ross (Ed.), *The psychology of learning and motivation* (pp. 183–215). Elsevier Academic Press. http://dx.doi.org/10.1016/bs.plm.2016.03.003

Kubik, V., Gaschler, R., & Hausman, H. (2021). Enhancing student learning in research and educational practice: The power of retrieval practice and feedback. *Psychology Learning and Teaching, 20*(1), 1-20. https://doi.org/10.1177/1475725720976462

Chapter I

Kuhbandner, C., & Emmerdinger, K. J. (2019). Do students really prefer repeated rereading over testing when studying textbooks? A reexamination. *Memory, 27*(7), 952–961. https://doi.org/10.1080/09658211.2019.1610177

McDaniel, M. A., & Little, J. L. (2019). Multiple-choice and short-answer quizzing on equal footing in the classroom: Potential indirect effects of testing. In J. Dunlosky & K. A. Rawson (Eds.), *The Cambridge handbook of cognition and education* (pp. 480–499). Cambridge University Press. https://doi.org/10.1017/9781108235631.020

Morris, C. D., Bransford, J. D., & Franks, J. J. (1977). Levels of processing versus transfer appropriate processing. *Journal of Verbal Learning and Verbal Behavior, 16*(5), 519–533. https://doi.org/10.1016/S0022-5371(77)80016-9

Pan, S. C. & Rickard, T. C. (2018). Transfer of test-enhanced learning: Meta-analytic review and synthesis. *Psychological Bulletin, 144*(7), 710–756. https://doi.org/10.1037/bul0000151

Pyc, M. A., & Rawson, K. A. (2010). Why testing improves memory: Mediator effectiveness hypothesis. *Science*, *330*(6002), 335. https://doi.org/10.1126/science.1191465

Ramsden, P. (2003). *Learning to teach in higher education*. Routledge.

Richter, T., Berger, R., Ebersbach, M., Eitel, A., Endres, T., Borromeo Ferri, R. et al. (2022). How to promote lasting learning in schools: Theoretical approaches and an agenda for research. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie/German Journal of Developmental Psychology and Educational Psychology, 54*(4), 135–141. https://doi.org/10.1026/0049-8637/a000258

Roediger, H. L. III, & Butler, A. C. (2011). The critical role of retrieval practice in long-term retention. *Trends in Cognitive Sciences, 15*(1), 20–27. https://doi.org/10.1016/j.tics.2010.09.003

Roediger III, H. L., & Karpicke, J. D. (2006). Test-enhanced learning: Taking memory tests improves long-term retention. *Psychological Science*, *17*(3), 249–255. https://doi.org/10.1111/j.1467-9280.2006.01693.x

Roediger, H. L. III, & Marsh, E. J. (2005). The Positive and Negative Consequences of Multiple-Choice Testing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*(5), 1155–1159. https://doi.org/10.1037/0278-7393.31.5.1155

Rowland, C. A. (2014). The effect of testing versus restudy on retention: A meta-analytic review of the testing effect. *Psychological Bulletin*, *140*(6), 1432-1436. https://doi.org/10.1037/a0037559

Schneider, W., Tibken, C. & Richter, T. (2022). The development of metacognitive knowledge from childhood to young adulthood: Major trends and educational implications. In J. Lockman (Ed.), *Advances in child development and behavior* (Vol. 63, pp. 273–307). Cambridge, MA: Elsevier. https://doi.org/10.1016/bs.acdb.2022.04.006

Schwieren, J., Barenberg, J., & Dutke, S. (2017). The testing effect in the psychology classroom: A meta-analytic perspective. *Psychology Learning & Teaching*, *16*(2), 179–196. https://doi.org/10.1177/1475725717695149

Soderstrom, N. C., & Bjork, R. A. (2014). Testing facilitates the regulation of subsequent study time. *Journal of Memory and Language, 73*, 99-115. https://doi.org/10.1016/j.jml.2014.03.003

Taraban, R., Maki, W. S. & Rynearson, K. (1999). Measuring study time distributions: Implications for designing computer-based courses. *Behavior Research Methods, Instruments & Computers, 31*(2), 263–269. https://doi.org/10.3758/BF03207718

Van Gog, T., & Sweller, J. (2015). Not new, but nearly forgotten: The testing effect decreases or even disappears as the complexity of learning materials increases. *Educational Psychology Review, 27*(2), 247-264. https://doi.org/10.1007/s10648-015-9310-x

Chapter I

Wong, J., Baars, M., Davis, D., Van Der Zee, T., Houben, G. J., & Paas, F. (2019).

Supporting self-regulated learning in online learning environments and MOOCs: A

systematic review. *International Journal of Human–Computer Interaction, 35*(4-5),

356-373. https://doi.org/10.1080/10447318.2018.1543084

Yang, C., Luo, L., Vadillo, M. A., Yu, R., & Shanks, D. R. (2021). Testing (quizzing) boosts

classroom learning: A systematic and meta-analytic review. *Psychological Bulletin,*

147(4), 399–435. https://doi.org/10.1037/bul0000309

Chapter II

The Testing Effect in the Lecture Hall: Does it Depend on Learner Prerequisites?

Study 1

A version of this chapter was published as:

The Testing Effect in the Lecture Hall: Does it Depend on Learner Prerequisites?

Julia Glaser & Tobias Richter

Abstract. The benefits of practice testing for long-term learning are well established in many contexts. However, little is known about learner characteristics that might moderate its effectiveness. The effects of practice tests might depend on individual prerequisites for learning, especially in real-world educational settings. We explored whether the effects of practice testing in a regular university lecture would depend on cognitive (e.g., prior knowledge), motivational (e.g., learning motivation), or emotional (test anxiety) dispositions. We implemented an experimental intervention design in psychology courses for teacher students ($N = 208$). One week before the lecture, focal learner characteristics were assessed. Immediately after the lecture, participants completed an online review session with short-answer questions (practice testing with corrective feedback) or summarizing statements (restudy), alternating within each participant. One week later, retention of learning contents was assessed with a criterial test containing short-answer and multiple-choice questions. A testing effect emerged ($\eta_p^2 = .07$), with better retention for the tested compared with the restudied contents. Some learner characteristics affected learning outcomes, but no interactions with testing vs. restudy occurred. These results suggest that the testing effect in the university classroom is a robust phenomenon that benefits learning irrespective of primary individual learning prerequisites.

The Testing Effect in the Lecture Hall: Does it Depend on Learner Prerequisites?

Learning in a university context is demanding. Students are required to understand and encode large amounts of information in a short period of time, retain the knowledge until the exam and hopefully after, and retrieve and apply the knowledge when it is needed. Learners receive little direct support and guidance in their learning activities, especially in larger courses. One promising way to foster long-term retention even in large-scale university courses is practice testing, that is, practicing the retrieval of learned information (Roediger & Butler, 2011). In typical studies on retrieval practice, a group of students that receive practice tests are compared to a group of students that are given the opportunity to restudy the material, a learning activity that, according to a survey by Karpicke et al. (2009), matches students' typical study behavior when they study on their own (but see Kuhbander & Emmerdinger, 2019, who report results showing that students use rereading purposefully for difficult passages and often use a combination of different strategies, including testing in later stages of learning). For example, in a classical study, Roediger and Karpicke (2006) asked students to read an expository text that was divided into subsections. After each section, participants were either given the opportunity to reread the text (restudy group) or were asked to recall as much information as possible and write it down (testing group). In a test immediately after learning, the restudy group performed slightly better than the testing group (81% vs. 75% retention), but in subsequent tests two days later and one week later, students who were asked to recall the learned information as part of their study activities remembered more information than students who had reread the text sections (two days: 68 % vs. 54%; one week: 56% vs. 42%).

The testing effect has been demonstrated by a wide variety of studies using different materials, periods of learning, and laboratory and field contexts. Meta-analytic results suggest that the testing effect is a robust phenomenon with medium to large effect sizes ($g = 0.51$,

31

Adesope et al., 2017; $g = 0.50$ Rowland, 2014), also in real-world educational settings ($g = 0.33$, Yang et al., 2021) and psychology classes ($d = 0.73$, Schwieren, Barenberg, & Dutke, 2017; all effect sizes for comparisons of testing vs. restudying). Likewise, the review by Agarwal et al. (2021), found that 57% out of 49 studies in real-world educational settings had medium or large effect sizes. These reviews also revealed important moderators of the testing effect. For example, retrievability is an important factor that can affect test performance. For retrieval attempts without feedback, high retrievability is a precondition of direct positive effects of testing because only successful retrieval attempts elicit retrieval practice (Greving, Lenhard & Richter, 2022; Greving & Richter, 2018). For retrieval attempts with feedback, retrievability seems to be less important (Kornell & Vaughn, 2016; Kornell et al., 2015). Feedback, specifically corrective feedback, might increase the testing effect by consolidating correct responses and alerting students to knowledge gaps and misconceptions, which might promote indirect positive effects of testing. Finally, testing seems to be superior to restudying especially when learning outcomes are assessed at longer retention intervals of one week or more compared to shortly after learning (Rowland, 2014).

Compared to other types of conditions, learner characteristics have received relatively little attention as potential moderators of the testing effect. Learner characteristics represent one theoretically important class of potential moderators of the testing effect, besides materials, the retrieval task, and contextual variables (Kubik, Gaschler, & Hausman, 2021). A number of learner characteristics from cognitive variables (e.g., prior knowledge) to noncognitive variables (e.g., motivational and emotional dispositions) might theoretically affect the effectiveness of practice testing, especially in self-regulated learning when learners' degrees of freedom are high. The present study examined the role of a range of learner characteristics for the testing effect in large psychology courses at the university, which is geared towards teacher students who typically exhibit a large variability of learner characteristics.

In the following section, we will briefly review potential mechanisms underlying the effects of practice tests on learning. This review provides the basis for a discussion of how various cognitive, motivational, and emotional learner dispositions might interact with practice testing during learning.

Cognitive Mechanisms Underlying the Testing Effect

Practice tests might benefit learning in two ways. First, practice tests have been shown to directly improve later retrieval of knowledge from long-term memory (*direct testing effect*). According to the elaborative retrieval hypothesis (Carpenter, 2009), retrieving the response to a question in the practice test leads to the activation of related information, which strengthens the integration of the tested information in long-term memory. A similar account is the mediator effectiveness hypothesis (Pyc & Rawson, 2010), which proposes that practice testing creates better mediators, that is, information linked to the target information, which can be activated by a cue. A third major theoretical account of the testing effect is transfer-appropriate processing (Morris, Bransford, & Franks, 1977), meaning that testing creates conditions during learning that are more similar to the later criterial test, compared to restudying. This similarity of learning and the test, in turn, is assumed to benefit performance in the criterial test. In sum, direct testing effects are assumed to occur because practice testing activates elaborative processes during learning, creating memory structures that foster the retrieval of information in a later test of the learning outcomes.

Second, practice testing might also benefit learning via an indirect route (*indirect* or *mediated testing effect*), especially when feedback is provided. In this scenario, practice tests inform test takers about their current performance level, preventing the illusion of knowing (Glenberg et al., 1982), which may occur during restudy. Accordingly, practice testing may improve metacognitive processes, such as judgements of learning, and also indirectly improve the regulation of subsequent self-directed learning activities (e.g., Barenberg & Dutke, 2019; Soderstrom & Bjork, 2014).

Learner Characteristics Potentially Moderating the Testing Effect

Even though positive effects of practice tests seem to be robust, their magnitude may still depend on moderating or boundary conditions (McDaniel & Butler, 2011; see also the meta-analysis by Rowland, 2014). Research exists on the moderating role of characteristics in educational settings such as the provision of feedback (e.g., Butler & Roediger, 2008), the type of questions used during practice testing (multiple choice vs. short-answer questions, e.g., Greving & Richter, 2018; McDaniel & Little, 2019), and the learning materials (e.g., its complexity, van Gogh & Sweller, 2015). In contrast, research on the moderating role of learner characteristics is still scarce. Such moderating effects of learner characteristics are especially likely when instructional measures are implemented in a self-regulated learning setting, which requires that learners monitor and regulate their learning activities (Alexander & Greene, 2017). Based on theoretical considerations, we selected a number of cognitive, motivational, and emotional dispositions for closer investigation in the present study. The possible role of these learner characteristics for the effectiveness of the testing effect and the relevant empirical research is discussed next.

Cognitive Dispositions. Practice testing requires learners to invest mental effort to be effective, especially to elicit direct testing effects. This assumption is inherent, for example, in the elaborative retrieval hypothesis or the mediator effectiveness hypothesis and it follows from the general characterized. Therefore, learners equipped with better cognitive abilities might benefit more from practice testing. However, Jonsson et al. (2020) found in two experiments that testing effects for learning Swedish-Swahili word pairs were independent of cognitive ability (assessed with a wide range of basic cognitive ability measures, from fluid intelligence to working memory). Similar results have been reported by several other studies (Bertilsson et al., 2017; Pan et al., 2015; Wiklund-Hörnquist et al., 2014), and other studies even suggest that learners with abilities in the lower range might profit more from practice testing than their high-ability counterparts (Agarwal et al., 2017; Brewer & Unsworth, 2012)

Apparently, the cognitive processes conducive to learning that are triggered by practice tests place no demands on general cognitive abilities.

Another cognitive learner characteristic that might affect the magnitude of the testing effect is domain-specific prior knowledge. Prior knowledge is relevant for the comprehension of learning materials, and it affects the retrievability of knowledge during practice testing. The complex learning materials typically used in learning settings at the university need to be comprehended to encode and integrate the to-be-learned information in existing knowledge structures, which affects retrievability. Retrievability, in turn, is important for direct testing effects to occur in the lab (Rowland, 2014) and in university classrooms (Greving & Richter, 2018; Greving et al., 2020) because only knowledge that is successfully retrieved can be consolidated by practice testing (especially if no feedback is given). Despite the theoretical importance of prior knowledge, research systematically examining its moderating role is scarce. For example, Carpenter et al. (2016) who showed that high-achieving students benefited more from retrieval practice than do middle- or low-achievers, which they explain by the fact that high-achievers have more prior knowledge and thus can process new knowledge more deeply, fostering elaborative retrieval (for similar results, see Francis et al., 2020; Marsh et al., 2009). In contrast, some studies that have employed a no-treatment control group (instead of the typical restudy control group) found learners with low prior knowledge to benefit more from retrieval practice than high prior knowledge learners (e.g., Cogliano et al., 2019). However, it seems questionable whether a no-treatment control group is an adequate point of comparison for assessing the effectiveness of retrieval practice and the role of prior knowledge. Finally, the only study in which prior knowledge was manipulated experimentally found no moderating effect of prior knowledge on the testing effect (Buchin & Mulligan, in press; for null results using a pre-experimental measure of prior knowledge, see also Xiaofeng et al., 2016).

Motivational Dispositions. Practice testing is a desirable difficulty (Bjork & Bjork, 2011) that makes learning more effective (especially long-term memory) but also subjectively more difficult. By judging retrieval practice as more difficult, learners misinterpret it as ineffective for learning and therefore use it less often in their self-regulated learning activities (misinterpreted-effort hypothesis; Kirk-Johnson et al., 2019). And, in contrast to restudying, learners often make errors and become aware of knowledge gaps and misconceptions when taking practice tests, which can be aversive for learners. Therefore, motivational dispositions and attitudes towards errors described in the literature might be crucial for whether learners can constructively use the negative feedback that inevitably occurs during practice testing. First, *achievement goals* are typically distinguished into mastery and performance goals (e.g., Dweck, 1986; Harackiewicz, Barron, & Elliot 1998). Mastery goals involve engaging in deep learning to understand the learning materials. Learners with strong mastery goals should be motivated to actively seek and use the feedback provided from practice tests. In contrast, performance goals might involve the desire to demonstrate ability to others and to receive extrinsic rewards for the learning activities. Strong performance goals might be associated with a tendency to avoid the potentially negative feedback provided by practice testing. In line with this reasoning, Weissgerber et al. (2016) found evidence that achievement goals are differentially related to the self-regulated use of learning strategies such as practice testing. In a similar vein, Yan et al. (2014) found a positive relationship between the belief that one's intelligence can be increased through effort and the use of beneficial learning strategies such as self-testing.

Second, in a similar vein, learners' *error orientation* (Rybowiak et al., 1999) might also affect how effectively they exploit the potential benefits of practice testing. Introduced originally in work and organizational psychology, error orientation refers to an individual's attitudes towards errors occurring at work, comprising different facets such as learning from errors, error strain, or thinking about errors. Given that errors can hardly be avoided during

practice testing and that especially mediated effects of practice testing depend on the ability to use errors constructively (that means recognizing and seizing the opportunity to learn from errors), learners' error orientation might affect the benefits they can gain from practice tests.

Third, learners *need for cognition* (NFC) might play a role in how effectively they use practice testing. The NFC denotes "an individual's tendency to engage in and enjoy effortful cognitive processing" (Cacioppo et al., 1984, p. 306). Schindler et al. (2019) demonstrated that NFC moderates the generation effect. A generation effect occurs when self-generated information is remembered better compared to information that is read. This effect is related to the testing effect and may also be categorized as a desirable difficulty in learning. Schindler et al. (2019) found that students with lower NFC benefitted more from generation than students with higher NFC. The authors explained this finding by arguing that individuals with higher NFC are already inclined to engage in generative learning activities without specific instructions, whereas individuals with lower NFC need the explicit instruction to generate information to include generation in their learning process. Thus, it seems plausible that individuals with higher cognition needs already use good learning strategies and therefore do not benefit as much from guided practice as individuals with lower cognition needs. But, only one study (Bertilsson et al. 2021) has examined the moderating role of NFC on the testing effect and found no evidence for such a role. However, Bertilsson et al. used relatively simple materials (vocabulary learning), and they suggested that a moderator effect of NFC might occur for more complex materials.

Emotional Dispositions. Affective dispositions are closely linked to motivational dispositions and might also affect whether learners use practice testing effectively. Practice tests create a test situation that includes the possibility of not knowing an answer and receiving negative feedback. This feedback might be more aversive for some learners than for others, depending on certain emotional dispositions. For example, Ramsden (2003) showed that students scoring high on fear of failure use challenging learning strategies less frequently,

possibly to avoid negative emotions as a result of mistakes. A similar reasoning applies to the related construct of test anxiety. For example, Weissgerber and Reinhard (2018) found that students scoring high on test anxiety reported to use self-testing strategies less frequently. Text-anxious students might also use self-testing less effectively when confronted with practice tests because the worries and physiological tension that characterize test anxiety prevent them from engaging in the cognitive processes conducive to learning. Previous studies have shown that performance pressure, which likely triggers test anxiety, reduces the beneficial effects of practice testing (Hinze & Rapp, 2011). Tse and Pu (2012) used the Swahili vocabulary learning paradigm introduced by Roediger and Karpicke (2006) to examine how trait test anxiety and working memory together influence the testing effect. They found the testing effect to be decreased in learners high in test anxiety and low in working memory, which may be explained by the fact test anxiety, especially its worry component, increases working memory load.

Rationale of the Present Study

The available research raises the question of how individual differences in cognitive, motivational, and emotional dispositions affect the magnitude of the testing effect, when practice testing is implemented in a typical self-regulated learning setting in higher education. This study examined this question by incorporating practice testing in optional review activities offered to complement compulsory psychology lectures for university teacher students. Restudy was used as the control condition, in line with the main body of research on retrieval practice and given its prevalence in self-regulated learning activities of university students. To maximize experimental control and statistical power, practice testing vs. restudy was implemented within-subjects and within each lecture topic by using alternating questions and summarizing statements. Consistent with previous theory and research on retrieval practice, we hypothesized that a positive testing effect would occur and that this effect would be robust across different lecture topics.

Additionally, we examined for a range of individual differences in cognitive dispositions (prior knowledge, mean retrievability during practice tests), motivational dispositions (learning goals, need for cognition) and emotional dispositions (text anxiety) whether they would moderate the testing effect in the conditions described above. Given the paucity and overall inconclusive nature of research on the role of learner characteristics during practice testing in real-world educational settings, we pursued this investigation as exploratory research questions.

Method

Participants

Participants were 208 students enrolled in the teacher-training program at the University of Würzburg. Seventy-six students studying to become teachers for elementary school (Grundschule), 6 for non-academic track middle-school teachers (Mittelschule), 12 for secondary school (Realschule), 29 for high school (Gymnasium), 63 for special education, and 5 to become speech therapists; 17 did not specify their field of study.

Participants were recruited from three different courses (Behavioral and Learning Disorders in Childhood and Youth, Development in Childhood and Youth, Psychology of Learning), all of which are part of the mandatory psychology curriculum for prospective teachers. They could choose to participate in one of five lecture topics that were part of the three courses: Reading and Writing Problems, Attentional Deficit Hyperactivity Disorder, Development of Thinking, Development of Intelligence, and Avoidance Learning.

Of the 327 students who initially signed up for the study, only 208 completed all parts of the study. The other 119 students were excluded from the analysis because of incomplete data. Data for the study were collected in two waves in subsequent semesters (November/December 2020, $n = 78$, April to June 2021, $n = 130$). The distribution of participants between topics, waves, and parts is shown in Table I.1.

Table I.1

Number of Participants per Lecture Topics in Waves 1 and 2 and Attrition During the Experiment

Topic	Part 1		Part 2		Part 3		Valid (cases)
	Wave 1	Wave 2	Wave 1	Wave 2	Wave 1	Wave 2	
Reading and writing problems	18	108	17	91	15	82	87
Attentional Deficit Hyperacticity Disorder	24	20	19	17	18	17	31
Development of Thinking	27	35	25	18	24	17	37
Development of Intelligence	26	42	19	33	17	33	42
Avoidance Learning	20	6	12	5	10	4	11
Total	115	212[a]	92	164	84	153	
	327		256		237		208

Note. Wave 1 took place in the winter semester 2020/2021. Wave 2 took place in summer semester 2021. Part 1 included the prior knowledge test and assessments of other learner characteristics. Part 2 was the learning phase and Part 3 the criterial test. Valid cases correspond to participants with complete data that were included in the analyses.
[a] One person did not choose a topic.

Most participants were female (89.9%) and received course credit for their participation (94.7%). Participants not receiving course credit took part in the study voluntarily without any compensation. Participants' age ranged from 18 to 41 years ($M = 20.99$, $SD = 3.45$). Only five participants reported another language than German as their first language.

Experimental Study Materials

We analyzed the contents of the lecture session for each topic and extracted 25 information units for each topic. For each of the five lecture sessions, 15 of the 25 information units were selected and presented either as a statement (restudy) or as a short answer question (practice testing). The statements were merely presented, and participants decided how long they wished to restudy them. The short answer questions were posed without a time limit. They required participants to answer with one sentence or to provide keywords. After

providing their answer, participants were shown the correct answer as feedback. A sample

short answer question and corresponding statement read as follows:

Short answer Question: In which order do the stages of writing acquisition occur?

Statement: Stages of writing acquisition in order:
- logographic stage
- alphabetic stage
- orthographic stage

The experimental study materials are available in the Appendix.

Measures

Criterial Test

The dependent variable was the performance in criterial tests with 20 questions for

each topic. Ten questions referred to the 15 information units presented in the experimental

study materials. The other ten questions were drawn from the remaining 25 information units

of each of the focal lecture sessions and served as filler items. The criterial tests for three of

the five topics (Attention Deficit Hyperactivity Disorder, Development of Intelligence, and

Avoidance Learning) consisted of multiple-choice questions with four response options, each

of which could be correct or incorrect. Partial credit (.25) was given for each response option

that was correctly ticked or correctly not ticked. The criterial tests for two of the five topics

(Reading and Writing Problems, Development of Thinking) consisted of short answer

questions. These were items also scored according to a partial credit scheme (i.e., 0-1 points

in intervals of .25). Again, two raters independently coded the responses for 320 questions (20

questions from 16 participants) to estimate inter-rater reliability of the short-answer questions.

Cohen's κ was .859 ($SE = .025$), indicating almost perfect agreement according to Landis and

Koch (1977). Again, given the high inter-rater reliability, the remaining answers to the short

answer questions were coded by only one coder.

Learner Characteristics

Table I.2 shows all descriptive measures of the assessed learner characteristics.

Table I.2

Descriptive Statistics of Learner Characteristics

	Study Sample				Reference Sample		
	M	*SD*	*min*	*max*	*N*	*M*	*SD*
Criterial Test	.68	0.19	.00	.99	-	-	-
Retrievability	.70	0.19	.00	1.00	-	-	-
Prior Knowledge	.71	0.09	.47	.98	-	-	-
Need for Cognition	4.57	0.78	2.13	6.38	91	4.83	0.78
Learning Motivation (SELLMO)							
Learning goals	34.25	4.07	11.00	40.00	1926	31.23	5.14
Performance goals - approach	21.92	4.53	7.00	35.00	1926	23.79	5.08
Performance goals – avoidance	20.84	6.56	8.00	38.00	1926	21.45	6.43
Avoidance of Work	17.45	5.13	8.00	36.00	1926	21.66	6.69
Test Anxiety (TAI-G)							
Interference	5.91	2.45	3.00	12.00	709	5.65	2.10
Worry	15.21	3.73	5.00	20.00	709	12.95	3.60
Arousal	9.44	3.30	4.00	16.00	709	8.45	2.95
Lack of Confidence	7.19	2.15	3.00	12.00	709	7.35	1.85
TAIG Total Value	37.75	9.37	19.00	66.0	709	34.45	8.05
Error Orientation (EOQ)							
Risk taking	3.56	0.76	1.75	5.00	478	3.84	0.73
Error strain	3.19	0.87	1.20	5.00	478	2.35	0.72
Error anticipation	3.27	0.62	1.60	5.00	478	2.23	0.65
Covering up errors	2.11	0.59	1.00	3.83	478	2.00	0.71
Error communication	3.89	0.60	1.75	5.00	478	3.41	0.67
Thinking about errors	3.85	0.70	2.00	5.00	478	3.37	0.70
Error competence	3.72	0.55	2.25	5.00	478	4.07	0.49
Learning from errors	3.70	0.62	1.75	5.00	478	4.11	0.53

Note. Retrievability: individual performance in the practice test. For all measures of the study sample N = 208. Data for the reference samples was taken form the following sources: Need for Cognition (Schindler et al., 2019), Learning Motivation (Spinath et al., 2002), Test Anxiety (Hodapp, 1991), Error Orientation (Rybowiak et al., 1999).

Retrievability. Retrievability was defined as the individual performance in the short-answer questions in the retrieval practice condition. Participants' answers were scored with partial credit, depending on how much of the answer was correct. The minimum score was 0 for incorrect answers, .25, .50, or .75 were scored for partially correct/complete answers, and the maximum 1 for fully correct/complete answers. For example, participants received 1 point if the answer to the example question was "logographic stage, alphabetic stage, orthographic stage." If they wrote, for example, "alphabetic stage, orthographic stage" they received .50 points because the answer was correct but incomplete. For incorrect or grossly imprecise

answers, such as "learning letters, learning sounds, learning spelling", participants received 0 points. Two raters independently coded <u>300</u> responses (7 or 8 responses from 40 participants). The inter-rater reliability (Cohen's κ) was .887 (*SE* = .024), which indicates almost perfect agreement (Landis & Koch, 1977). Given the high inter-rater reliability, the remaining answers were coded by only one coder.

Prior Knowledge. For each of the five topics, we constructed a prior knowledge test. Based on 15 information units, multiple-choice questions with four alternatives were constructed, each of which could be correct or false. The prior knowledge questions tapped into knowledge that was relevant for understanding the lecture topics. Their content did not overlap with the statements and short answer-questions used in the restudy/practice testing phase. The following is an example of a prior knowledge question for the topic reading and writing problems:

> What does the term "precursor skills" refer to?
> a) Skills that require a certain time before they can be acquired or performed well.
> b) Skills that need to be acquired before another skill.
> c) An example of precursor skills to literacy acquisition might be nonsense rhymes.
> d) An example of a precursor skill might be writing since it takes some time to master the skill fluently.

Need for Cognition. Need for Cognition was assessed with the short form of the Need for Cognition Scale (Bless et al., 1994). The scale contains 16 statements assessing inclination and enjoyment to engage in effortful thinking. The items are rated on a 7-point Likert scale (sample item: "I really enjoy a task that involves coming up with new solutions to problems."). The internal consistency of the questionnaire was satisfactory (Cronbach's α = .840, *N* = 208 participants).

Learning and Performance Motivation. Participants' learning and performance motivation was assessed with the SELLMO-ST inventory (Spinath et al., 2002) that contains 31 self-describing statements about learning motivation rated on a 5-point scale. The items are sorted into four subscales: Learning Goals, Approach Performance Goals, Avoidance

Performance Goals, and Tendency to Avoid Work. The internal consistencies (Cronbach's α) for the four subscales ($N = 208$) were satisfactory, with .811 for the subscale Learning Goals (8 items), .772 for Approach Performance Goals (7 items), .877 for Avoidance Performance Goals (8 items), and .838 for Tendency to Avoid Work (8 items).

Test Anxiety. Test anxiety was measured by the short version of the TAI-G (Spinath et al., 2002; Wacker et al., 2008), which contains 15 items referring to feelings and thoughts concerning test situations rated on a 5-point scale. The internal consistency of the scale was excellent (Cronbach's α = .913, $N = 208$; sample item: "I think about what happens if I do poorly").

Error Orientation. Facets of error orientation were assessed with the Error Orientation Questionnaire (German version, Rybowiak et al., 1999), which contains 33 items expressing different attitudes towards errors rated on a 5-point scale. The internal consistencies (Cronbach's α) of the seven subscales were .624 for Error Competence (4 items, sample item: "When I have made a mistake, I immediately know how to correct it"), .683 for Learning From Errors (4 items, "My mistakes help me to improve my work"), .790 for Error Risk Taking (4 items, "I would prefer to err than to do nothing at all"), .845 for Error Strain (5 items, "I feel embarrassed when I make an error"), .674 for Error Anticipation (5 items, "Whenever I start a piece of work, I am aware that mistakes occur"), .744 for Covering up Errors (6 items, "I'd rather keep my mistakes to myself"), .618 for Error Communication (4 items, "When I have done something wrong, I ask others how to do it better"), and .824 for Thinking About Errors (5 items, "When a mistake occurs, I analyze it thoroughly") ($N = 208$).

Procedure

The study was conducted online via the online survey system Unipark (www.unipark.com). Figure I.1 provides an overview of the general procedure. Participants selected one specific lecture from the course that they attended and registered via an online university platform, after which they received a link to the first part of the experiment. The

first part of the experiment included the prior knowledge test on the chosen topic, the

assessment of other learner characteristics, and a questionnaire about basic demographic

information.

Figure I.1

Design and General Procedure of the Experiment
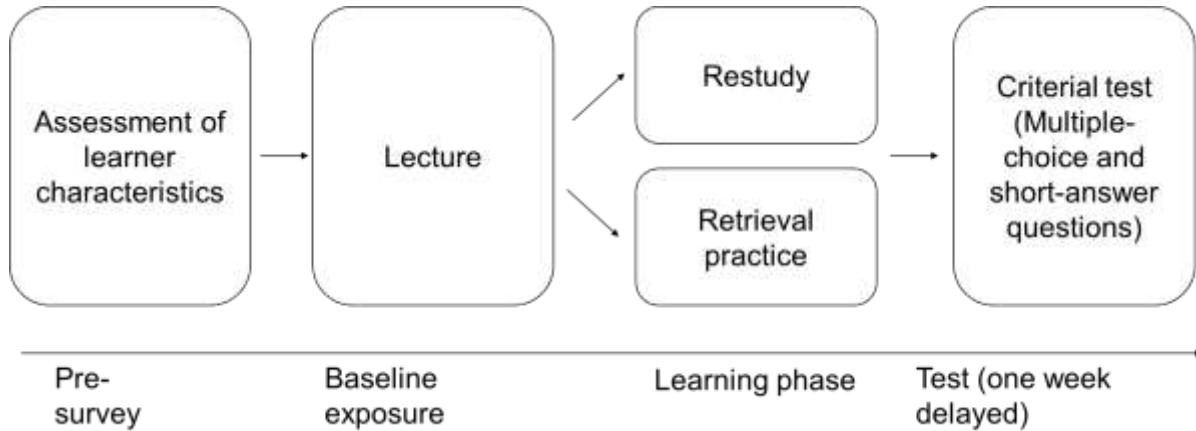
Note. Participants signed up for participation for one lecture topic. They then received a pre-survey (online) for assessing learner characteristics (potential moderators of the testing effect), followed 2 to 14 days later by the lecture as the baseline exposure to the learning materials and a learning phase with retrieval practice and restudy (varied within-subjects). One week later the criterial test was provided with multiple-choice or short-answer questions, depending on topic.

Between 2 and 14 days after completing the first part, participants attended the lecture

on the selected topic. They received an email with a link to the second part of the experiment

and were asked to complete this part immediately after the lecture. About one third (34.0%) of

the participants completed it within 1 hour after the lecture, 21.5% within one to 4 hours,

11.3% within four to 12 hours after the lecture, 10.9% on the same day, and 22.3% more than

24 hours later. The second part included the experimental intervention (retrieval practice vs.

restudy). Participants answered seven or eight short-answer questions and read seven or eight

statements about the lecture content (altogether 15 items). After completion of the test,

participants were asked to report on the extent that they were distracted during the processing

of the test and whether they used additional study materials. Most of participants reported a

low level of distraction during completion of the test ($M = 2.37$, $SD = 1.40$; range from 1 =

not at all to 7 = strongly distracted), and 88.5% reported not using additional study materials to complete the task ($N = 208$).

The third part of the experiment was scheduled one week after the completion of the second part. Again, participants received a link to the third part via email. About two thirds of participants (66.8%) completed the third part exactly one week after the second part, 13.9% after 8 days, 5.3% after 9 days, and 5.3% after 10 or more days, and 8.7% earlier than 7 days ($N = 208$). In the third part, participants responded to the questions of the criterial test that included 20 multiple choice or 20 short-answer questions about the contents of the chosen lecture session. Participants reported a low level of distraction during the completion of the criterial test ($M = 2.09$, $SD = 1.32$; range from 1 = not at all to 7 = strongly distracted), and 94.7% reported that they had not used any study materials to complete the task ($N = 208$). Finally, feedback regarding the perceived comprehensibility and usefulness of the study was obtained from participants.

Design

We employed one-factorial designs with the factor learning condition (practice testing vs. restudy) varied within-subjects. The dependent variable was the performance in the critical test. Learning condition was varied within each review session so that each participant reviewed half of the contents as statements and the other half as short-answer questions plus feedback. The assignment of contents to the two learning conditions was counterbalanced across participants by using two different lists. Each participant was randomly assigned to one of the two lists.

Sensitivity

A sensitivity analysis performed with GPower 3 (Faul et al., 2009) revealed that given the sample size of $N = 208$, the design was sensitive enough to detect a main effect for learning condition with an effect size of $f = .102$ ($d = 0.204$) or higher, assuming a correlation of .59 between the levels of the factor learning condition (which corresponds to the actual

correlation in the present study), an α-level of .05 and a power (1-β) of .90. For the detection of an interaction effect of any of the learner characteristics with learning condition, the design was sensitive enough to detect an effect with an effect size of $f = .051$ ($d = 0.102$) or higher, again assuming an α-level of .05 and a power (1-β) of .90.

Availability of Data and Materials

The data files and analysis syntax underlying the analyses reported in this study are available in the Appendix and via Open Science Framework (https://osf.io/9hw2z/?view_only=e56cf9fcb7844469a1f05e7b3cbcc2d8). Materials are available from the authors upon request.

Results

We first estimated a linear model with learning condition as experimental factor (varied within-subjects) and the lecture topics (varied between-subjects) and their interactions. In a second step, we estimated a series of linear models that additionally included each of the learner characteristics as continuous predictor and its interaction with learning condition (Judd, Kenny & McClelland, 2001) to explore the role of learner characteristics. We performed separate analyses for each learner characteristic to maximize the chance to detect a moderator effect. In a final step, we estimated An α-level of .05 (two-tailed) was used for all statistical tests. We report partial η^2 as the effect size measure.

Main Effect of Testing

The model with learning conditions (with no learner characteristics) revealed better learning outcomes in the criterial test for testing ($M = .779$, $SE = .018$) compared to restudying ($M = .707$, $SE = .019$), $F(1, 203) = 15.91$, $p < .001$, $\eta_p^2 = .073$. Therefore, Hypothesis 1 that predicted a positive testing effect was supported. The interaction of testing and lecture topics was not significant, $F(4, 203) = 0.84$, $p = .499$, $\eta_p^2 = .016$, suggesting that

the testing effect was independent of materials and contents. Figure I.2 displays the average

testing effect across topics and within each topic.

Figure I.2

The Testing Effect Across and Within Lecture Topics
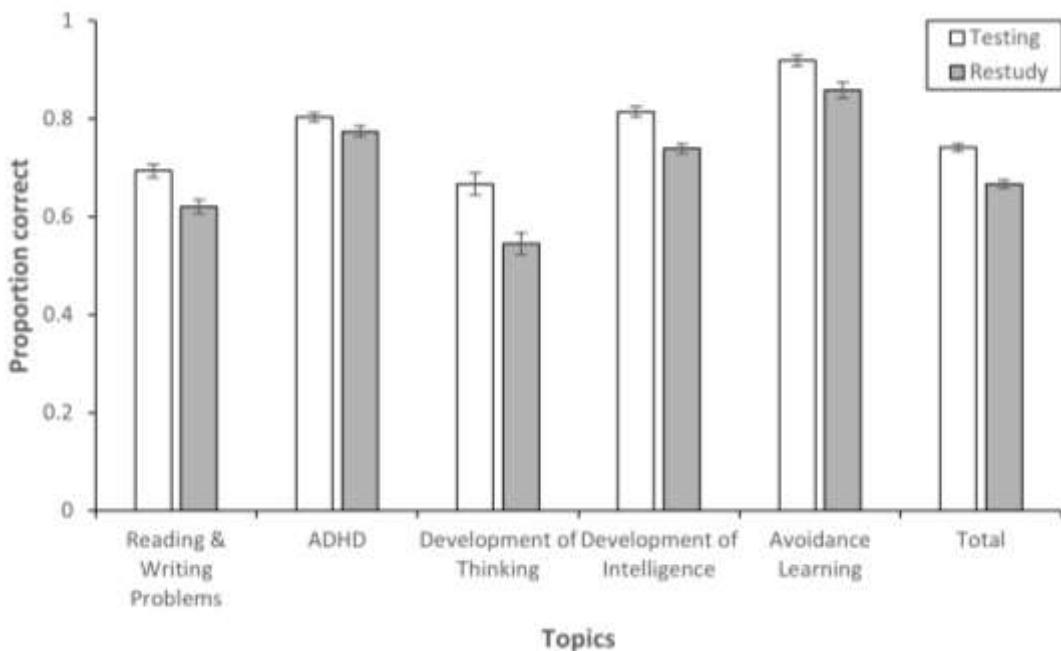
Note. Error bars represent the standard error of the mean. The *y*-axis shows the proportion of correct responses in the criterial test.

Main Effects of Learner Characteristics

The results for main effects of learner characteristics are shown in Table I.3.

Significant positive main effects were found for Learning from Errors, $F(1, 202) = 4.34$, $p =$

.038, $\eta_p^2 = .021$, and Thinking About Errors, $F(1, 202) = 11.76$, $p = .001$, $\eta_p^2 = .055$, from the

Error Orientation Questionnaire, meaning that students who understand mistakes as learning

opportunities and students who think about how errors occur and that they can be prevented in

the future achieved better learning outcomes than those scoring lower on these variables.

Moreover, the Tendency to Avoid Work scale from the SELLMO-ST showed a significant

negative effect, $F(1, 202) = 7.29$, $p = .008$, $\eta_p^2 = .035$, meaning that students scoring high on

this scale had lower learning outcomes. Finally, prior knowledge, $F(1, 202) = 16.78$, $p < .001$,

$\eta_p{}^2 = .077$, and retrievability, $F(1, 202) = 66.59$, $p < .001$, $\eta_p{}^2 = .248$, had positive main effects

on learning outcomes.

Table I.3

Main Effects of the Individual Characteristics

Factor	F (1, 202)	p	$\eta_p{}^2$	β
Retrievability	66.593	< .001	.248	.458
Prior knowledge	16.775	< .001	.077	.255
Need for Cognition	3.387	.067	.016	.118
Learning and performance motivation				
Learning goals	3.836	.052	.019	.126
Avoidance performance goals	.010	.921	.000	-.006
Approach performance goals	.191	.663	.001	-.028
Tendency to avoid work	7.294	.008	.035	-.172
Test Anxiety				
Interference	2.630	.106	.013	-.104
Worry	.197	.657	.001	.029
Arousal	.047	.828	.000	-.014
Lack of Confidence	1.414	.236	.007	-.077
Error orientation				
Learning from Errors	4.341	.038	.021	.133
Error strain	.206	.651	.001	-.029
Risk taking	1.124	.290	.006	.068
Error anticipation	.296	.587	.001	-.035
Covering up errors	3.022	.084	.015	-.112
Error communication	2.243	.136	.011	.096
Thinking about errors	11.762	.001	.055	.216

Note. Dependent variable was in all cases the criterial test performance. Retrievability: individual performance in the practice test.

Learner Characteristics as Moderators of the Testing Effect

      None of the interactions between learning condition and any of the learner

characteristics was significant (Table I.4). Thus, we found no evidence that the testing effect

depends on learner characteristics.

Table I.4

Interactions Between Learning Condition and Learner Characteristics With Simple Slopes in the Testing and Restudy Conditions

Learner Characteristic	$F(1, 202)$	p	$\eta^2_{partial}$	$\beta_{Testing}$	$\beta_{Restudy}$
Retrievability	1.442	.231	.007	.462	.359
Prior Knowledge	2.119	.147	.010	.188	.264
Need for Cognition	0.047	.829	.000	.116	.095
Learning and Performance Motivation					
Learning Goals	1.257	.264	.006	.152	.074
Avoidance Performance Goals	0.033	.855	.000	-.012	.000
Approach Performance Goals	0.007	.935	.000	-.023	-.027
Tendency to Avoid Work	0.001	.975	.000	-.157	-.149
Test Anxiety					
Interference	1.400	.238	.007	-.135	-.054
Worry	2.553	.112	.012	.078	-.024
Arousal	1.030	.311	.005	.020	-.043
Lack of Confidence	0.931	.336	.005	-.039	-.096
Error Orientation					
Learning From Errors	0.128	.721	.001	.135	.104
Error Strain	2.437	.120	.012	-.078	.022
Risk Taking	2.023	.156	.101	.109	.016
Error Anticipation	0.368	.545	.002	-.013	-.049
Covering up Errors	1.424	.234	.007	-.142	-.060
Error Communication	1.033	.311	.005	.122	.052
Thinking About Errors	0.016	.899	.000	.203	.183

Note. Dependent variable was in all cases the criterial test performance. Interactions were estimated in separate linear models for each learner characteristic. Retrievability: individual performance in the practice test. $\beta_{Testing}$: Simple slope in the testing condition (standardized coefficient). $\beta_{Restudy}$: Simple slope in the restudy condition (standardized coefficient).

Perceived Comprehensibility and Usefulness of the Study

Participants also provided feedback on the comprehensibility of the tasks and the usefulness of the study, that is, how useful participants judged the study for their personal learning. Most participants (72.6%) perceived the study as very comprehensible, 24.1% as acceptable, and 3.4% saw need for improvement ($N = 237$). Most participants also found the study to be useful (49.4% useful, 15.6% very useful), and 30.0% rated it as medium useful. Only 3.8% of participants found the study not very useful and 1.3% not useful at all.

Discussion

The present study examined the testing effect and potential moderating effects of cognitive, motivational, and emotional dispositions of learners in a higher education setting. We implemented a minimal intervention in three regular university courses, with multiple topics and a within-participant and within-topic variation of practice testing vs. restudying. The main finding was that a testing effect emerged across all topics, meaning that tested knowledge was retained better over one week than knowledge that was merely studied. No evidence was found for a moderating effect of any learner characteristic on the strength of the testing effect.

Our study contributes to the literature on real-life educational applications of retrieval practice by providing a rigorous experimental test of the testing effect with actual study materials in the university classroom. Even though a growing number of studies have examined the benefits of retrieval practice on the university or college level with actual course materials (Yang et al., 2021, included 225 effects in their meta-analysis), many of these studies are based on non-experimental designs, such as quasi-experimental designs with a nonrandom allocation of groups of students to learning conditions. Moreover, many studies in real-world educational settings (about 60% in the meta-analysis by Yang et al., 2021) have compared practice testing to no activity at all or restudying the learning materials, which is not a very informative comparison regarding the usefulness of practice testing in higher education. Our study used a within-subjects design to compare practice testing with short-answer questions and corrective feedback to restudying the same contents. The design was unique in that practice testing vs. restudying was varied also within topics. Therefore, any confounds of learning condition and topic can be ruled out. The fact that the testing effect emerged in all five lecture topics included in the experiment supports the generalizability of the positive testing effect found in our study. With a partial η^2 of .07 (which corresponds to f

51

= .27 or $d = 0.56$, Lenhard & Lenhard, 2016), the size of the testing effect was medium to large and in the order of magnitude of testing effects reported for laboratory studies. It is also comparable to or even exceeds the reported effect sizes for classroom studies (psychological classes: $d = 0.73$, Schwieren et al., 2017; classroom studies in general: $g = 0.33$, Yang et al., 2021). The size of the effect is remarkable considering the minimalistic nature of the intervention and the variety of influences that can contribute to the results and increase the error variance in a field experiment such as in the current study. One likely explanation for the testing effect, in line with major theoretical accounts of the direct testing effect (e.g., Bjork, 1975; Carpenter, 2009; Pyc & Rawson, 2009), is that practice testing increased cognitive effort and prompted participants to process the learned more deeply. This explanation is backed up by the observation that participants spend more time on practice test items ($M = 79.45$ seconds, $SD = 51.79$ per item) than on restudy items ($M = 13.91$ seconds, $SD = 19.43$). Feedback reading time was relatively short ($M = 12.10$ seconds, $SD = 11.98$), which indicates that participants did not put much emphasis on feedback (in line with the findings by Greving et al., 2022, on the use of feedback during practice testing in the psychology classroom).

In sum, our results provide strong support for the idea that practice testing with short-answer questions and corrective feedback can foster learning when it is implemented as an additional study opportunity in a university course. Given the minimalistic nature of the intervention, which was completed online by participants within a few minutes and in a single session, practice testing has also proven to be a didactic tool that can be implemented in a highly efficient way.

A novel research question addressed by the present experiment was to explore whether a broad array of cognitive, motivational, and emotional dispositions would moderate the testing effect. Despite the fact that each learner characteristic included in our study was selected based on theoretical considerations, no evidence for a moderating role of any learner characteristic was found. To be sure, the logic of the significance precludes the conclusion

from a null effect that the effect in question does not exist in the population. However, the study was sufficiently powered to find even a small moderator effect, and we used mostly established tests with a high reliability to assess the learner characteristics and conducted a number of significance tests without correcting the alpha-level for multiple tests. Moreover, several of the learner characteristics, such as retrievability and prior knowledge, exhibited main effects on learning that were well interpretable. Given these facts, it seems quite likely that the testing effect that we examined was indeed not moderated by any of the learner characteristics included in the experiment. This conclusion would be in line with findings by Jonsson et al. (2020), who concluded that the testing effect is "a learning method for all," which they found to be independent of cognitive ability. Our results broaden this conclusion to a wider range of cognitive, motivational, and emotional characteristics which have not been investigated in previous research on the testing effect. Theoretically, our findings suggest that the mechanisms described in major theories of learning, especially the increased cognitive effort induced by retrieval practice, do not depend on specific learner prerequisites. Apparently, when implemented in a typical psychology course, practice tests are per se strong enough to induce cognitive processes conducive to learning, such as elaborative retrieval, at least if short-answer questions are used and corrective feedback is given.

Despite its informative and clear results, our study also has limitations that need to be discussed. One limitation is that the composition of the sample might have been too homogeneous to elicit moderator effects of learner characteristics. All participants were students pursuing similar fields of study, which might have reduced the variance in those characteristics. Effect could emerge in a more heterogeneous sample with a greater variability of learner characteristics. As the comparison with other samples show (Table 2), some (although not all) of the learner characteristics, most notable the facets of test anxiety, had considerably lower standard deviations in the present sample than in the comparison sample. A similar argument can be made about the sample of items selected for the practice tests and

the criterial tests. The lecture topics were not very complex, the questions for all the topics were deliberately kept easy (to ensure high retrievability), and most students likely possessed the prior knowledge necessary for understanding and actively encoding the learning contents. All these characteristics maximize the chances that most learners can profit from the testing effect but concurrently help to keep the potential influence of learner characteristics small. Future research could examine this issue directly by examining the moderating role of learner characteristics with materials of varying difficulty.

Another potential limitation is that lecture topics were not manipulated experimentally. Instead, participants could choose the topic. On consequence of this procedure is that the groups assigned to each lecture vary in size, with some groups being quite small. Also, the students potentially chose the topic they were most interested in and therefore might have been more motivated to learn and to improve their learning outcomes (Triarisanti & Purnawarman, 2019). On the positive side, however, the feature that students could choose the topic for which they wanted to study more approximates a typical self-regulated learning setting, enhancing the ecological validity of the study.

For practical reasons, we were able to include only five items per learning condition (testing vs. restudying) in the criterial test. The relatively low number of items permitted using aggregated measures of learning outcomes combined with an ANOVA approach. The alternative would be using a crossed-classified design with random effects of participants and items in a Generalized Linear Mixed Model (GLMM, Dixon, 2008). Such a design has the advantage of testing whether an effect generalizes not only to population of participants but also to populations of items and materials, but it requires larger numbers of items to achieve sufficient statistical power (Westfall, Judd, & Kenny, 2014). Based on the ANOVA approach used in the present study, we cannot conclude from our findings whether they generalize to larger population of items and contents. However, we varied lecture topics between-subjects, and the findings were extremely consistent across all five lecture topics. This fact makes us

optimistic that the conclusions that can be drawn from our results are not confined to the specific materials used in our study.

As a final limitation, practice testing was limited to short-answer questions and corrective feedback provided online to participants shortly after the actual lectures and an assessment of learning outcomes one week after learning. Therefore, predicting the extent that our results, especially the absence of a moderating role of learner characteristics, generalize to other forms of practice testing (such as practice tests with multiple-choice questions or free recall) and to other practice schedules or retention intervals is difficult.

To conclude, this research contributes to the available body of research by providing a strict experimental test of the effect of practice testing in a higher education context. Our result suggest that practice testing can be implemented effectively and efficiently as an online complement of regular university courses. Individual differences in cognitive, motivational, or emotional dispositions seem not to affect the magnitude of the testing effect, suggesting that it is a "learning method for all" (Jonsson et al., 2020) that potentially benefits all students.

References

Adesope, O. O., Trevisan, D. A., & Sundararajan, N. (2017). Rethinking the use of tests: A meta-analysis of practice testing. *Review of Educational Research, 87*(3), 659-701. https://doi.org/10.3102/0034654316689306

Agarwal, P. K., Finley, J. R., Rose, N. S., & Roediger III, H. L. (2017). Benefits from retrieval practice are greater for students with lower working memory capacity. *Memory, 25*(6), 764-771. https://doi.org/10.1080/09658211.2016.1220579

Agarwal, P. K., Nunes, L. D., & Blunt, J. R. (2021). Retrieval practice consistently benefits student learning: A systematic review of applied research in schools and classrooms. *Educational Psychology Review*, 33(4), 1409-1453. https://doi.org/10.1007/s10648-021-09595-9

Chapter II

Alexander, P. A. & Greene, J. A. (2017). *Self-regulation in education*. Routledge.

Artelt, C. & Lompscher, J. (1996). Lernstrategien und Studienprobleme bei Potsdamer

    Studierenden [Learning strategies and study problems among Potsdam Students]. In J.

    Lompscher & H. Mandl (Ed.), *Lehr- und Lernprobleme im Studium. Bedingungen und*

    Veränderungsmöglichkeiten (pp. 161-184). Bern: Huber.

Barenberg, J., & Dutke, S. (2019). Testing and metacognition: Retrieval practise effects on

    metacognitive monitoring in learning from text. *Memory, 27*(3), 269-279.

    https://doi.org/10.1080/09658211.2018.1506481

Bertilsson, F., Wiklund-Hörnqvist, C., Stenlund, T., & Jonsson, B. (2017). The testing effect

    and its relation to working memory capacity and personality characteristics. *Journal of*

    Cognitive Education and Psychology, 16(3), 241–259. https://doi.org/10.1891/1945-

    8959.16.3.241

Bertilsson, F., Stenlund, T., Wiklund-Hörnqvist, C., & Jonsson, B. (2021). Retrieval practice:

    Beneficial for all students or moderated by individual differences? *Psychology*

    Learning & Teaching, *20*(1), 21–39. https://doi.org/10.1177%2F1475725720973494

Bjork, R. A. (1975). Retrieval as a memory modifier: An interpretation of negative recency

    and related phenomena. In R. Solso (Ed.), *Information processing and cognition: The*

    Loyola Symposium (pp. 123-144). Lawrence Erlbaum Associates.

Bjork, E. L., & Bjork, R. A. (2011). Making things hard on yourself, but in a good way:

    Creating desirable difficulties to enhance learning. In M. A. Gernsbacher, R. W. Pew,

    L. M. Hough, J. R. Pomerantz (Eds.) & FABBS Foundation, *Psychology and the real*

    world: Essays illustrating fundamental contributions to society (pp. 56–64). Worth

    Publishers.

Bless, H., Wänke, M., Bohner, G., Fellhauer, R. F., & Schwarz, N. (1994). Need for

    Cognition: Eine Skala zur Erfassung von Engagement und Freude bei Denkaufgaben

[Need for Cognition: A scale measuring engagement and enjoyment in cognitive tasks]. *Zeitschrift für Sozialpsychologie*, *25*, 147-154.

Brewer, G. A., & Unsworth, N. (2012). Individual differences in the effects of retrieval from long-term memory. *Journal of Memory and Language, 66*(3), 407-415. https://doi.org/10.1016/j.jml.2011.12.009

Buchin, Z., & Mulligan, N. W. (in press). Retrieval-based learning and prior knowledge. *Journal of Educational Psychology*.

Butler, A. C., & Roediger, H. L. (2008). Feedback enhances the positive effects and reduces the negative effects of multiple-choice testing. *Memory & Cognition, 36*(3), 604-616. https://doi.org/10.3758/MC.36.3.604

Cacioppo, J. T., Petty, R. E., & Kao, C. F. (1984). The efficient assessment of need for cognition. *Journal of Personality Assessment*, *48*(3), 306–307. https://doi.org/10.1207/s15327752jpa4803_13

Carpenter, S. K. (2009). Cue strength as a moderator of the testing effect: The benefits of elaborative retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35*(6), 1563–1569. https://doi.org/10.1037/a0017021

Carpenter, S. K., Lund, T. J., Coffman, C. R., Armstrong, P. I., Lamm, M. H., & Reason, R. D. (2016). A classroom study on the relationship between student achievement and retrieval-enhanced learning. *Educational Psychology Review*, *28*(2), 353-375. https://doi.org/10.1007/s10648-015-9311-9

Cogliano, M., Kardash, C. 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

Dixon, P. (2008). Models of accuracy in repeated-measures designs. *Journal of Memory and Language, 59*(4), 447-456. https://doi.org/10.1016/j.jml.2007.11.004

Chapter II

Dweck, C. S. (1986). Motivational processes affecting learning. *American Psychologist*, *41*(10), 1040-1048. https://doi.org/10.1037/0003-066X.41.10.1040

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*(2), 175-191. https://doi.org/10.3758/BF03193146

Francis, A. P., Wieth, M. B., Zabel, K. L., & Carr, T. H. (2020). A classroom study on the role of prior knowledge and retrieval tool in the testing effect. *Psychology Learning & Teaching, 19*(3), 258–274. https://doi.org/10.1177/1475725720924872

Glenberg, A.M., Wilkinson, A.C. & Epstein, W. (1982). The illusion of knowing: Failure in the self-assessment of comprehension. *Memory & Cognition 10*(6), 597–602. https://doi.org/10.3758/BF03202442

Greving, S., Lenhard, W., & Richter, T. (2020). Adaptive retrieval practice with multiple-choice questions in the university classroom. *Journal of Computer Assisted Learning, 36*(6), 799-809. https://doi.org/10.1111/jcal.12445

Greving, S., Lenhard, W., & Richter, T. (2022). The testing effect in university teaching: Using multiple-choice testing to promote retention of highly retrievable information. *Teaching of Psychology*. Advance online publication. https://doi.org/10.1177/00986283211061204

Greving, S., & Richter, T. (2018). Examining the testing effect in university teaching: Retrievability and question format matter. *Frontiers in Psychology*, *9*: 2412. https://doi.org/10.3389/fpsyg.2018.02412

Harackiewicz, J. M., Barron, K. E., & Elliot, A. J. (1998). Rethinking achievement goals: When are they adaptive for college students and why? *Educational Psychologist, 33*(1), 1-21. https://doi.org/10.1207/s15326985ep3301_1

Hasselhorn, M., & Gold, A. (2022). *Pädagogische Psychologie: Erfolgreiches Lernen und Lehren*. Kohlhammer.

Hinze, S. R., & Rapp, D. N. (2014). Retrieval (sometimes) enhances learning: Performance pressure reduces the benefits of retrieval practice. *Applied Cognitive Psychology, 28*(4), 597–606. https://doi.org/10.1002/acp.3032

Hodapp, V. (1991). Das Prüfungsängstlichkeitsinventar TAI-G: Eine erweiterte und modifizierte Version mit vier Komponenten [The Test Anxiety Inventory TAI-G: An expanded and modified version with four components]. *Zeitschrift für Pädagogische Psychologie / German Journal of Educational Psychology, 5*(2), 121–130.

Jonsson, B., Wiklund-Hörnqvist, C., Stenlund, T., Andersson, M., & Nyberg, L. (2021). A learning method for all: The testing effect is independent of cognitive ability. *Journal of Educational Psychology*, *113*(5), 972–985. https://doi.org/10.1037/edu0000627

Judd, C. M., Kenny, D. A., & McClelland, G. H. (2001). Estimating and testing mediation and moderation in within-subject designs. *Psychological Methods, 6*(2), 115–134. https://doi.org/10.1037/1082-989X.6.2.115

Karpicke, J. D., Butler, A. C., & Roediger III, H. L. (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

Kirk-Johnson, A., Galla, B. M., & Fraundorf, S. H. (2019). Perceiving effort as poor learning: The misinterpreted-effort hypothesis of how experienced effort and perceived learning relate to study strategy choice. *Cognitive Psychology*, *115*, 101237. https://doi.org/10.1016/j.cogpsych.2019.101237

Kornell, N., Klein, P. J., & Rawson, K. A. (2015). Retrieval attempts enhance learning, but retrieval success (versus failure) does not matter. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*(1), 283. https://doi.org/10.1037/a0037850

Kornell, N., & Vaughn, K. E. (2016). How retrieval attempts affect learning: A review and synthesis. In B. H. Ross (Ed.), *The psychology of learning and motivation* (pp. 183–215). Elsevier Academic Press. http://dx.doi.org/10.1016/bs.plm.2016.03.003

Chapter II

Kubik, V., Gaschler, R., & Hausman, H. (2021). Enhancing student learning in research and educational practice: The power of retrieval practice and feedback. *Psychology Learning and Teaching, 20*(1), 1-20. https://doi.org/10.1177/1475725720976462

Kuhbandner, C., & Emmerdinger, K. J. (2019). Do students really prefer repeated rereading over testing when studying textbooks? A reexamination. *Memory, 27*(7), 952–961. https://doi.org/10.1080/09658211.2019.1610177

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*(1), 159–174. https://doi.org/10.2307/2529310

Lenhard, W., & Lenhard, A. (2016). *Berechnung von Effektstärken* [Computation of effect sizes] [Online resource]. https://www.doi.org/10.13140/RG.2.2.17823.92329

Marsh, E. J., Agarwal, P. K., & Roediger III, H. L. (2009). Memorial consequences of answering SAT II questions. *Journal of Experimental Psychology: Applied, 15*(1), 1-11. https://doi.org/10.1037/a0014721

McDaniel, M. A., & Butler, A. C. (2010). A contextual framework for understanding when difficulties are desirable. In A. S. Benjamin (Ed.), *Successful remembering and successful forgetting: A Festschrift in honor of Robert A. Bjork* (pp. 175-198). Psychology Press.

McDaniel, M. A., & Little, J. L. (2019). Multiple-choice and short-answer quizzing on equal footing in the classroom: Potential indirect effects of testing. In J. Dunlosky & K. A. Rawson (Eds.), *The Cambridge handbook of cognition and education* (pp. 480–499). Cambridge University Press. https://doi.org/10.1017/9781108235631.020

Morris, C. D., Bransford, J. D., & Franks, J. J. (1977). Levels of processing versus transfer appropriate processing. *Journal of Verbal Learning and Verbal Behavior, 16*(5), 519-533.

Pan, S. C., Pashler, H., Potter, Z. E., & Rickard, T. C. (2015). Testing enhances learning across a range of episodic memory abilities. *Journal of Memory and Language, 83*, 53-61. https://doi.org/10.1016/j.jml.2015.04.001

Pyc, M. A., & Rawson, K. A. (2010). Why testing improves memory: Mediator effectiveness hypothesis. *Science*, *330*(6002), 335. https://doi.org/10.1126/science.1191465

Ramsden, P. (2003). *Learning to teach in higher education*. Routledge.

Roediger III, H. L., & Butler, A. C. (2011). The critical role of retrieval practice in long-term retention. *Trends in Cognitive Sciences, 15*(1), 20-27. https://doi.org/10.1016/j.tics.2010.09.003

Roediger III, H. L., & Karpicke, J. D. (2006). Test-enhanced learning: Taking memory tests improves long-term retention. *Psychological Science*, *17*(3), 249–255. https://doi.org/10.1111/j.1467-9280.2006.01693.x

Rowland, C. A. (2014). The effect of testing versus restudy on retention: A meta-analytic review of the testing effect. *Psychological Bulletin*, *140*(6), 1432-1436. https://doi.org/10.1037/a0037559

Rybowiak, V., Garst, H., Frese, M., & Batinic, B. (1999). Error orientation questionnaire (EOQ): Reliability, validity, and different language equivalence. *Journal of Organizational Behavior*, *20*(4), 527–547. https://doi.org/10.1002/(SICI)1099-1379(199907)20:4%3C527::AID-JOB886%3E3.0.CO;2-G

Schindler, J., Schindler, S., & Reinhard, M.-A. (2019). Effectiveness of self-generation during learning is dependent on individual differences in Need for Cognition. *Frontline Learning Research*, *7*(2), 23–39. https://doi.org/10.14786/flr.v7i2.407

Schwieren, J., Barenberg, J., & Dutke, S. (2017). The testing effect in the psychology classroom: A meta-analytic perspective. *Psychology Learning & Teaching*, *16*(2), 179–196. https://doi.org/10.1177/1475725717695149

Spinath, B., Stiensmeier-Pelster, J., Schöne, C., & Dickhäuser, O. (2002). *Skalen zur*

    Erfassung der Lern-und Leistungsmotivation: SELLMO [Scales for assessing learning

    and performance motivation]. Hogrefe.

Soderstrom, N. C., & Bjork, R. A. (2014). Testing facilitates the regulation of subsequent

    study time. *Journal of Memory and Language, 73*, 99-115.

    https://doi.org/10.1016/j.jml.2014.03.003

Triarisanti, R., & Purnawarman, P. (2019). The influence of interest and motivation on

    college students' language and art appreciation learning outcomes. *International*

    Journal of Education, *11*(2), 130-135. http://dx.doi.org/10.17509/ije.v11i2.14745

Tse, C. S., & Pu, X. (2012). The effectiveness of test-enhanced learning depends on trait test

    anxiety and working-memory capacity. *Journal of Experimental Psychology: Applied*,

    18(3), 253. https://doi.org/10.1037/a0029190

Van Gog, T., & Sweller, J. (2015). Not new, but nearly forgotten: The testing effect decreases

    or even disappears as the complexity of learning materials increases. *Educational*

    Psychology Review, 27(2), 247-264. https://doi.org/10.1007/s10648-015-9310-x

Wacker, A., Jaunzeme, J., & Jaksztat, S. (2008). Eine Kurzform des

    Prüfungsängstlichkeitsinventars TAI-G [A short form of the test anxiety inventory

    TAI-G]. *Zeitschrift für Pädagogische Psychologie*, *22*(1), 73–81.

    https://doi.org/10.1024/1010-0652.22.1.73

Weissgerber, S. C., & Reinhard, M.-A. (2018). Pilot study on the relationship of test anxiety

    to utilizing self-testing in self-regulated learning. *International Journal of*

    Psychological Studies, *10*(4), 95-109. https://doi.org/10.5539/ijps.v10n4p95

Weissgerber, S. C., Reinhard, M.-A., & Schindler, S. (2016). Study harder? The relationship

    of achievement goals to attitudes and self-reported use of desirable difficulties in self-

    regulated learning. *Journal of Psychological and Educational Research*, *24*(1), 42-60.

Westfall, J., Kenny, D. A., & Judd, C. M. (2014). Statistical power and optimal design in experiments in which samples of participants respond to samples of stimuli. *Journal of Experimental Psychology: General*, *143*(5), 2020–2045. https://doi.org/10.1037/xge0000014

Wiklund-Hörnqvist, C., Jonsson, B., & Nyberg, L. (2014). Strengthening concept learning by repeated testing. *Scandinavian Journal of Psychology*, *55*(1), 10-16. https://doi.org/10.1111/sjop.12093

Xiaofeng, M., Xiao-e, Y., Yanru, L., & AiBao, Z. (2016). Prior knowledge level dissociates effects of retrieval practice and elaboration. *Learning and Individual Differences, 51*, 210-214. https://doi.org/10.1016/j.lindif.2016.09.012

Yan, V. X., Thai, K. P., & Bjork, R. A. (2014). Habits and beliefs that guide self-regulated learning: Do they vary with mindset? *Journal of Applied Research in Memory and Cognition*, *3*(3), 140-152. https://doi.org/10.1016/j.jarmac.2014.04.003Y

Yang, C., Luo, L., Vadillo, M. A., Yu, R., & Shanks, D. R. (2021). Testing (quizzing) boosts classroom learning: A systematic and meta-analytic review. *Psychological Bulletin, 147*(4), 399–435. https://doi.org/10.1037/bul0000309

Chapter III

The Testing Effect in the Lecture Hall: Does it Transfer to Content

Studied but not Practiced?

Study 2

The Testing Effect in the Lecture Hall: Does it Transfer to Content Studied but not Practiced?

Julia Glaser & Tobias Richter

Abstract. Practice tests have been shown to be an effective means to foster long-term retention in higher education, at least compared to restudying (i.e., the testing effect). The present study replicated and extended prior research by examining whether and to what extent the positive effects of testing on long-term retention in a typical psychology lecture transfers to content presented only during initial learning (and not practiced). Using a within-subjects design, we alternated post-lecture multiple-choice practice tests and restudying opportunities in two psychology classes ($N = 67$). One week after the final lecture session of a cycle of six weekly lecture sessions, retention of learning content was assessed by comparing performance on questions referring to content practiced via testing, encountered via restudying, or unreviewed. We found a testing effect for practiced content, whereas no transfer effect occurred for untested content from the same lecture sessions. These results show that the testing effect is a powerful learning tool, but also suggest a possible boundary condition pertaining only to explicitly tested content. Practice testing should be integrated regularly in higher-education courses to foster long-term retention for a final test. However, educators should take care that important content is fully covered in practice tests.

The Testing Effect in the Lecture Hall: Does it Transfer to Content Studied but not Practiced?

One goal of formal education from kindergarten to higher education is to foster the acquisition of lasting knowledge (Richter et al., 2022). Test-enhanced learning, consisting of an initial study phase followed by practice tests administered in class or via self-testing, has been shown to be an effective and efficient way to foster long-term retention more broadly (Roelle et al., 2022; Rowland, 2014; Yang et al., 2021) and more specifically in psychology courses (see Schwieren et al., 2017, for a meta-analysis). In a typical implementation of retrieval practice in a university classroom, students attend a class and are provided with a practice test afterwards with or without feedback. Researchers would then compare studying plus retrieval to studying plus restudying on a retention test administered sometime after learning, usually after one week or even later (Roediger & Butler, 2011).

In the research partially replicated and extended in the current study, we asked introductory psychology students to complete short-answer questions for half of the content in the lecture session with corrective feedback (practice test) and summarizing statements (restudy) for the other half of the content; this took place during review sessions administered online in the week after the lecture session (Glaser & Richter, 2023a). Across six different lecture topics, a positive testing effect emerged in a final test one week after learning. Content for which short-answer questions were posed was remembered better than content for which summarizing statements were presented, with a medium effect size (Cohen's $d = 0.55$) (Glaser & Richter).

Meta-analytic reviews suggest that positive testing effects such as the one obtained in our recent study (Glaser & Richter, 2023a) are a very robust phenomenon. Practice testing seems to be more effective than restudying in different age groups, with different types of learning material, and in the laboratory, with an overall medium effect size (Hedge's $g = 0.50$, Rowland, 2014) and classroom settings ($g = 0.33$; Yang et al., 2021; see also Agarwal et al.,

2021). The meta-analyses and additional research also point to relevant moderators. Practice testing seems to be more effective for longer retention intervals, stressing its usefulness for fostering lasting learning (Rowland, 2014). Corrective feedback for responses in the practice test can increase the positive effect of testing by strengthening existing knowledge and allowing learners to detect knowledge gaps and to improve metacognitive calibration, which benefits future learning activities (*indirect testing effect*: Arnold & McDermott, 2013). If no feedback is given, the retrievability of the learned content is crucial because successful retrieval is a precondition for retrieval practice to be effective (*direct testing effect*; Greving et al., 2022; Greving & Richter, 2018). The questions in the practice test must stimulate active retrieval, which can be achieved through cued-recall (short-answer) questions or appropriately designed multiple-choice questions (Butler, 2018). Finally, repeated testing has been shown to be particularly beneficial for learning (Butler, 2010).

Given the robust learning effects of test-enhanced learning, practice testing is considered a silver bullet for promoting lasting knowledge (Dunlosky et al., 2013). However, the evidence for the applicability of test-enhanced learning in educational contexts is somewhat limited because in most experimental studies, the criterial test covers exactly the practiced or reviewed content after learning. Although demonstrating the effectiveness of practice tests for the explicitly tested content is important, the applicability of test-enhanced learning would be greatly expanded if practice tests would also benefit the retention of other content encountered during the initial learning session, even if the information was not explicitly tested. For instance, psychology students attending a lecture are typically expected to acquire broad knowledge of the topics covered in the lecture, and psychology educators using low-stakes quizzes in their classroom usually do so in the hope to make all the important content stick, not just the subset of facts that are covered in the practice test.

A meta-analysis by Pan and Rickard (2018) that included 192 effect sizes from studies with different types of learning materials (often expository texts) revealed a medium-sized

positive effect of practice testing across different types of transfer, including application and inference questions, transfer from one question format to another, and the transfer to content encountered during initial learning but not covered by the practice test. However, for the latter type of transfer, the research is especially inconclusive. The present field experiment replicated our previous research (Glaser & Richter, 2023a) and extended it to provide a clarification of this question in a typical lecture setting. To ground our research questions, we briefly review theoretical explanations of the testing effect that suggest practice testing might extend to content learned but not explicitly tested. We will then discuss the available empirical studies that have addressed this phenomenon in more detail.

Transfer of Test-enhanced Learning to Content Studied but not Practiced

Transfer occurs when "learning in one context or with one set of materials impacts on performance in another context or with other related materials" (Perkins & Salomon, 1992, p. 425). In the present study, we examined a specific form of transfer effects that is potentially elicited by test-enhanced learning. Specifically, we were interested in whether the beneficial effects of retrieval practice extend to content learned in the same lecture session but not explicitly covered in the practice test. Theoretical accounts of the testing effect suggest that such transfer effects are possible. For example, Carpenter (2009) proposed elaborative retrieval as a mechanism that underlies direct testing effects. According to her account, (successful) retrieval not only involves the activation of the retrieved information but also the activation of related concepts through a spreading activation mechanism. Citing an example from Anderson (1976), Carpenter illustrates the mechanism of elaborative retrieval by the attempt to learn the association between the words *dog* and *chair* by thinking of a dog who loved to sit on his master's chair but was scolded by the master for leaving his hairs on the chair. This brief story creates an elaborative retrieval structure that provides multiple pathways for future retrieval attempts (e.g., *dog-scold-chair*, *dog-master-chair*, *dog-sit-chair*, *dog-master-scold-chair*, ...), thus facilitating long-term retention. Once formed, the

elaborative structure should also facilitate retrieval of other elements of the structure such as the concept *chair* when cued with *master* or *master* when cued with *chair*.

In line with this assumption, Carpenter (2011) found that retrieval practice with word pairs not only improved cued recall of the practiced target words but also recall of words semantically associated with the word pairs. These semantic associates act as mediators and can enhance the recall of the learned materials (see also Carpenter & Yeung, 2017; Pyc & Rawson, 2010) but in this process, the retrievability of the mediators is also enhanced. In educational settings, the concepts taught within a lesson on a specific topic are usually coherent with each other, which is a favorable condition for the formation of elaborative structures. Therefore, if practice testing elicits elaborative retrieval, the memory traces of information contained in the lesson but not tested in the practice test might also be strengthened if this information is associated with the practiced content. In addition to elaborative retrieval, mechanisms assumed to underly indirect testing effects might also contribute to this specific type of transfer effect, which has also been called *retrieval-induced facilitation* (Chan et al., 2006).

In their meta-analysis of transfer effects, however, Pan and Rickard (2018) found no significant overall transfer effects for studies that examined transfer effects to the retention of untested materials seen during initial study ($d = 0.16$, $p = .20$, $k = 17$). In other words, the hypothesis of retrieval-induced facilitation received no support. However, the analysis also revealed a considerable heterogeneity of effects that warrants a closer look at the specifics of the primary studies.

Several studies included in Pan and Rickard's (2018) meta-analysis provide support of transfer effects to untested content, most of them based on expository texts (Butler, 2010, Experiments 1b, 2, and 3; Chan et al., 2006; Chan, 2010; Hinze et al., 2013, Experiment 2). In these studies, great care was taken that the transfer questions were indeed semantically highly coherent with the practiced content, which might be a critical condition for transfer effects to

occur. Other studies used specific tasks in the practice tests that might facilitate transfer effects to untested contents. For example, Hinze et al. (2013, Experiment 3) found that transfer effects occurred only if the practice test included the task to write an explanation, which is likely to foster elaborative (or constructive) retrieval, but not if the task was free recall. Another example is a classroom experiment conducted by Balch (1998) in an introductory psychology course. In this study, the students who took the practice tests additionally scored the practice test of another participant before they received feedback on their own answers. This additional task might have enhanced elaborative retrieval, thus contributing to a transfer effect. Finally, McDaniel et al. (2012) reported transfer effects in two field experiments conducted with undergraduates as part of a web-based class on biopsychological topics, one experiment (based on a very small sample) providing support, a second experiment (based on a slightly larger sample) failing to provide support for transfer effects of testing compared to restudying. Thus, the results by McDaniel et al. (2012) are inconclusive regarding the focal question.

Other studies included in the meta-analysis by Pan and Rickard (2018) also provided evidence against transfer effects of testing, at least against transfer to contents not strongly related to the practice questions. For example, in three experiments reported by Butler (2010), performance in control questions that referred to the expository text read in the study phase was worse in the practice test conditions compared to the restudy condition that included the possibility to reread whole passages (Experiments 1a, 1b, and 2). In still other studies, such as a study by Nungester and Duchastel (1982) with high-school students who read a history text, or classroom experiments by Wooldridge et al. (2014) and La Porte and Voss (1975) with undergraduates, transfer effects of testing on studied but untested content were not significant. However, like the experiments by McDaniel et al. (2012), several of these studies were based on small sample sizes and are likely to be underpowered, which limits the conclusions that can be drawn from their results. Another classroom experiment by Pilotti et al. (2009) is

likewise inconclusive regarding the focal question, as participants in the restudy condition (i.e., the control condition) received the practice questions with answers and were prompted to think about why the answer was correct. In other words, participants in the control condition engaged in an elaborative activity that might have enhanced transfer.

Rationale of the Present Study

The studies reviewed in the preceding section rely on a broad variety of practice tasks, use very different types of materials and practice tests, and are overall inconclusive on the question of whether practice tests improve long-term retention beyond the content that is explicitly tested. Given the practical relevance of this question for psychology educators, the current study addressed transfer effects of practice testing to untested contents in typical psychology courses, covering several weeks of a semester and with a broad range of topics, to explore the generalizability of results. As a replication and extension of our prior work (Glaser & Richter, 2023a), we examined whether short practice tests that supplemented six regular lecture sessions in two different psychology lectures during one semester would benefit long-term retention, measured after a retention interval of one to several weeks post-learning (one week after the review session following the last of six consecutive weekly lecture sessions). In addition to examining testing effects for the knowledge targeted by the practice questions, we focused on the question whether the positive effects of practice tests on learning would extend to other content from the same lecture, which was neither restudied nor explicitly tested. In terms of direct effects of testing, such an effect might occur when practice testing encourages elaborative retrieval, which includes the activation of content linked to the content directly targeted by the question (Chan et al., 2006).

We compared the effects of practice testing with multiple-choice questions to the effects of restudying summary statements that were equivalent to the correct responses in the practice tests (within-subjects, alternating weekly). Both the testing condition and the restudying condition were offered as online learning materials and implemented in a highly

economical way, taking only a few minutes to complete. The questions in the practice tests were presented with corrective feedback and repeated once if participants failed to provide the correct answer in the first run, which should create favorable conditions for testing effects to occur. In the final test presented one week after the final review session, participants received either multiple-choice or short-answer questions.

In this setting, we tested the following hypotheses. First, we expected a testing effect. Content from lecture sessions accompanied by practice tests should be remembered better than content from lecture sessions accompanied by restudy materials (Hypothesis 1). However, this testing effect should be stronger for questions that are directly tested in the practice tests compared to questions that refer to content from the same lecture sessions, but which are not also part of the practice tests or the restudy material (Hypothesis 2). Third, we expected learning outcomes to be better for questions that directly refer to reviewed (practiced or restudied) content than learning outcomes assessed with new questions that refer to content that was not reviewed (Hypothesis 3).

Additionally, we examined the following exploratory research questions. First, we examined whether the testing effect would be stronger for learners who performed better in the practice tests, as reflected in a higher average retrievability in the practice tests. This research question is based on theory and research showing that retrievability is crucial for retrieval practice to be effective, at least when no feedback is presented (for evidence in a lecture context, see Greving & Richter, 2018). Moreover, we explored whether the question format in the final test would make a difference in the testing effect. Insofar as transfer-appropriate processing plays a role for the testing effect (Veltre et al., 2015), beneficial effects of retrieval practice might be larger if the same question format (multiple-choice) is used in the practice and the final test compared to different question formats (multiple-choice vs. short-answer questions). Finally, by employing two types of questions in the final test, we were able to investigate whether the testing effect varies as a function of whether the learning

outcomes are assessed in the same or a different response format as in the practice test. The principle of transfer-appropriate processing (Morris et al., 1977) suggests that the congruence of response format might be important. Nevertheless, testing effects with incongruent response formats in practice and final tests have been found in previous studies (yielding a mean effect size of $d = 0.28$ in the meta-analysis by Pan & Rickard, 2018), even though they seem to be smaller than testing effects found with congruent response formats ($d = 0.58$).

Method

Participants

We conducted an online study with students enrolled in the teacher-training program at the University of Würzburg. Of the 97 students who signed up for the study, 67 completed all parts of the study. The remaining 30 students had to be excluded from the analysis because they did not participate in all parts of the study or skipped large parts of the practice tests (more than 50%). Participants whose data were excluded did not differ significantly from those whose data were included in relevant learner characteristics such as gender, age, study performance and prior knowledge (see Table II.1).

Table II.1

Comparison between Included and Excluded Participants

		Age	Grade	NFC	Prior Knowledge
Final sample	*M (SD)*	20.63 (2.92)	2.11 (0.54)	4.46 (0.89)	2.47 (0.70)
Dropped-out participants	*M (SD)*	22.03 (5.08)	2.17 (0.53)	4.81 (0.75)	2.45 (0.73)
t		1.415	0.513	1.907	-0.127
p		.165	.609	.059	.899
Cohen's d		0.38	0.11	0.42	-0.03

Note. Final sample $n = 67$. Dropped-out participants: $n = 30$. Age in years. Grade: German A-Levels (ranging from 1 = very good to 6 = deficient). NFC: Need for cognition score (ranging from 1 = low NFC to 5 = high NFC). Prior Knowledge: self-rated prior knowledge (ranging from 1 = hardly any prior knowledge to 5 = extensive prior knowledge).

Participants were recruited from two different courses (*Behavioral and Learning Disorders in Childhood and Youth* and *Developmental Psychology in Childhood and Youth*), both of which are part of the mandatory psychology curriculum for prospective teachers in their first year of study. Twenty-five students were studying to become elementary school teachers (Grundschule), four were non-academic track middle-school teachers (Mittelschule), three were studying for secondary school (Realschule), 21 for high school (Gymnasium), 11 for special education (Sonder-/Förderschule), and two were studying pedagogy (missing data from 1 student).

Most of the 67 participants included in the analysis identified as female (86.6%) and their age ranged from 18 to 37 years ($M = 20.63$, $SD = 2.92$). Only one participant reported a native language other than German. All participants received course credits for their participation in the study. Table II.2 shows detailed information on the sample.

Table II.2

Detailed information on the sample

	Age	Grade	NFC	Prior Knowledge
M	20.63	2.11	4.46	2.47
SD	2.92	0.54	0.88	0.70
min	18	1.00	2.31	1.00
max	37	3.10	6.25	4.83

Note. $N = 67$ for age, NFC, Prior Knowledge; $N = 66$ for Grade.

A sensitivity analysis revealed that assuming a power (1-β) of .90, a Type I error probability $\alpha = .05$ and a correlation $\rho = .561$ between the levels of the repeated-measures factor (the median of the observed correlations), the design was sensitive enough to detect a small testing effect of $f = 0.154$ (corresponding to $\eta^2 = .023$ or Cohen's $d = 0.308$; power analysis performed with GPower, Faul et al., 2007; effect size transformations computed with the tool provided by Lenhard & Lenhard, 2016).

Materials

Practice Tests and Summary Statements

Within each of the two courses, there were 12 sessions, for each of which 15 information units were identified. For every information unit, we constructed a summarizing statement and a short-answer or a multiple-choice question. Summarizing statements were created by summarizing the main ideas of the information unit in one short sentence (e.g., *Dyslexia is stable during development, but therapy can have a positive influence*). Short-answer questions were created by asking for the main ideas out of the information unit (e.g., *What can be said about the stability of the development of dyslexia, especially in relation to therapy? Give your answer in 1-2 short sentences*). Multiple-choice questions were created by adding four possible answers to the short-answer question with varying numbers of correct alternatives (e.g., response options: *The development of dyslexia ... (a) is stable and cannot be influenced by therapy; (b) is stable, but therapy can have a positive influence; (c) is unstable depending on the child's development in other areas; (d) is unstable, and the problem disappears in some cases without therapy*). The questions including the sample solution for all topics can be viewed in OSF in German as well as in an English translation.

For each lecture session, 10 of the 15 information units were selected and presented either as a summary statement or as a multiple-choice question. The statements were presented, and participants had up to 60 seconds to restudy them but could move on to the next item earlier. For the multiple-choice questions, participants also had up to 60 seconds to provide their responses. After their answer, they were shown corrective feedback (*correct/incorrect*) and the correct answers. If participants chose a wrong answer or missed a correct option for a question, they received the question again in the end. As before, they received feedback on their answers.

Final Test

The dependent variable was based on the performance on a final test with a total of 60 questions, 10 for each of the previously learned topics. For each topic, five questions referred to information units that had been presented or tested in the practice phase, the other five referred to information units not presented in the practice phase.

The final tests were either presented with multiple-choice questions ($n = 30$) or with short-answer questions ($n = 37$). The distribution of participants among question format was randomized and balanced within the two lectures (see Table II.3).

Table II.3

Distribution of Participants between Lecture Topics and Means and Standard Deviations for their self-rated Prior Knowledge for each Topic

Lecture	Pre-survey	Session 1		Session 2		Session 3		Session 4		Session 5		Session 6		Criterial Test		Valid (cases)
		Re-read	Test	Re-read	Test	Re-read	Test	Re-read	Test	Re-read	Test	Re-read	Test	MC	SA	
Behavioral and Learning Disorders	54	26	20	24	21	21	24	22	21	21	20	22	20	20	24	39
Self-rated Prior Knowledge: *M (SD)*		2.66 (0.97)		1.85 (0.94)		2.76 (0.89)		2.71 (0.96)		2.44 (0.95)		3.15 (0.88)				
Development in Childhood & Youth	43	21	11	20	10	11	19	19	10	10	18	19	10	15	14	28
Self-rated Prior Knowledge: *M (SD)*		2.31 (0.97)		2.24 (0.88)		2.23 (0.95)		1.62 (0.98)		2.77 (0.91)		2.42 (1.03)				
Total		47	31	44	31	32	43	41	31	31	38	41	30	35	38	
	97	78		75		75		72		69		71		73		67

Note. Participants received a link and should fill in the next questionnaire or re-read statements. Some participants copied links from other participants, so they participated in the wrong learning condition. Therefore, the figures for the individual conditions vary. Valid cases are participants with complete data that were included in the analyses.

The multiple-choice questions were scored by partial credit, with 0.25 points given for each response option that was correctly ticked or correctly not ticked. The short-answer questions were also scored according to a partial credit scheme (0-1 points in intervals of 0.25). Two raters independently coded the responses to 120 questions from 12 participants (60

questions for each participant) to estimate inter-rater reliability of the short-answer questions. Cohen's κ was .816 (*SE* = .017), indicating almost perfect agreement according to Landis and Koch (1977). The interrater agreement for each topic is provided in Table II.4. Given the high inter-rater reliability, the remaining answers to the short-answer questions were coded by only one coder.

Table II.4

Values of Cohen's K ordered by Lecture and Topics

Lecture	Topic	Cohen's K	*SE*
Behavioral and Learning Disorders	Reading and writing problems	.796	.061
	Maths	.786	.064
	ADHD	.703	.071
	Aggression	.703	.074
	Violence	.768	.074
	Fear	.798	.062
	Average	.709	.029
Development in Childhood and Youth	Development of Thinking	.876	.051
	Development of Intelligence	.764	.066
	Development of Memory	.788	.063
	Development of Meta-memory	.950	.034
	Development of Language	.872	.055
	Development of Motivation	.921	.044
	Average	.865	.022
Overall		.816	.017

Retrievability

For each participant, a mean retrievability score was computed based on the sum of correctly answered questions in the practice tests. The minimum score for each question was 0 (none of the four response options correctly ticked or left unticked), the maximum score was 1 (all four response options correctly ticked or left unticked). Each correctly ticked or not ticked option was scored with 0.25 points.

Additional Measures

For exploratory purposes, we also assessed need for cognition with the short form of the Need for Cognition Scale (Bless et al., 1994). The exploratory analyses with this scale yielded no interesting insights. Therefore, the results will not be reported. Participants also provided single-item ratings of their prior knowledge for the six topics covered in the chosen lecture (used for the description of the sample) and the comprehensibility and usefulness of the study for their own learning activities on 5-point scales. Participants also indicated whether they had used other learning methods than those implemented in the review sessions. Materials and results concerning these measures are provided in the Appendix and in the online supplement (Glaser & Richter, 2023b).

Design

We investigated the effect of learning condition (restudy vs. testing) on the final test performance (either questions referring to restudied or tested content vs. new content). Learning type and type of test question were varied within-subjects. Each participant received practice tests for three out of six lecture topics and summarizing statements to restudy three other lecture topics. The assignment of lecture topics to either the testing or the restudy condition as well as the sequence of conditions was counterbalanced between participants. Each participant was randomly assigned to one of the two sequences. The final test, which covered all previous lecture topics, was presented as either short-answer or multiple-choice questions and included questions that had previously been asked in the same wording as well as questions that had previously been presented as a statement and questions that addressed related, untested knowledge.

In sum, the study was based on a 2 (learning condition: testing vs. restudy) X 2 (question type in the final test: restudied/tested vs. new content) X 2 (question format in the final test: short-answer vs. multiple-choice tests) within-subjects design. Moreover, the lecture

varied between participants as a quasi-experimental factor with two levels (*Behavioral and Learning Disorders in Childhood and Youth* vs. *Development in Childhood and Youth*).

Procedure

Data for the study were collected from November 2021 until January 2022 with the online survey tool SoSci Survey (www.soscisurvey.de). Across all topics, participants estimated their prior knowledge on the lecture topic as moderate ($M = 2.47$, $SD = 0.70$, on a 5-point scale from 1 to 5). The distribution of participants between lecture topics and means and standard deviations for their self-rated prior knowledge for each topic are provided in Table II.3 above.

After choosing the course they attended (*Behavioral and Learning Disorders in Childhood and Youth* or *Development in Childhood and Youth*), participants provided their demographic information and received the *Need for Cognition* scale, and the self-assessment of prior knowledge. They also provided their email address to register for the subsequent parts of the study.

During the course, students attended their classes as usual. Two days after the first lecture lesson, which was part of the study (*Reading and Writing Problems* or *Development of Thinking*), participants received their first link with the review task via email. They were asked to respond to the practice questions or restudy the summarizing statements to review the core contents of the lecture session. Participants were randomly assigned to either the restudy condition or the testing condition. One week later, two days after the second thematic lesson, they received the second link with a new review task via email. Those participants assigned to the testing condition in the first review session now received the summarizing statements, whereas those participants assigned to the restudy condition in the first review session now received the practice test. This procedure was repeated for Sessions 3 and 4 and again for Sessions 5 and 6. Therefore, each participant received the restudy condition and the practice session each three times, in an alternating sequence.

The final part of the study took place one week after the completion of the final review session. Participants now received the final test with 60 questions (either multiple-choice or short-answer questions), covering all six themes of the previous lessons and containing questions referring to practiced/restudied content and questions referring to lecture content that was not also part of the review session. Finally, participants rated the comprehensibility and usefulness of the study for their own learning activities and indicated whether they had used other learning methods than those implemented in the review sessions. Figure II.1 illustrates the design and procedure of the experiment.

Figure II.1

Flow Diagram Illustrating the Design and Procedure of the Study
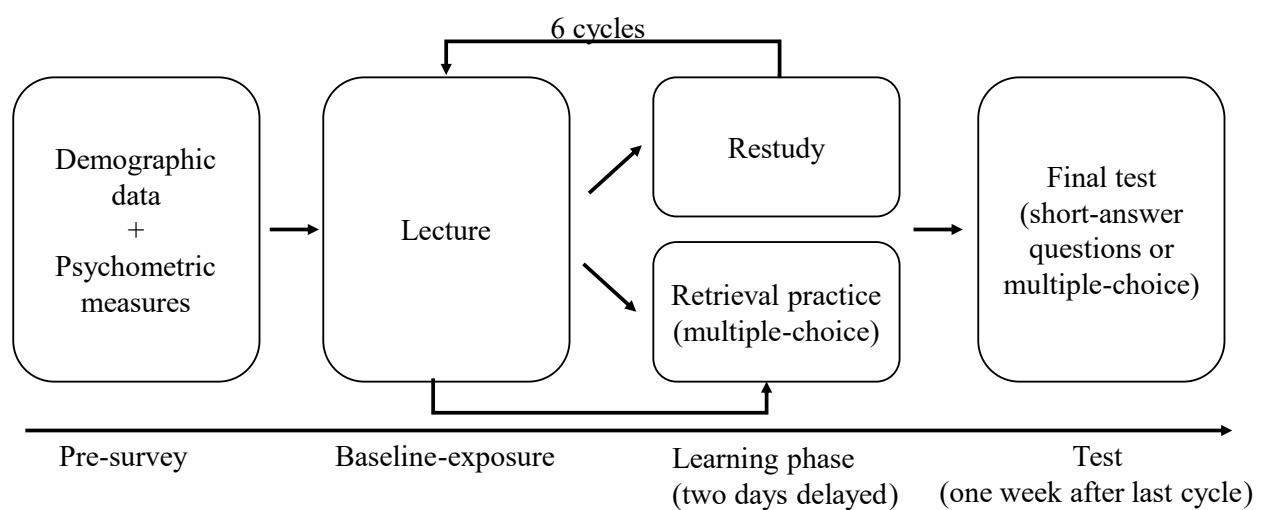
Note. Participants from two different lectures could volunteer to participate in the experiment. In each lecture, demographic data and psychometric measures were collected first, followed by the baseline exposure to one of the lecture sessions. Two days after the session, the review phase took part with either retrieval practice or restudy (within-subjects, alternating weekly). These two steps were repeated weekly for 6 weeks. One week after the final lecture session, a test was administered, which covered all former lecture topics. The test was either presented with short-answer or multiple-choice questions.

Results

We estimated a linear model (mixed ANOVA) with learning condition and question type as experimental factors varied within-subjects. Moreover, type of final test and lecture were included as between-subject factors. The model also included the interactions of these

variables. We used an α-level of .05 for all statistical tests. We report partial η^2 as the effect size measure. All data files and analysis scripts can be found in the Appendix and in the online repository (Glaser & Richter, 2023b).

Effects of Learning Condition, Question Type, and Question Format

We found no main effect for learning condition, $F(1, 64) = 0.63$, $p = .429$, $\eta_p^2 = .01$. Thus, Hypothesis 1, predicting an overall testing effect, could not be supported. Importantly, however, the interaction of learning condition and question type was significant, $F(1, 64) = 5.18$, $p = .026$, $\eta_p^2 = .08$. Figure II.2 displays the interaction of learning condition and question type. Follow-up tests revealed a testing effect for questions referring to reviewed content. Final test performance in questions that referred to tested contents ($M = .70$, $SE = .020$) was better than performance in questions that referred to restudied contents ($M = .66$, $SE = .017$), $F(1, 64) = 5.33$, $p = .024$, $\eta_p^2 = .08$. In contrast, for questions referring to lecture content that was not reviewed, we found no significant difference between sessions followed by testing ($M = .62$, $SE = .018$) compared to restudying ($M = .63$, $SE = .017$) , $F(1, 64) = 0.74$, $p = .393$, $\eta_p^2 = .01$. Thus, Hypothesis 2 that testing would especially benefit the tested content was supported. More specifically, a testing effect occurred only for content explicitly tested and no transfer effect to untested content from the same lecture occurred.

Figure II.2

Interaction of Learning Condition and Question Type
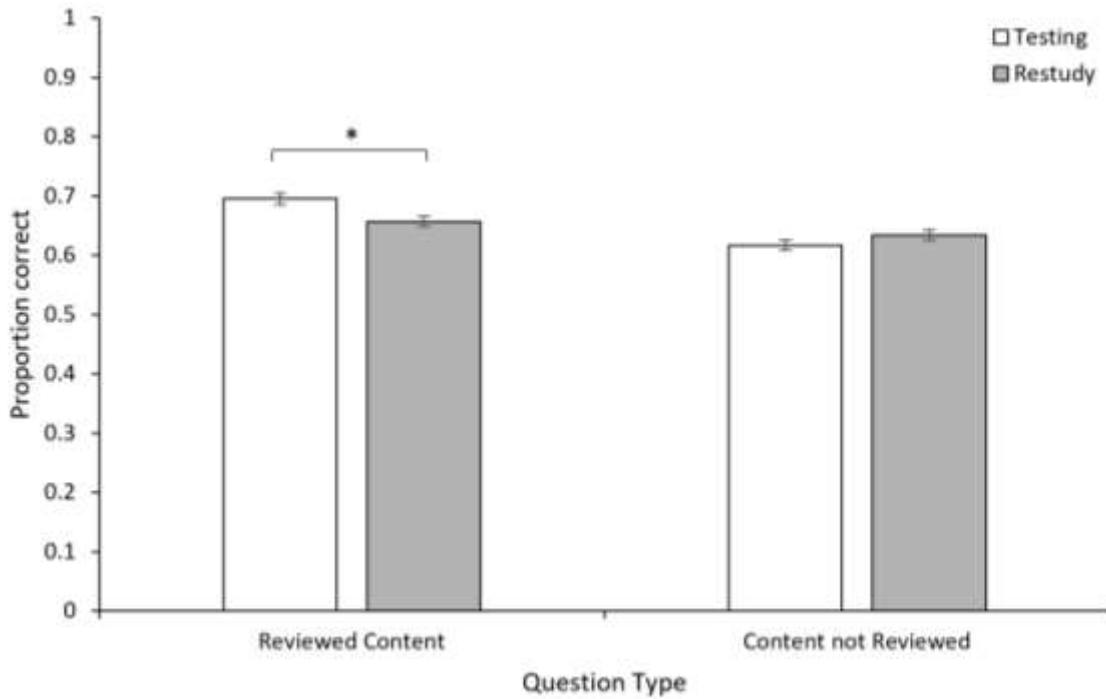
Note. Error bars represent the standard error of the mean. The *y*-axis shows the proportion of correct responses in the final test.

The interaction of learning condition and question type was significant, $p = .026$. We found a strong main effect for question type, with overall better learning outcomes in the final test for questions referring to reviewed (i.e., tested or restudied) content ($M = .68$, $SE = .017$) compared to questions referring to content not reviewed after the lecture ($M = .63$, $SE = .014$), $F(1, 64) = 19.09$, $p < .001$, $\eta_p^2 = .23$. Therefore, Hypothesis 3 predicting a better learning outcome for questions referring to reviewed (tested or restudied) content was basically supported. However, the interaction effect of question type and question format was also significant, $F(1, 64) = 5.83$, $p = .019$, $\eta_p^2 = .08$. Figure II.3 displays the interaction of question type and question format. Follow-up tests revealed that the better learning outcome for reviewed content was significant only with multiple-choice questions (reviewed content: $M = .75$, $SE = .024$; content not reviewed: $M = .67$, $SE = .021$), $F(1, 64) = 21.19$, $p < .001$, $\eta_p^2 = .25$. In contrast, performance on short-answer questions was not significantly different

between the two question types (reviewed content: $M = .60$, $SE = .022$; content not reviewed: $M = .58$, $SE = .019$), $F(1, 64) = 2.16$, $p = .147$, $\eta_p^2 = .03$.

**Figure II.3**

Interaction of Question Type and Question Format
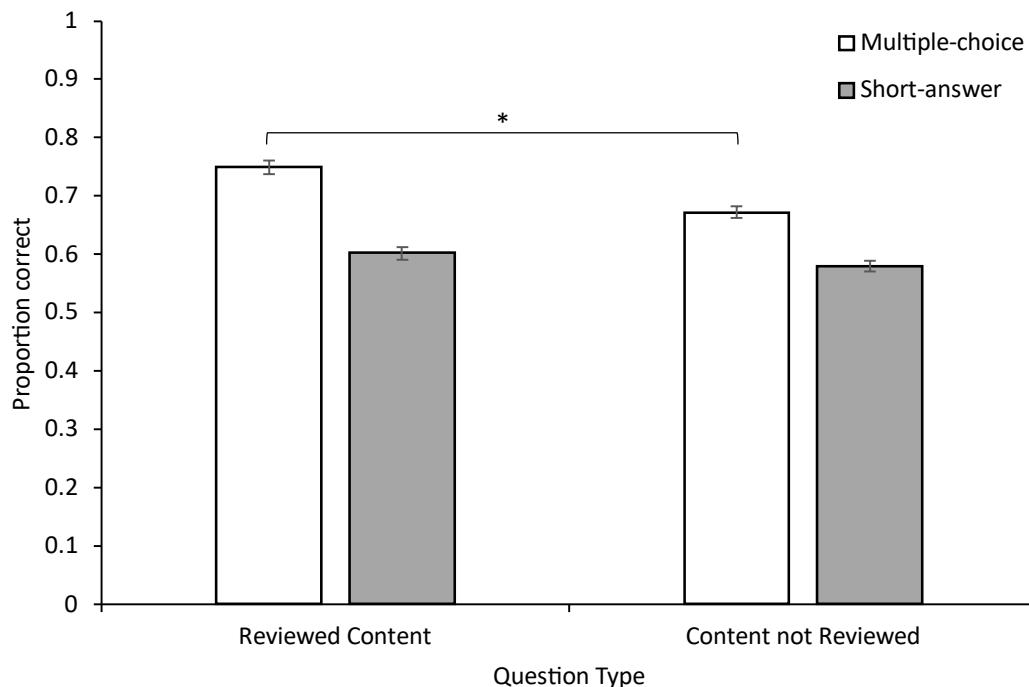
Note. Error bars represent the standard error of the mean. The *y*-axis shows the proportion of correct responses in the final test.

We also tested whether lecture topic would moderate the hypothesized effects, which would provide information about the generalizability of the effects. Only the interaction of lecture topic with question type was significant, $F(1, 64) = 6.67$, $p = .012$, $\eta_p^2 = .09$. Follow-up tests revealed that learning outcomes for reviewed content compared to lecture content that was not reviewed was significant only in the *Behavioral and Learning Disorders in Childhood and Youth* lecture (reviewed content: $M = .71$, $SE = .021$; content not reviewed: $M = .63$, $SE = .019$), $F(1, 64) = 28.60$, $p < .001$, $\eta_p^2 = .31$, but not in the *Development in Childhood and Youth* lecture, $F(1, 64) = 1.38$, $p = .244$, $\eta_p^2 = .02$ (reviewed content: $M = .64$, $SE = .025$; content not reviewed: $M = .62$, $SE = .022$). None of the other interactions of lecture

topic with any of the other independent variables was significant (for all tests, $p > .103$). An overview of all effects and interactions is provided in Table II.5.

Table II.5

Overview of all main effects and interactions

Effect	F (1, 64)	p	$\eta_p{}^2$
Learning Condition	0.63	.429	.01
Learning Condition * Lecture	2.72	.104	.04
Learning Condition *Question Format	0.37	.548	.01
Question Type	19.09	< .001	.23
Question Type* Lecture	6.67	.012	.09
Question Type * Question Format	5.83	.019	.08
Learning Condition * Question Type	5.18	.026	.08
Learning Condition * Question Type * Lecture	1.99	.163	.03
Learning Condition * Question Type * Question Format	0.77	.384	.01

Moderating Effect of Retrievability

In an additional step, we estimated an expanded linear model to explore whether the testing effect found in the primary analysis would be moderated by the average retrievability, operationalized as performance in the practice tests. When adding retrievability to the model, we found a significant main effect on final test performance, $F(1, 63) = 10.81$, $p = .002$, $\eta_p{}^2 = .15$, with the average higher retrievability being associated with better learning outcomes ($r = .40$). However, none of the interactions with learning condition was significant (for all tests of two- and three-way interactions), and all the hypothesis-relevant effects reported above remained basically unchanged after including retrievability in the model (see Table II.6).

Table II.6

Overview of all main effects and interactions concerning Retrievability

Effect	$F (1, 63)$	p	η_p^2
Retrievability	10.81	.002	.15
Learning Condition * Retrievability	0.69	.411	.01
Question Type * Retrievability	0.25	.616	< .01
Learning Condition * Question Type * Retrievability	0.35	.557	.01

Discussion

The present study replicated and extended our prior research (Glaser & Richter, 2023a), specifically examining whether short practice tests would benefit long-term retention compared to restudying the information and whether this effect would extend to non-reviewed content that was featured in the same lecture session. We created a minimal intervention over a period of 6 weeks in two regular university lectures, with six topics each and a within-participant and within-topic variation of practice testing versus restudying. Learning outcomes were assessed after the review session following the final lecture in a cycle of six consecutive weekly lecture sessions, implying a retention interval of 1 to 6 weeks after learning, depending on when in the cycle the topic was learned. The main finding was that a testing effect occurred only for content that was part of the practice test but not for content that was featured in the lecture but not tested. We also found that reviewed content (i.e., tested or restudied) was overall remembered better than content that was only encountered during the lecture, at least for multiple-choice questions. Finally, we found no moderating effect of response congruency in the practice tests (containing short-answer questions) nor the final test (containing short-answer questions versus multiple-choice).

These results contribute to the literature on transfer effects of testing by providing clear evidence against the assumption of transfer effects to untested material. The significant interaction obtained in the present study demonstrates that the testing effect for questions that appeared in the practice effect and the final test is larger than the transfer effect for final test questions that refer to untested content. Moreover, our study was adequately powered, suggesting that the null effect for the transfer effect of testing was not due to a low sensitivity. Apart from these statistical arguments, a strength of the present study is its implementation in two psychology lectures based on curricular content distributed over 12 different topics. With these features, the results are informative for the application of practice testing in typical psychology courses in higher education. Our results suggest, in line with numerous other studies conducted in the university classroom (see Yang et al., 2021), that practice tests are an effective way to increase retention of the tested information. This effect did not depend on whether the question format in the practice tests (multiple-choice) matched the question format in the final test (multiple-choice versus short-answer). Our results also suggest more clearly than previous research that teachers should not expect that these positive effects extend to untested information from the same course. One practical recommendation based on these findings is that teachers should attempt to cover the central content of their course in a practice test, to maximize the chance that the relevant knowledge lasts.

Restudied contents were also remembered better than contents not covered in the review, which indicates that restudying after a delay also has beneficial effects for long-term retention (see also Rawson & Kintsch, 2005). However, this positive effect was not as large as the effect of practice testing and occurred only in one of the two lectures and only for multiple-choice items in the final test, suggesting that teachers should use quizzing rather than restudy opportunities to support student learning in their courses.

Boundary Conditions of Transfer Effects

Do our results refute the results from previous studies that suggest a transfer effect of testing to untested information? Certainly not. A closer look at the studies providing clear positive evidence for this kind of transfer effect reveals that these studies combined the review session with specific additional tasks (explanation tasks or expectation of inference questions in the final test, Hinze et al., 2013; scoring another participant's practice test, Balch, 1998), or that they used carefully controlled materials to ensure that the untested content had indeed strong semantic associations with the content covered in the practice test (e.g., Chan et al., 2006). In our experiment, tested and untested materials were always associated by being presented within one thematically coherent unit, but the associations were otherwise likely to vary as it is typical for lecture content. Thus, our results do not rule out the possibility of transfer effects of testing to content learned but not practiced. Such effects seem to depend on certain boundary conditions, which should be systematically examined in future research. One such condition is that the strength of the semantic association of information or the coherence of the topics presented in one learning session is crucial for a transfer effect of testing to untested content, as highlighted by the elaborative processing account of the testing effect (Carpenter, 2009) or the notion of retrieval-induced facilitation (Chan et al., 2006). Another condition that could favor transfer effects is the combination of practice testing with additional tasks or instructions that stimulate elaborative processing or generative learning (Richter et al., 2022; Roelle et al., 2022). Such measures have been shown to enhance the long-term effects of other desirable difficulties in authentic educational contexts. For example, Ziegler and Stern (2014) have shown that instructional support that enhances comparing and contrasting enhances the effectiveness of interleaving in mathematics education.

Chapter III

Limitations and Future Directions

Despite the clear pattern of results and the practical implications, the present study also has some limitations. One limitation already stated is that because of the implementation in the context of a regular university lecture, the strength of the semantic associations between different content featured in one lecture session could not be controlled, which limits the theoretical value of the study. Other limitations, include the lack of counterbalancing the order of topics, the varying retention interval (1-6 weeks) for each lecture topic, and the possibility that students engaged in additional learning activities during the study. Importantly, however, none of these methodological characteristics represent a confound that would undermine the internal validity of the experiment and its central conclusions.

In addition, the dropout rate in the present study was high due to the voluntary character of the study and the multiple sessions that were required for completing the experiment. Future research should find ways to keep the rate lower, for example by making study participation a mandatory part of class assignments. Finally, even though we took efforts to maximize generalizability of results, by including two different lectures, which are typical for introductory psychology lectures, and 12 different lecture topics in total, the present effects might still depend on the domain, the complexity of the topic, or didactical features of the lecture. Likewise, the population of teacher trainees at a German university that our sample was based on is relatively heterogeneous regarding their characteristics such as interests and skills, which depend, among other things, on their field of studies. Nevertheless, whether and to what extent our findings generalize to other populations of learners with different backgrounds and different cognitive and motivational prerequisites is unclear. Future studies should address this issue by examining more heterogeneous samples.

Conclusion

In summary, the results of the present study underscore once more the utility that practice tests can have in university teaching for consolidating acquired knowledge. However,

our findings clarify an important limitation of this benefit. The practice tests seem to selectively promote only the retention of explicitly tested content and not the retention of other content presented in the same learning unit that was not tested. To promote comprehensive learning in this context, practice tests arguably need to be supplemented by other instructional measures that promote elaborative processing and the construction of integrated mental representations of the learning content. Psychology instructors are encouraged to include all important topics in the quizzes they provide as learning opportunities for their students.

References

Agarwal, P. K., Nunes, L. D., & Blunt, J. R. (2021). Retrieval practice consistently benefits student learning: A systematic review of applied research in schools and classrooms. *Educational Psychology Review, 33*(4), 1409–1453. https://doi.org/10.1007/s10648-021-09595-9

Anderson, J. R. (1976). *Language, memory, and thought.* Lawrence Erlbaum.

Arnold, K. M., & McDermott, K. B. (2013). Test-potentiated learning: Distinguishing between direct and indirect effects of tests. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39*(3), 940–945. https://doi.org/10.1037/a0029199

Balch, W. R. (1998). Practice versus review exams and final exam performance. *Teaching of Psychology, 25*(3), 181–185. https://doi.org/10.1207/s15328023top2503_3

Bless, H., Wänke, M., Bohner, G., Fellhauer, R. F., & Schwarz, N. (1994). Need for Cognition: eine Skala zur Erfassung von Engagement und Freude bei Denkaufgaben [Need for Cognition: A scale measuring engagement and enjoyment in cognitive tasks]. *Zeitschrift für Sozialpsychologie*, *25*, 147-154.

Butler, A. C. (2010). Repeated testing produces superior transfer of learning relative to repeated studying. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36*(5), 1118–1133. https://doi.org/10.1037/a0019902

Chapter III

Butler, A. C. (2018). Multiple-choice testing in education: Are the best practices for

    assessment also good for learning? *Journal of Applied Research in Memory and*

    Cognition, 7(3), 323-331. https://doi.org/10.1016/j.jarmac.2018.07.002

Carpenter, S. K. (2009). Cue strength as a moderator of the testing effect: The benefits of

    elaborative retrieval. *Journal of Experimental Psychology: Learning, Memory, and*

    Cognition, 35(6), 1563–1569. https://doi.org/10.1037/a0017021

Carpenter S. K. (2011). Semantic information activated during retrieval contributes to later

    retention: Support for the mediator effectiveness hypothesis of the testing effect.

    Journal of experimental psychology. Learning, memory, and cognition, *37*(6), 1547–

    1552. https://doi.org/10.1037/a0024140

Carpenter, S. K., & Yeung, K. L. (2017). The role of mediator strength in learning from

    retrieval. *Journal of Memory and Language, 92*, 128-141.

    https://doi.org/10.1016/j.jml.2016.06.008

Chan, J. C. (2010). Long-term effects of testing on the recall of nontested materials. *Memory,*

    18(1), 49-57. https://doi.org/10.1080/09658210903405737

Chan, J. C. K., McDermott, K. B., & Roediger, H. L. III. (2006). Retrieval-induced

    facilitation: Initially nontested material can benefit from prior testing of related

    material. *Journal of Experimental Psychology: General, 135*(4), 553–571.

    https://doi.org/10.1037/0096-3445.135.4.553

Dunlosky, J., 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

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G* Power 3: A flexible statistical

    power analysis program for the social, behavioral, and biomedical sciences. *Behavior*

    Research Methods, 39(2), 175-191. https://doi.org/10.3758/BF03193146

Glaser, J., & Richter, T. (2023a). The testing effect in the lecture hall: Does it depend on learner prerequisites? *Psychology Learning & Teaching, 22*(2), 159–178. https://doi.org/10.1177/14757257221136660

Glaser, J., & Richter, T. (2023b, October 24). *The testing effect in the lecture hall: Transfer to Untested Content?* [Supplemental online material]. https://osf.io/wc3kh/?view_only=23bc6a041d4a467b87788fbe2cc3bd9b

Greving, S., & Richter, T. (2018). Examining the testing effect in university teaching: Retrievability and question format matter. *Frontiers in Psychology, 9*, Article 2412. https://doi.org/10.3389/fpsyg.2018.02412

Greving, S., Lenhard, W., & Richter, T. (2023). The testing effect in university teaching: Using multiple-choice testing to promote retention of highly retrievable information. *Teaching of Psychology, 50*(4), 332–341. https://doi.org/10.1177/00986283211061204

Hinze, S. R., Wiley, J., & Pellegrino, J. W. (2013). The importance of constructive comprehension processes in learning from tests. *Journal of Memory and Language, 69*(2), 151–164. https://doi.org/10.1016/j.jml.2013.03.002

LaPorte, R. E., & Voss, J. F. (1975). Retention of prose materials as a function of postacquisition testing. *Journal of Educational Psychology, 67*(2), 259–266. https://doi.org/10.1037/h0076933

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*(1), 159–174. https://doi.org/10.2307/2529310

Lenhard, W., & Lenhard, A. (2016). *Computation of effect sizes* [Online tool]. Psychometrica. https://www.psychometrica.de/effect_size.html

McDaniel, M. A., Wildman, K. M., & Anderson, J. L. (2012). Using quizzes to enhance summative-assessment performance in a web-based class: An experimental study. *Journal of Applied Research in Memory and Cognition,1*(1), 18–26. https://doi.org/10.1016/j.jarmac.2011.10.001

Chapter III

Morris, C. D., Bransford, J. D., & Franks, J. J. (1977). Levels of processing versus transfer appropriate processing. *Journal of Verbal Learning and Verbal Behavior, 16*(5), 519–533. https://doi.org/10.1016/S0022-5371(77)80016-9

Nungester, R. J., & Duchastel, P. C. (1982). Testing versus review: Effects on retention. *Journal of Educational Psychology, 74*(1), 18-22. https://doi.org/10.1037/0022-0663.74.1.18

Pan, S. C., & Rickard, T. C. (2018). Transfer of test-enhanced learning: Meta-analytic review and synthesis. *Psychological Bulletin, 144*(7), 710–756. https://doi.org/10.1037/bul0000151

Perkins, D. N., & Salomon, G. (1992). Transfer of learning. In T. Husén & T. N. Postlethwaite (Eds.), *The international encyclopedia of education* (2nd ed., pp. 425-441). Pergamon.

Pilotti, M., Chodorow, M., & Petrov, R. (2009). The usefulness of retrieval practice and review-only practice for answering conceptually related test questions. *The Journal of General Psychology, 136*(2), 179-204. https://doi.org/10.3200/GENP.136.2.179-204

Pyc, M. A., & Rawson, K. A. (2010). Why testing improves memory: Mediator effectiveness hypothesis. *Science, 330*(6002), 335. https://doi.org/10.1126/science.1191465

Rawson, K. A., & Kintsch, W. (2005). Rereading effects depend on time of test. *Journal of Educational Psychology, 97*(1), 70-80. https://doi.org/10.1037/0022-0663.97.1.70

Richter, T., Berger, R., Ebersbach, M., Eitel, A., Endres, T., Borromeo Ferri, R., Hänze, M., Lachner, A., Leutner, D., Lipowsky, F., Nemeth, L., Renkl, A., Roelle, J., Rummer R., Scheiter, K., Schweppe J., von Aufschnaiter, C., & Vorholzer, A. (2022). How to promote lasting learning in schools: Theoretical approaches and an agenda for research. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie/German Journal of Developmental Psychology and Educational Psychology, 54*(4), 135-141. https://doi.org/10.1026/0049-8637/a000258

Roediger, H. L. III, & Butler, A. C. (2011). The critical role of retrieval practice in long-term retention. *Trends in Cognitive Sciences, 15*(1), 20–27. https://doi.org/10.1016/j.tics.2010.09.003

Roelle, J., Schweppe, J., Endres, T., Lachner, A., von Aufschnaiter, C., Renkl, A., Eitel, A., Leutner, D., Rummer, R., Scheiter, K., & Vorholzer, A. (2022). Combining retrieval practice and generative learning in educational contexts: Promises and challenges. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie/German Journal of Developmental Psychology and Educational Psychology, 54*(4), 142–150. https://doi.org/10.1026/0049-8637/a000261

Rowland, C. A. (2014). The effect of testing versus restudy on retention: A meta-analytic review of the testing effect. *Psychological Bulletin, 140*(6), 1432–1463. https://doi.org/10.1037/a0037559

Schwieren, J., Barenberg, J., & Dutke, S. (2017). The testing effect in the psychology classroom: A meta-analytic perspective. *Psychology Learning & Teaching, 16*(2), 179-196. https://doi.org/10.1177/1475725717695149

Veltre, M. T., Cho, K. W., & Neely, J. H. (2015). Transfer-appropriate processing in the testing effect. *Memory, 23*(8), 1229-1237. https://doi.org/10.1080/09658211.2014.970196

Wooldridge, C. L., Bugg, J. M., McDaniel, M. A., & Liu, Y. (2014). The testing effect with authentic educational materials: A cautionary note. *Journal of Applied Research in Memory and Cognition, 3*(3), 214–221. https://doi.org/10.1037/h0101801

Yang, C., Luo, L., Vadillo, M. A., Yu, R., & Shanks, D. R. (2021). Testing (quizzing) boosts classroom learning: A systematic and meta-analytic review. *Psychological Bulletin*, *147*(4), 399–435. https://doi.org/10.1037/bul0000309

Chapter III

Ziegler, E., & Stern, E. (2014). Delayed benefits of learning elementary algebraic

transformations through contrasted comparisons. *Learning and Instruction, 33*, 131-

146. https://doi.org/10.1016/j.learninstruc.2014.04.006

Chapter IV

The Testing Effect in the Lecture Hall: Does Metacognitive

Activation influence its Strength?

Study 3

The Testing Effect in the Lecture Hall: Does Metacognitive Activation influence its Strength?

Julia Glaser

Abstract. One of the most effective ways to learn sustainably is practice testing, as has been proven in many studies. But perhaps its effectiveness depends on how students use it. Above all, active, intentional use could be beneficial - metacognitive activation of students during learning could make a positive contribution to this. Within this study, we investigated the question if minimal metacognitive activation of learners influences the strength of the testing effect, especially when learners did not yet know the learning strategy of testing. We conducted an experimental intervention in psychology courses for student teachers ($N = 157$). One week before the lecture, the main learners' characteristics were recorded. Immediately after the lecture, participants completed an online review with multiple-choice questions (practice test with corrective feedback and optional revision) or summary statements (restudy), alternating within each participant. Half of the participants received an instruction at the beginning of the practice test, which was intended to serve metacognitive activation, and reminder prompts during the session. One week later, retention was assessed with a criterial test of short-answer and multiple-choice questions. A testing effect emerged ($\eta_p^2 = .136$), with better retention for the tested compared with the restudied contents. No effect of the metacognitive activation could be observed, neither for participants knowing the testing effect nor for those not knowing. Individual learner characteristics also did not affect the testing effect. The results suggest that the testing effect is a robust phenomenon whether one is aware of its positive learning effect or not. It can – and should - therefore be easily integrated into learners' learning practices without having to explain it in detail.

The Testing Effect in the Lecture Hall: Does Metacognitive Activation influence its Strength?

Self-regulated learning plays an important role in today's educational world. Especially in online environments where there is no instructor present, it is important that students pick the right strategies and use them correctly to learn successfully and sustainably (Wong et al., 2019). One promising strategy for sustained learning is test-enhanced learning. Consisting of two phases – initial study and practice test – it has been shown to be an effective and efficient way to foster long-term retention (Roelle et al., 2022; Rowland, 2014; Yang et al., 2021). Test-based learning, even in real-world educational contexts, has been well researched and is usually studied with a similar structure: students attend a course, learn the information there, then take a timely practice test, and are tested a while later to see how much they still know. Rereading is usually used as the control group (Roediger & Butler, 2011). As an illustration, Glaser and Richter (2023) conducted an experiment involving students enrolled in introductory psychology lectures. In this study, half of the lecture content was provided to the students in the form of short-answer questions accompanied by corrective feedback (referred to as the practice test), while the other half was presented through summarizing statements (referred to as restudy). These review sessions, involving both practice testing and restudying, were administered online during the week following the lecture session. Across five distinct lecture topics, a positive testing effect became evident when assessing the students' performance in a final test conducted one week after the initial learning session. Notably, the content that had been delivered through short-answer questions was retained more effectively compared to the content presented through summarizing statements, demonstrating a medium-sized effect (Cohen's $d = 0.55$).

This positive testing effect has been found in a variety of studies, conducted with participants of different age-groups, with different kinds of study materials and in laboratory

($g = 0.50$, Rowland, 2014) as well as in educational contexts ($g = 0.33$; Yang et al.. 2021; see also Agarwal et al., 2021). The testing effect is mediated by two mechanisms: the indirect and the direct testing effect (Arnold & McDermott, 2013). The indirect testing effect can be enhanced by corrective feedback, as this reveals knowledge gaps and thus influences future learning activities. In the absence of feedback, the direct testing effect works best when the learned content has high retrievability (Greving et al., 2022; Greving & Richter, 2018). Repeated testing (Butler, 2010), and especially with challenging questions that encourage active recall (Butler, 2018), favors the testing effect.

Although the testing effect and also some moderators are already well researched, there are still gaps. The effectiveness of practice tests as a learning strategy that promotes sustained learning may depend on how learners use them (Soderstrom & Bjork, 2014). In analogy to other learning strategies, an active, intentional use could be particularly effective, in which learners invest more cognitive resources, reflect on the practice questions and thus retrieve more content from long-term memory than with a more passive or incidental use (see also the concept of "constructive retrieval", Hinze, Wiley & Pellegrino, 2013). If this reasoning is correct, metacognitive activation that makes learners aware of the potential benefits of practice tests and the underlying psychological mechanisms should further enhance the testing effect. Some studies have generally found that metacognitive activation in the form of prompts has positive effects on the effectiveness of learning strategies (e.g. Biwer et al., 2020; Schwonke et al., 2013) but these are mostly more general studies looking at different learning strategies.

The present field experiment replicated the study by Glaser and Richter (2023) discussed above and extended it to provide a clarification of this question in a typical lecture setting. To ground the research question of our study, we briefly review theoretical explanations of the testing effect that suggest metacognitive activation might influence the

strength of testing effect. We will then discuss the available empirical studies that have addressed this phenomenon in more detail.

Metacognitive Activation and its importance for Learning

People often have a deficient mental model of how they learn and recall, which leads them to both wrongly assess and misalign their own learning. Often due to a lack of education in learning strategies, unfavorable and flawed beliefs about what constitutes effective learning thus develop over a period of years (Bjork et al., 2013). The most common wrong belief about learning is, that it has to be easy, meaning that one gets a feeling of fluency (Biwer et al., 2020). Research shows that for effective and sustainable learning, the opposite is true: learning has to be challenging (as known as desirable difficulties, Bjork & Bjork, 2011).

But even if students know about effective strategies, they often do not execute them spontaneously (Bannert, 2009). The reason for this is usually the additional learning effort associated with the strategy and the benefits, which are considered to be low. In some cases, students also lack the necessary resources (like time or transfer knowledge) to implement the strategy (Foerst et al., 2017). This leads to a worse performance then possible. Therefore, knowledge about learning strategies and their benefits as well as metacognitive activation of these in the learning phase are important. Biwer and colleagues (2020) showed that a training of learning strategies improved the metacognitive knowledge of the students, lead to more spontaneous use of effective strategies and less dysfunctional beliefs about the effectiveness of strategies (Biwer et al., 2020). When learned, the strategies have to be present in the moment of learning. For this, prompting can be a useful aid to activate the metacognitive knowledge. Berthold and colleagues found out that especially cognitive prompts (*directly supporting a student's processing of information*; Bannert, 2009) or a combination of cognitive and metacognitive prompts (*supporting a student's monitoring and control of their information*; Bannert, 2009) activates learning strategies and can lead to improved learning outcomes (Berthold et al., 2007). And Schwonke and colleagues (2013) conducted a study in

an online environment in which they either gave subjects info on learning strategies or they did not. The information, presented as prompts, increased learning efficiency, for example by saving time, and helped people without prior metacognitive knowledge to gain a deeper understanding of the strategy (Schwonke et al., 2013). So even a minimal prompt, as used by Schwonke and colleagues (2013), could potentially strengthen the effect of good learning strategies like the testing effect.

Rationale of the Present Study

The present study examines the extent to which minimal metacognitive activation affects the strength of the testing effect. This question is assessed by incorporating practice tests into optional retrieval activities offered as a supplement to mandatory psychology lectures for student teachers in universities. Consistent with the majority of research on retrieval practice and given its prevalence in self-regulated learning activities for university students, repetition was used as a control condition. To maximize experimental control and statistical power, the comparison between practice and repetition within subjects and within each lecture topic was conducted using alternating questions and summary statements. Consistent with previous theory and research on recall practice, we hypothesized that there is a significant testing effect of the learning condition with better outcomes in the criterial test for items of the testing condition and that this effect would be robust across lecture topics.

Further, participants who received the metacognitive activation instruction should yield a stronger testing effect than those who didn't. Especially those participants within the experimental group, who reported not knowing about the effectiveness of testing, should score higher than the ones knowing about it.

Additionally, we pursed a range of individual differences (NFC, Positive and Negative Affect, Prior knowledge) as exploratory research questions regarding their influence on the testing effect.

Method

Participants

Participants were 157 students enrolled in the teacher-training program at the University of Würzburg. 29 students studying to become teachers for elementary school (Grundschule), 10 for non-academic track middle-school teachers (Mittelschule), 5 for secondary school (Realschule), 60 for high school (Gymnasium), 41 for special education, 6 are majoring in pedagogy, and 1 is studying to become a speech therapist; 5 did not specify their field of study.

Participants were recruited from two different courses (Behavioral and Learning Disorders in Childhood and Youth, Development in Childhood and Youth), all of which are part of the mandatory psychology curriculum for prospective teachers. They could choose to participate in one of six lecture topics that were part of the three courses: Reading and Writing Problems, Mathematic Problems, Attentional Deficit Hyperactivity Disorder, Development of Thinking, Development of Intelligence, and Development of Memory.

Of the 194 students who initially signed up for the study, only 157 completed all parts of the study. The other 37 students were excluded from the analysis because of incomplete data. Data for the study were collected in two waves in subsequent semesters (November/December 2022, $n = 53$, April to June 2023, $n = 104$). The distribution of participants between topics, waves, and parts is shown in Table III.1.

Table III.1

Number of Participants per Lecture Topics in Waves 1 and 2 and Attrition During the Experiment

Topic	Part 1		Part 2		Part 3		Valid (cases)
	Wave 1	Wave 2	Wave 1	Wave 2	Wave 1	Wave 2	
Reading and writing problems	18	29	15	26	15	26	41
Mathematic problems	8	23	7	21	6	20	26
Attentional Deficit Hyperacticity Disorder	13	47	12	36	12	34	46
Development of Thinking	9	8	7	7	7	6	13
Development of Intelligence	6	12	5	10	5	8	13
Development of Memory	10	11	8	10	8	10	18
Total	64	130	54	110	53	104	
	194		164		157		157

Note. Wave 1 took place in the winter semester 2022/2023. Wave 2 took place in summer semester 2023. Part 1 included the assessments of learner characteristics. Part 2 was the learning phase and Part 3 the criterial test. Valid cases correspond to participants with complete data that were included in the analyses.

The majority of participants was female (80.9 %) and their age ranged from 18 to 42 years ($M = 20.66$, $SD = 3.300$). Only six participants reported another language than German as their first language. All participants received course credit for their participation. Table III.2 shows the characteristics of the sample in detail.

Table III.2

Details of the sample's characteristics

	Age	Grade	NFC	Prior Knowledge	PANAS practice	PANAS criterial	PANAS general
M	20.66	2.06	3.65	2.36	3.74	3.68	3.64
SD	3.30	0.52	0.71	0.87	0.43	0.41	0.43
Min	18	1.00	2.00	1.00	2.30	2.65	2.15
Max	42	3.30	5.25	5.00	4.85	4.55	4.40

Experimental Study Materials

We analyzed the contents of the lecture session for each topic and extracted 15 information units for each topic. Within the repetition session, participants were alternately presented the 15 information units either as a statement (restudy), which summarized the main ideas of the information unit and which participants could restudy for maximum 60 seconds or as a multiple-choice question (practice testing) asking for the main idea for which they also had a time limit of 60 seconds to fill it out. In both cases if they wanted, they could go on to the next item faster. Each question contained four response options, each of which could be correct or incorrect and assessed with .25 points for correctly ticking. After providing their answer, participants were shown the correct answer as feedback. A sample multiple-choice question and corresponding statement read as follows:

Multiple-Choice Question: In which order do the stages of writing acquisition occur? Logographic phase -> alphabetic phase -> orthographic phase; Alphabetic phase -> logographic phase -> orthographic phase; Orthographic phase -> alphabetic phase -> logographic phase; Alphabetic phase -> orthographic phase -> logographic phase

Statement: Stages of writing acquisition in order: logographic stage -> alphabetic stage -> orthographic stage

Half of the participants received a minimally modified instruction to activate their metacognitive awareness of learning strategies. Therefore, the usual instruction informing participants about the procedure was supplemented with the reference to the effectiveness of testing and the request to seriously learn with these materials. Figure III.1 shows a comparison of the two instructions. The experimental study materials are available in the Appendix.

Figure III.1

Comparison of instructions (between-manipulation)

Instruction without metacognitive activation

In the following, you will be asked **15 multiple choice questions or statements** on the topic "Attention and concentration" of the lecture "Abnormalities in the experience and behavior of children and adolescents".

You have **max. 1 minute processing time / reading time** per question - after that you will be automatically forwarded to the next question. You can also continue on your own by clicking the Next button. Jumping back is not possible.

Read the statements carefully and remember their contents. You are already learning for the exam!
Read the questions carefully and mark all correct answer alternatives.
At least one alternative is correct; all of them could also be correct.

After the question, you will receive feedback on whether your answer was correct and, if you answered incorrectly or incompletely, what the correct answers would have been. Questions that you answered incorrectly will be asked again at the end of the questionnaire. Some questions will be asked until you have answered them correctly.

The processing time will be about 15min. Answer the questions **without any aids** - this is the only way to have a learning effect for the exam.

At the end of the questionnaire, you can expect some organizational questions, for which you will need approx. 3min.

Have fun & success!

Instruction with metacognitive activation

In the following, you will be asked **15 multiple choice questions or statements** on the topic "Attention and concentration" of the lecture "Abnormalities in the experience and behavior of children and adolescents".

You have **max. 1 minute processing time / reading time** per question - after that you will be automatically forwarded to the next question. You can also continue on your own by clicking the Next button. Jumping back is not possible.

Read the statements carefully and remember their contents. You are already learning for the exam!
Read the questions carefully and mark all correct answer alternatives.
At least one alternative is correct; all of them could also be correct.

After the question, you will receive feedback on whether your answer was correct and, if you answered incorrectly or incompletely, what the correct answers would have been. Questions that you answered incorrectly will be asked again at the end of the questionnaire. Some questions will be asked until you have answered them correctly.

Important note: learning with questions and feedback is an **effective learning strategy** that promotes sustainable recall[1]. Think carefully about the questions and take the time to read the feedback carefully as well. The more you think about the necessary and also related knowledge, the more solid the memory of it will be and you will be able to recall it more easily in an exam situation. So it pays to read carefully and think hard.

Review question: Did you already know that self-testing is an effective learning strategy? (Yes / No)

The processing time will be about 15min. Answer the questions **without any aids** - this is the only way to have a learning effect for the exam.

At the end of the questionnaire, you can expect some organizational questions, for which you will need approx. 3min.

Have fun & success!

1) Dunlosky, J., 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

Note. Instructions were presented within the practice session right before starting the task. At the end of the instruction, participants had to indicate whether they were already aware of the testing effect. 85.4% stated that they were aware of it.

Measures

Criterial Test

The dependent variable was the performance in a criterial test with a total of 15 questions to the chosen lecture topic, all of them seen before in the practice test.

The criterial tests were either presented with multiple-choice questions ($N = 81$) or with short answer questions ($N = 76$). The distribution of participants among the question format was random and balanced within the six topics.

The multiple-choice questions were scored by partial credit (.25), given for each response option that was correctly ticked or not correctly not ticked. The short answer questions were also scored according to a partial credit scheme (0-1 points in intervals of .25). Two raters independently coded the responses from 32 participants (15 questions for each participant) to estimate inter-rater reliability of the short-answer questions. An overall value of Cohen's κ was .798 (SE = .056), indicating substantial agreement according to Landis and Koch (1977). The values for every topic can be seen in Table III.3. Again, given the high inter-rater reliability, the remaining answers to the short answer questions were coded by only one coder.

Table III.3

Values of Cohen's κ ordered by lecture and topics

Lecture	Topic	Cohen's κ	SE
Behavioral and Learning Disorders	Reading and writing Problems	.870	.043
	Maths Problems	.777	.055
	ADHD	.881	.045
	Total	.843	.048
Development in Childhood and Youth	Development of Thinking	.726	.072
	Development of Intelligence	.793	.058
	Development of Memory	.741	.063
	Total	.753	.064
Overall		.798	.056

Learner Characteristics

Need for Cognition. Need for Cognition was assessed with the German short form of the Need for Cognition Scale (NFC-K; Beißert et al., 2015). The scale contains 4 statements concerning thinking preferences, which had to be rated on a seven-point Likert scale.

Positive and Negative Affect. Positive and Negative Affect was assessed with the German Version of the Positive and Negative Affect Schedule (PANAS; Breyer & Bluemke, 2016). Participants were presented 10 items measuring positive (e.g., excited, inspired) and 10

measuring negative affect (e.g., upset, afraid) and should rate them on a five-point Likert scale the extent to which this emotion is currently occurring (1 very slightly – 5 extremely).

Qualitative Query of Prior Knowledge. To assess how much participants already know about the topics asked, they were asked to provide qualitative ratings to judge their level of knowledge. Therefore, five-point scales with the options *nothing, hardly, a little, much, intensive* were offered and the participants had to tick for their chosen topic how much they have dealt with.

Demographic Data. Further, participants were asked to provide their demographic data. This included gender, age, mother tongue, high school graduation grade as a performance indicator, and field of study (type of school, combination of subjects). To be able to assign the persons to the correct mail distribution list, they were asked for their seminar group and to give their mail address. A later allocation of these data to a name or a mail address was prevented by separate storage.

Data to ensure internal validity. Finally, some measures to ensure the internal validity of the study, were assessed. Therefore, participants were asked to indicate whether they used any auxiliary means in completing the practice test or the criterial test, which aids they used, and to what extent. Further, they were also asked whether and, if so, how long they had been studying additionally before the practice test or between the two tests. Finally, they had to rate how distracted they were during the practical test and the critical test. For this they could mark on a scale from 1 - hardly - to 5 - strongly - what they experienced. Based on this, the students were divided into 2 groups: those who were not very distracted (1-3) and those who were very distracted (4-5).

Design

We investigated the effect of the independent variable learning condition (restudy or testing) on the dependent variable performance in the critical test. Special about the testing was the within-variation of the learning condition and the between-variation of the

metacognitive awareness (normal vs. increased by instruction). Each participant was randomly assigned to one of the groups by filling in the pre-survey.

Procedure

Data collection for the study took part from October 2022 until December 2022 and from May 2023 until July 2023 and processed via the online system SoSciSurvey (www.soscisurvey.de).

Participants had to sign up for the study on university's registration system SONA. With their registration, they received the link for the pre-survey, where they first had to choose the lecture and topic they wanted to participate in. Afterwards they had to fill in a questionnaire containing their demographic data, the Need for Cognition questionnaire, the PANAS questionnaire and the qualitative query of their prior knowledge. By giving their mail address, they automatically registered for the upcoming parts of the study, for which they received the links via the named mailing.

Then the experimental phase began. Students attended their lecture as usual. Directly after the chosen thematic lesson, participants received their first link with the task. Half of the participants now received the normal instruction for the task, the other half was presented with the special instruction with focus on metacognitive activation. Then all participants were asked to read the statements or answer in the multiple-choice questions to recapture the knowledge learned. If answered incorrectly, the questions had to be repeated later up to a maximum of five times. Participants within the special instruction group received a reminder to recapture knowledge seriously after every fifth item.

The last part of the study took place one week after the lecture respectively the learning phase. Participants now received a questionnaire with 15 questions covering all information units of the lecture session. The questionnaire was either designed with multiple-choice questions or short answer questions, varied, and balanced between participants and topics. The questions were all presented either as a statement or a question during the learning

phase. After filling in the questionnaire, feedback regarding some points of the organization of the study, like its level of comprehensibility and usefulness, was obtained from the participants and they had to indicate whether they used other learning methods than in the study during the experimental phase.

After completing all parts of the intervention, the students received their course credit. A visual presentation of the process can be seen in Figure III.2.

Figure III.2

Visual Presentation of Design and General Procedure.
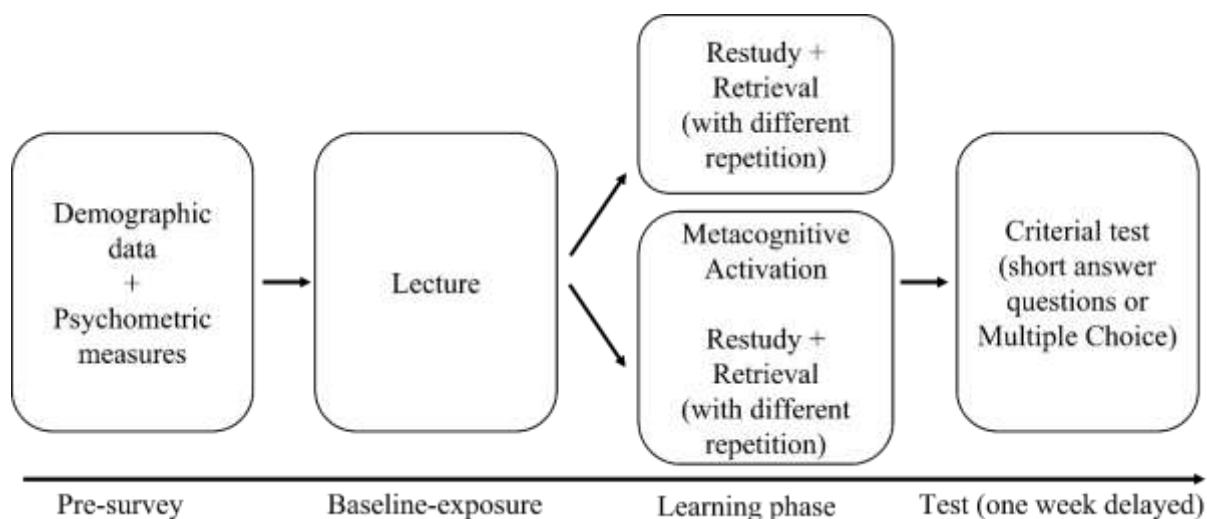
Note. Using a mixed design, participants decided for one topic of a module (between-subjects) and run the same procedure: demographic data and psychometric measures were followed by the baseline-exposure as the normal lecture. Directly afterwards the learning phase took part using either retrieval practice or restudy (within-subjects), alternating for each information unit. Retrieval practice means therefore practice questions with the option to repeat wrongly answered questions. A between-subjects manipulation of the instruction varied the metacognitive activation of participants. One week after the last lecture, there was a test covering all information units presented in the learning phase. The test was either presented with short answer or multiple-choice questions.

Results

We first estimated a linear model with the learning condition as the experimental factor (varied within-subjects) and the lecture topics (varied between-subjects) and their interactions. In a second step, we estimated a linear model with the instruction as the experimental factor (varied between-subjects) and its interaction with the learning condition

and topic. In a final step, we estimated a series of linear models that additionally included each of the learner characteristics as a continuous predictor and its interaction with the learning condition (Judd et al., 2001) to explore the role of learner characteristics. We performed separate analyses for each learner characteristic to maximize the chance to detect a moderator effect. We used an α-level of .05 for all statistical tests. We report partial η^2 as the effect size measure. All data files and analysis scripts can be found in the Appendix and in the repository of the Open Science Framework (https://osf.io/ncvy9/?view_only=83c6e1cf2c4c442bb16f4ec8de4a9a7b).

Main Effect of Testing

The model with learning conditions (with no learner characteristics) revealed better learning outcomes in the criterial test for testing ($M = .658$, $SD = .23$) compared to restudying ($M = .584$, $SD = .22$), $F(1, 155) = 24.431$, $p < .001$, $\eta_p^2 = .136$. Therefore, Hypothesis 1 that predicted a positive testing effect was supported. The interaction of testing and modules was not significant, $F(1, 155) = .226$, $p = .635$, $\eta_p^2 = .001$, suggesting that the testing effect was independent of materials and contents. Figure III.3 displays the average testing effect across and within modules.

Effect of the Metacognitive Instruction

The interaction between learning condition and instruction was not significant, $F(1, 154) = .081$, $p = .776$, $\eta_p^2 = .001$, arguing against Hypothesis 2. Figure III.4 displays the average testing effect for both study groups.

Figure III.3

The testing effect across and within modules
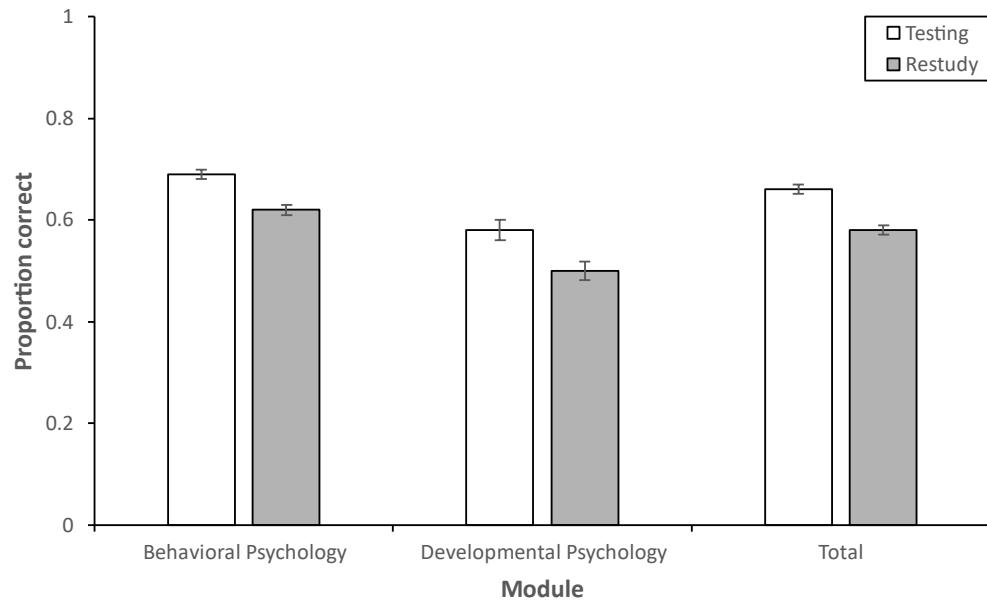
Note. Error bars represent the standard error of the mean. The y-axis shows the proportion of correct responses in the criterial test.

Figure III.4

Interaction of Learning Condition and Experimental Condition
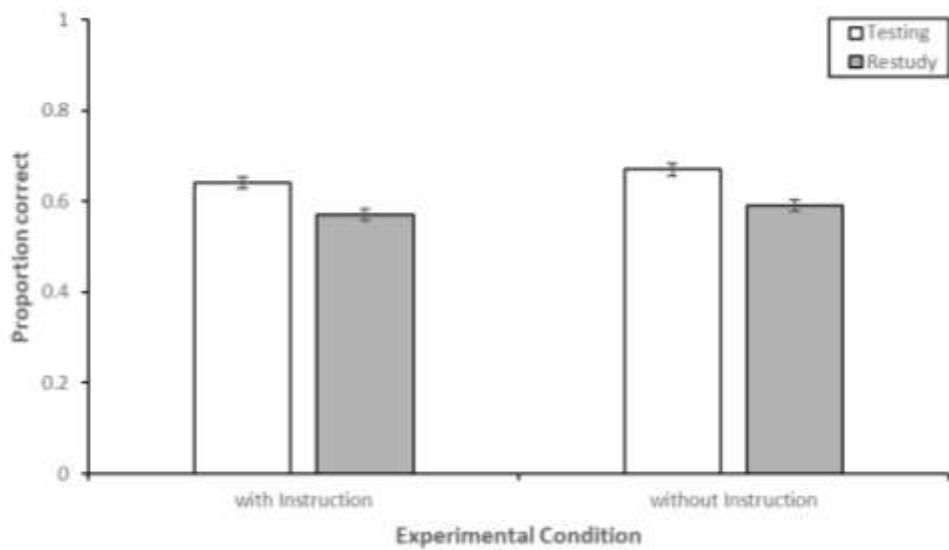
Note. Error bars represent the standard error of the mean. The *y*-axis shows the proportion of correct responses in the criterial test.

Effects of Metacognitive Knowledge and Learner Characteristics

There is no significant difference in efficacy of testing for people who reported to know the testing effect vs. people who did not at first ($F(1, 79) = .000$, $p = .986$, $\eta_p^2 = .000$). Therefore, Hypothesis 3 that predicted a stronger testing effect for people with less metacognitive knowledge was not supported. Table III.4 shows the interactions concerning metacognitive activation.

Table III.4

Overview of all main effects and interactions concerning Metacognitive Activation

Effect	df	F (1, df)	p	η_p^2
Instruction	155	.490	.485	.011
Learning Condition *Instruction	154	.081	.776	.001
Metacognitive Knowledge	80	.008	.931	.000
Learning Condition * Metacognitive Knowledge	79	.000	.986	.000

None of the interactions between learning condition and any of the learner characteristics was significant (Table III.5). Thus, we found no evidence that the testing effect depends on learner characteristics.

Table III.5

Overview of all main effects and interactions concerning learner characteristics

Effect	df	F (1, df)	p	η_p^2
Need for Cognition	156	.607	.437	.013
Learning Condition * Need for Cognition	155	.975	.325	.006
Grade	145	.296	.587	.002
Learning Condition * Grade	144	.003	.957	.000
Prior Knowledge	146	.041	.840	.000

Learning Condition * Prior Knowledge	145	.035	.851	.000
PANAS Practice	155	3.282	.072	.038
Learning Condition * PANAS Practice	154	.038	.845	.000
PANAS General	155	.000	.997	.000
Learning Condition * PANAS General	154	.548	.460	.004
PANAS Criterial	155	.559	.456	.012
Learning Condition * PANAS Criterial	154	.003	.958	.000

Measures on distraction, auxiliary tools, and additional leaning

To monitor the general conditions of the study, some control measures were collected. Students Learning activity before the learning phase and between practice testing and criterial testing was assessed. There was no effect neither of pre-learning ($p = .490$), which 6.4% of the students did, nor of learning between the tests ($p = .361$), which 26.8% of students did. Also, the use of auxiliary means was controlled. 3.2% of study participants reported using auxiliary means during the practice testing. However, this had no significant influence on the results ($p = .799$). The situation was similar for the criterial tests: 7.0% used aids, but there was no significant difference from the others ($p = .846$). A last measure we assessed was the felt distraction. Students were (as described in the method section) either less or more distracted. 28.4% of participants in the practice test and 20.8% participants in the criterial test reported being highly distracted. Nevertheless, there was no difference between the severely and the less severely leveled individuals, neither in the practice test ($p = .559$), nor in the criterial test ($p = .184$). Given these results, internal validity of the study can be confirmed. Details can be seen in Table III.6.

The Testing Effect in the Lecture Hall: Does Metacognitive Activation influence its Strength?

Table III.6

Control measures in detail

Effect	df	$F(1, df)$	p	η_p^2
Pre-Learning * Learning Condition	154	.478	.490	.003
Further Learning * Learning Condition	152	.841	.361	.006
Auxiliary in Practice * Learning Condition	155	.065	.799	.000
Auxiliary in Criterial * Learning Condition	150	.038	.846	.000
Distraction while Practice * Learning Condition	153	.342	.559	.002
Distraction while Criterial * Learning Condition	147	1.785	.184	.012

Discussion

The present study examined whether short practice tests would benefit long-term retention compared to restudying the information and whether this effect would be strengthened by metacognitive awareness of the effectiveness of this learning strategy. Learning outcomes were assessed after a retention interval of 1 week after learning. We created a minimal intervention in two regular university lectures, with three topics each and a within-participant and within-topic variation of practice testing versus restudying. The main finding was that a testing effect occurred, regardless of the instruction for metacognitive awareness. We also did not find a difference between people already knowing about the testing effect and its effectiveness and people not knowing about it. This result suggests that the testing effect is an effective means of enhancing long-term memory without incorporating metacognitive knowledge.

These results contribute to the literature on real-life educational practices by providing an experimental test of the testing effect in the university classroom using actual study materials. Even if there is already much literature on testing in the classroom context (Yang et al., 2021, included 225 effects in their meta-analysis), this study brings new aspects, as it is a

true experimental study (compared to many quasi-experimental studies), compared testing to restudy (instead of the comparison to "no learning", as in other studies) and was broadly thematic to ensure transferability to other topics. Further on, it is a within-design so that any confounds of learning condition and topic can be ruled out.

A novel research question addressed by the present experiment was to explore whether metacognitive awareness of the effectiveness of testing would moderate the strength of the testing effect. We did not find any effect of our instruction for metacognitive activation on the strength of the testing effect, neither for persons knowing nor for persons not knowing about its effectiveness. These findings are in contradiction with previous research on metacognitive activation and its effect on learning strategy use (e.g. Biwer et al., 2020; Schwonke et al., 2013). But do our results invalidate the results of others? Probably not. The other studies are structured differently, have different target groups, and are used in different contexts and with different strategies. For example, Schwonke and colleagues' study (2013) was conducted with high school students (Schwonke et al., 2013). It can be assumed that they have a significantly lower knowledge of learning strategies and therefore the effectiveness of the intervention is quite different than in our study, where the knowledge is already widely spread. And Biwer and colleagues (2020) used an intervention program spread over several weeks with a much greater intensity than our minimal intervention (Biwer et al., 2020). It is not excluded that metacognitive activation has an influence on the testing effect. However, there seems to be no effect under the general conditions we chose in this study.

However, it is also conceivable that the implementation of the practice tests chosen here with multiple testing in the case of incorrect answers, and corrective feedback directly after the test question already made the testing so effective that the metacognitive activation in the sense of a ceiling effect could no longer have any additional benefit. This is indicated by the very strong testing effect that was found. Without mandatory repetition of the incorrectly answered questions, people without metacognitive activation would probably

benefit less from the test and group differences would possibly emerge. Future research could implement metacognitive activation in a design where multiple testing is excluded, and feedback is given in a self-regulated way at the end of the session to give any effects greater scope to unfold.

Limitations of the Present Study

Despite its informative results, our study also has limitations in need of discussion. A limitation, but at the same time the greatest strength of the study, is that it took place in the classroom. This makes the conditions less controlled and less actually detectable than in a lab study. For example, it is not certain that the students really did not use any auxiliary tools or engage in any additional learning activities. Also, the environment of the students, which was not standardized, could influence the results of the study. As we saw in the results, some students did use auxiliary means or felt very distracted during the participation. There were no significant differences and internal validity of the study was given, but some results descriptively showed that it would negatively affect the testing effect. Therefore, this could be one of the reasons why the intervention did not work as well as expected.

Another limitation could be that the intervention is too small. One simple statement about the effectiveness of testing is eventually too little or the manner of presentation is too inconspicuous, so students may not have perceived this thoroughly. There was a hurdle that it couldn't just be clicked on, but students are experienced in filling out surveys while only skimming. In addition, the testing effect is a very famous phenomenon widely known among students, so only few (env. 15%) did not know about it before. If the information is so widespread, participants may already be using the testing effect in their everyday learning anyway and show less strong effects here than in groups that have less experience with the testing effect.

And lastly, generalizability of the study is not fully given due to the features of the sample. Participants were recruited from student teachers at the University of Würzburg from

two courses. It can be assumed that student teachers differ less in personality, interests, intelligence, and other characteristics than the cross-section of society. In addition, they all have a similar amount of experience with learning, since students are expected to meet higher standards than other groups in society, and they have also already dealt with what effective learning strategies are in their psychology studies. Another sample could possibly raise other results, so the study should be repeated with a more diverse and potentially less experienced group of learners.

Conclusion and Further Directions

In summary, the results of this study emphasize the usefulness of the testing effect even without much instruction. This makes it easier for teachers to integrate the strategy of testing into their students' lessons or home learning without much effort. The strategy of testing should be used as a helpful strategy wherever possible.

References

Agarwal, P. K., Nunes, L. D., & Blunt, J. R. (2021). Retrieval practice consistently benefits student learning: A systematic review of applied research in schools and classrooms. *Educational Psychology Review, 33*(4), 1409–1453. https://doi.org/10.1007/s10648-021-09595-9

Arnold, K. M., & McDermott, K. B. (2013). Test-potentiated learning: Distinguishing between direct and indirect effects of tests. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39*(3), 940–945. https://doi.org/10.1037/a0029199

Bannert, Maria (2009): Promoting Self-Regulated Learning Through Prompts. *Zeitschrift für Pädagogische Psychologie 23*(2), 139-145. https://doi.org/10.1024/1010-0652.23.2.139

Beißert, H., Köhler, M., Rempel, M., & Beierlein, C. (2015). Deutschsprachige Kurzskala zur Messung des Konstrukts Need for Cognition NFC-K. *Zusammenstellung sozialwissenschaftlicher Items und Skalen (ZIS)*. https://doi.org/10.6102/zis230

Berthold, K., Nückles, M., & Renkl, A. (2007). Do learning protocols support learning strategies and outcomes? The role of cognitive and metacognitive prompts. *Learning and Instruction*, *17*(5), 564-577. https://doi.org/10.1016/j.learninstruc.2007.09.007

Biwer, F., Egbrink, M. G. A. o., Aalten, P., & de Bruin, A. B. H. (2020). Fostering effective learning strategies in higher education—A mixed-methods study.*Journal of Applied Research in Memory and Cognition, 9*(2), 186–203. https://doi.org/10.1016/j.jarmac.2020.03.004

Bjork, E. L., & Bjork, R. A. (2011). Making things hard on yourself, but in a good way: Creating desirable difficulties to enhance learning. In M. A. Gernsbacher, R. W. Pew, L. M. Hough, & J. R. Pomerantz (Eds.), FABBS Foundation, psychology and the real world: Essays illustrating fundamental contributions to society (pp. 56–64). Worth Publishers.

Bjork, R. A., Dunlosky, J., & Kornell, N. (2013). Self-regulated learning: Beliefs, techniques, and illusions. *Annual Review of Psychology*, *64*, 417-444. https://doi.org/annurev-psych-113011-143823

Breyer, B., & Bluemke, M. (2016). Deutsche Version der Positive and Negative Affect Schedule PANAS (GESIS Panel). *Zusammenstellung sozialwissenschaftlicher Items und Skalen (ZIS)*. https://doi.org/10.6102/zis242

Butler, A. C. (2010). Repeated testing produces superior transfer of learning relative to repeated studying. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36*(5), 1118–1133. https://doi.org/10.1037/a0019902

Chapter IV

Butler, A. C. (2018). Multiple-choice testing in education: Are the best practices for

assessment also good for learning? *Journal of Applied Research in Memory and

Cognition, 7*(3), 323-331. https://doi.org/10.1016/j.jarmac.2018.07.002

Foerst, N. M., Klug, J., Jöstl, G., Spiel, C., & Schober, B. (2017). Knowledge vs. Action:

Discrepancies in University Students' Knowledge about and Self-Reported Use of

Self-Regulated Learning Strategies. *Frontiers in Psychology, 8,* 1288.

https://doi.org/10.3389/fpsyg.2017.01288

Glaser, J., & Richter, T. (2023). The testing effect in the lecture hall: Does it depend on

learner prerequisites? *Psychology Learning & Teaching, 22*(2), 159–178.

https://doi.org/10.1177/14757257221136660

Greving, S., & Richter, T. (2018). Examining the testing effect in university teaching:

Retrievability and question format matter. *Frontiers in Psychology, 9*, 2412.

https://doi.org/10.3389/fpsyg.2018.02412

Greving, S., Lenhard, W., & Richter, T. (2023). The testing effect in university teaching:

Using multiple-choice testing to promote retention of highly retrievable information.

Teaching of Psychology, 50(4), 332–341. https://doi.org/10.1177/00986283211061204

Hinze, S. R., Wiley, J. & Pellegrino, J. W. (2013). The importance of constructive

comprehension processes in learning from tests. *Journal of Memory and Language,

69*(2), 151–164. https://doi.org/10.1016/j.jml.2013.03.002

Judd, C. M., Kenny, D. A., & McClelland, G. H. (2001). Estimating and testing mediation and

moderation in within-subject designs. *Psychological Methods, 6*(2), 115–134.

https://doi.org/10.1037/1082-989X.6.2.115

Roediger, H. L. III, & Butler, A. C. (2011). The critical role of retrieval practice in long-term

retention. *Trends in Cognitive Sciences, 15*(1), 20–27.

https://doi.org/10.1016/j.tics.2010.09.003

Roelle, J., Schweppe, J., Endres, T., Lachner, A., von Aufschnaiter, C., Renkl, A., Eitel, A., Leutner, D., Rummer, R., Scheiter, K., & Vorholzer, A. (2022). Combining retrieval practice and generative learning in educational contexts: Promises and challenges. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie/German Journal of Developmental Psychology and Educational Psychology, 54*(4), 142–150. https://doi.org/10.1026/0049-8637/a000261

Rowland, C. A. (2014). The effect of testing versus restudy on retention: A meta-analytic review of the testing effect. *Psychological Bulletin, 140*(6), 1432–1463. https://doi.org/10.1037/a0037559

Schwonke, R., Ertelt, A., Otieno, C., Renkl, A., Aleven, V., & Salden, R. J. C. M. (2013). Metacognitive support promotes an effective use of instructional resources in intelligent tutoring. *Learning and Instruction, 23*, 136–150. https://doi.org/10.1016/j.learninstruc.2012.08.003

Soderstrom, N. C. & Bjork, R. A. (2014). Testing facilitates the regulation of subsequent study time. *Journal of Memory and Language*, *73*(1), 99–115. https://doi.org/10.1016/j.jml.2014.03.003

Wong, J., Baars, M., Davis, D., Van Der Zee, T., Houben, G. J., & Paas, F. (2019). Supporting self-regulated learning in online learning environments and MOOCs: A systematic review. *International Journal of Human–Computer Interaction, 35*(4-5), 356-373. https://doi.org/10.1080/10447318.2018.1543084

Yang, C., Luo, L., Vadillo, M. A., Yu, R., & Shanks, D. R. (2021). Testing (quizzing) boosts classroom learning: A systematic and meta-analytic review. *Psychological Bulletin*, *147*(4), 399–435. https://doi.org/10.1037/bul0000309

Chapter V

The Testing Effect in the Lecture Hall: How do Students judge this

Learning Method?

Additional Evaluation of Participants' Feedback and Practical
Implications

The Testing Effect in the Lecture Hall: How do Students judge this Learning Method?

Julia Glaser

Abstract. Students often falsely put into perspective the effort required by a learning strategy compared to its benefits and therefore use ineffective ones. The purpose of this evaluation is to examine how students view the strategy of testing when they are required to do so in class. To this aim, a survey was conducted on several study characteristics at the end of two testing interventions ($N = 224$). Descriptive observation shows that students find the intervention appropriate in principle. However, differences were found with regard to the length, the organization and the transparency of the implementation in favor for the shorter, better organized study. These observations can be used to optimize and adapt interventions to students' needs to make the use of strategies as easy as possible. Whenever possible, interventions should be evaluated by students, to ensure the adequacy and adapt where needed.

The Testing Effect in the Lecture Hall: How do Students judge this Learning Method?

Students often do not use effective learning strategies spontaneously, either because they lack knowledge about it (Bjork et al., 2013), or because they misanticipate effort and benefit (Bannert, 2009; Foerst et al., 2017). Said the other way round: if believing in its effectiveness and having enough resources to implement it, students use a strategy more often and especially more spontaneous. Thus, when learning strategies are used in the classroom, it is useful to discover how useful in terms of various aspects students perceived them to be. This can be used to better tailor classroom interventions to the needs of students, especially in terms of the available resources, so that it is more likely that they will take the strategy seriously and use it over the long term.

Rationale of the Evaluation

The purpose of this evaluation is to show how students perceive practice tests when they are provided to them as a learning strategy along with all the necessary resources. This matter will be considered by evaluating the feedback questions that students completed at the end of Study 2 and Study 3 and drawing conclusions about the appropriateness of practice testing in the classroom. We hypothesize that if the barrier to use the strategy is low because all resources are already available, it will be perceived as useful for the own learning. We also included other aspects of the feedback that we would like to look at in an exploratory way.

Method

Participants

Participants were the 224 students participating in Study 2 (long-term intervention) and Study 3 (short-term intervention) above. Taken together, they have an average age of 20.65 years ($SD = 3.18$), ranging from 18 to 42 years. Most participants were female (83.75%) and only 7 reported to have another mother tongue than German.

Measures

At the end of the criterial test in Study 2 and Study 3, participants had to answer a feedback questionnaire concerning the following aspects of the study: length of study, information density of study, usability of study, comprehensibility of study, organization of study, transparency of study goal. On a 5-point scale, they could rate whether the aspects were appropriate (1) or inappropriate (5). Additionally, they had the possibility to give an individual, additional feedback on further aspects not named. The questionnaire is available in the Appendix.

Results

We first looked descriptively at how satisfied students were with aspects of the study in general. We then calculated a mean comparison to relate participants from Study 2, which was longer-term, to participants from Study 3, which was shorter-term. The data file and analysis script can be found in the Appendix.

Descriptive view of the feedback

Table IV.I shows the descriptives for the entire sample. Most aspects were rated as appropriate or little lower in average.

Table IV.1

Descriptives for the entire sample.

	min	*max*	*M*	*SD*
Feedback about length of study	1	4	1.43	0.71
Feedback about information density of study	1	5	1.72	0.82
Feedback about usability of study	1	5	1.93	0.95
Feedback about comprehensibility of study	1	5	1.30	0.61
Feedback about organization of study	1	4	1.60	0.90
Feedback about transparency of study goal	1	5	2.03	1.01

Note. Students rated on a 5-point-scale with 1 meaning appropriate and 5 meaning inappropriate.

Test of mean differences between Study Groups 2 and 3

As viewable in Table IV.2, there were differences between Study 2 and Study 3 concerning some aspects. We therefore tested the statistical differences between them. Significant mean differences were found for the length of the study, $t(106) = 2.30$, $p = .023$, $\eta_p{}^2 = .109$, the organization of the study, $t(91) = 4.55$, $p = .007$, $\eta_p{}^2 = .152$, and the transparency of the study goal, $t(102) = 2.24$, $p = .014$, $\eta_p{}^2 = .200$. On all these aspects, Study 3 was judged to be more appropriate. Statistical values for all aspects of the feedback are shown in Table IV.3.

Table IV.2

Descriptives split by Study Group.

Group		*min*	*max*	*M*	*SD*
Study 2	Feedback about length of study	1	4	1.61	0.80
	Feedback about information density of study	1	4	1.69	0.80
	Feedback about usability of study	1	4	2.04	0.88
	Feedback about comprehensibility of study	1	3	1.30	0.52
	Feedback about organization of study	1	4	2.06	1.10
	Feedback about transparency of study goal	1	5	2.28	1.18
Study 3	Feedback about length of study	1	4	1.36	0.65
	Feedback about information density of study	1	5	1.73	0.83
	Feedback about usability of study	1	5	1.88	0.97
	Feedback about comprehensibility of study	1	5	1.31	0.64
	Feedback about organization of study	1	4	1.40	0.71
	Feedback about transparency of study goal	1	5	1.92	0.91

Note. Students rated on a 5-point-scale with 1 meaning appropriate and 5 meaning inappropriate. Study 2 was the long-term intervention; Study 3 was the short-term intervention.

Table IV.3

Statistical Values for the mean differences.

	T	df	p	η_p^2
Feedback about length of study	2.30	106	.023	.109
Feedback about information density of study	-.39	219	.696	.145
Feedback about usability of study	1.22	219	.225	.183
Feedback about comprehensibility of study	-.08	219	.940	.084
Feedback about organization of study	4.55	91	.007	.152
Feedback about transparency of study goal	2.24	102	.014	.200

Summary of the open answers

Some participants provided additional feedback that supplemented the evaluation questions. For example, some participants wished that the question format in the practice and criterial test (respectively the real exam) should match in order to provide the most benefit. Furthermore, feedback on the correctness of the answer would also be desirable in the critical test and not only in the practice test (which was not given in study 2) and a sufficient processing time for open questions (which was too short in study 2).

Discussion

The present evaluation aimed to find out how students perceive the method of practice testing in the classroom. Feedback on different aspects of the intervention carried out in the context of study 2 (long-term intervention) and 3 (short-term intervention) was evaluated and considered descriptively as well as in mean comparison. We found a consistently positive evaluation of the intervention. Furthermore, there were significant differences between length, organization, and transparency between the two studies in favor for the short-term intervention. This suggests that testing is well accepted as a strategy, but students prefer to participate in shorter interventions, and it requires good organization and transparency for high commitment.

These results contribute to the research on real-life educational practices by providing information on the preferences of students. Even if there is much research on the testing effect (Yang et al., 2021, included 225 effects in their meta-analysis), only few researchers have reported student feedback. Yet this is precisely what is important to tailor interventions to students' needs and preferences in order to increase their willingness to use strategies.

The fact that the feedback criteria were generally rated as adequate or just below suggests that students are comfortable using the strategy of testing for learning in the way implemented in our intervention (questions provided by instructor, feedback on answer, opportunity to repeat if errors are made).

The differences between the groups result from experimental differences, since the two studies were not performed identically. Study 2, compared to study 3, was long-term (7 weeks vs. 1 week). Students in study 3 rated the duration as more appropriate, so it can be assumed that shorter interventions in general are preferred. In addition, there were start-up difficulties in study 2 due to technical problems, so that the organization of study 2 was rated slightly lower. In general, however, the satisfaction is in the upper range, so that the procedure of our intervention (link to the exercise questions by mail on the day of the lecture) seems appropriate.

Additionally, to the feedback for the instructors, the evaluation could have positive effects for the students filling it in. Regardless of the voluntariness of using the learning strategy, evaluation at the end of a learning process has the benefit of allowing students to reflect on their learning and thus strengthen their metacognitive knowledge and awareness. For example, Biwer and colleagues were able to show that training in learning strategies strengthens metacognitive awareness and makes the spontaneous use of effective learning strategies more frequent (Biwer et al., 2020). So, letting students evaluate an intervention in the end has a double benefit and should be used where possible.

Limitations and Further Directions

Beside the information the evaluation provides, there are some limitations in interpreting it. The two studies evaluated differ in some characteristics (e.g., length, organization, content), therefore summarizing them may lead to inaccuracies. In the future, studies of identical structure should be compared in order to be able to make reliable statements. In addition, the feedback was rather general, so the informational value is limited. Finally, the evaluation applies only to this particular intervention setting. In other settings, evaluation is needed to say something about the strategy of testing and its perception.

Conclusion

The present evaluation shows that the intervention design as used in the two studies is considered appropriate by the students and thus can be used well in the classroom. In the future, evaluations should be used more frequently in order to be able to continuously improve interventions on the one hand and to offer students the opportunity to expand their metacognition on the other hand.

References

Bannert, Maria (2009): Promoting Self-Regulated Learning Through Prompts. *Zeitschrift für Pädagogische Psychologie 23*(2), 139-145. https://doi.org/10.1024/1010-0652.23.2.139

Biwer, F., Egbrink, M. G. A. o., Aalten, P., & de Bruin, A. B. H. (2020). Fostering effective learning strategies in higher education—A mixed-methods study.*Journal of Applied Research in Memory and Cognition, 9*(2), 186–203. https://doi.org/10.1016/j.jarmac.2020.03.004

Bjork, R. A., Dunlosky, J., & Kornell, N. (2013). Self-regulated learning: Beliefs, techniques, and illusions. *Annual Review of Psychology*, *64*, 417-444. https://doi.org/annurev-psych-113011-143823

Foerst, N. M., Klug, J., Jöstl, G., Spiel, C., & Schober, B. (2017). Knowledge vs. Action: Discrepancies in University Students' Knowledge about and Self-Reported Use of Self-Regulated Learning Strategies. *Frontiers in Psychology, 8,* 1288. https://doi.org/10.3389/fpsyg.2017.01288

Yang, C., Luo, L., Vadillo, M. A., Yu, R., & Shanks, D. R. (2021). Testing (quizzing) boosts classroom learning: A systematic and meta-analytic review. *Psychological Bulletin*, *147*(4), 399–435. https://doi.org/10.1037/bul0000309

Chapter VI

General Discussion

The mail goal of this dissertation was to investigate how practice testing as a learning strategy of self-directed learning can be adapted to individual characteristics or situational circumstances to generate best possible learning progress in a real-educational online setting. In Chapter 1, I summarized and discussed relevant research on the testing effect, especially with regard to the research questions. It became apparent that the testing effect itself is well researched, but that some potential influencing factors are largely unnoticed or have ambiguous findings. I also presented the main theoretical mechanisms of the testing effect and showed how these predict the influence of metacognitive activation or the effectiveness of testing in transfer tasks. Furthermore, I outlined the moderators of the testing effect that have already been researched and argued which other moderators could be envisaged on the basis of theoretical considerations.

In the following I will first summarize and discuss the main findings of the three empirical studies and then discuss the theoretical and practical implications that can be drawn from all the studies.

Summary and Discussion of Study 1

The first empirical study presented in Chapter II examined the testing effect and potential moderating effects of various individual factors like cognitive ability, motivational, and emotional dispositions in a real-educational setting. We provided students in regular lecture sessions with online practice tests shortly after the lessons, consisting of short answer questions with corrective feedback and review using summary statements, and compared how recall differs in a one-week delayed final test.

The main finding was a medium to strong positive testing effect ($\eta^2 = .07$) across all topics, meaning that tested knowledge was better retained then restudied knowledge. This is in line with the huge body of research on the direct testing effect (e.g. Bjork, 1975; Carpenter,

2009; Pyc & Rawson, 2009). The generalizability of the effect is ensured by the within-study design and the consistently positive result across all topics.

As a further result, we did not find any evidence for a moderating effect of any learner characteristic on the effectiveness of the testing effect. As this broad range of learner characteristics had not previously been considered as moderators of the testing effect, this finding extends the available literature and reinforces the assumption that the testing effect is "a learning method for all" (Jonsson et al., 2021). To sum up, the testing effect seems to be a great didactic tool to foster long-term retention which can be implemented in a highly efficient way regardless of individual differences.

Summary and Discussion of Study 2

The second empirical study presented in Chapter III replicated and extended the first study, investigating whether practice tests would not only benefit retrieval of directly tested contents, but also of non-tested, but thematically linked contents of the same learning nugget. Similar to Study 1, we offered students online practice sessions, but this time with a 2-day delay and the chance to repeat falsely answered question once, and after a six-week experimental phase with 6 different lectures, we conducted a final test across all topics to uncover the testing effect.

The main finding was again a strong testing effect ($\eta^2 = .08$), only occurring for content included in the practice test but not for content only presented in the lecture. This shows, clearer than other studies before, that the testing effect is limited on tested content and not transferable to untested content and teachers should take this into account when handing out learning materials.

Also, restudying the content after a two-day delay had a smaller, but positive effect on long-term retention compared to just studying content during the lesson once. So restudying is not valueless, but testing should, whenever possible, be preferred.

The results of this study confirm the strong testing effect found in Study 1. The learner characteristics surveyed, which also showed no effects here, also confirm the previous results.

But even if the results seem to be clear, the possibility of transfer effects of testing to content learned but not practiced is not necessarily ruled out. This study rather shows up that under certain boundary conditions, the testing effect cannot help transfer – details have to be examined in future research.

Summary and Discussion of Study 3

The third empirical study presented in Chapter IV, another replication and extension of Study 1, examined the testing effect and whether metacognitive activation in form of prompts influences its effectiveness. Similar to Study 1, we offered students online practice sessions, but this time with half of the participants receiving an instruction serving metacognitive activation and for all with the chance to repeat falsely answered question up to five times, and after one week, we conducted the final test.

The main result was a very strong testing effect ($\eta^2 = .136$), occurring regardless of the metacognitive activation of the learners with benefit to the tested items. There was also no difference in the final test performance of people knowing the testing effect before or not. It seems that, contrary to the claims of other studies (e.g. Biwer et al., 2020; Schwonke et al., 2013), metacognitive awareness has no influence on the effectiveness of learning with self-testing, neither for people with good metacognitive knowledge nor for people without. However, it is also possible that the opportunity to repeat falsely answered questions up to

five times has evoked a ceiling effect and therefore the metacognitive activation has not brought any further benefit. Further examination of this is necessary.

The results of this study confirm the strong testing effect found in studies 2 and 3.

Theoretical Implications

The first aim of this dissertation was to explore the effectiveness of testing in digital, self-directed learning environments. Positive effects of testing in the classroom context were demonstrated in all empirical studies. The experimental design of all studies with a within-manipulation of the learning condition and the use of different question formats and learning topics extends the previous literature to the fact that previously mostly quasi-experimental and between-designs were used (e.g. see meta-analysis by Yang et al., 2021). The strong testing effect despite only minimal intervention and use in the field (with many potential sources of variance) shows that testing in the form used here is highly effective.

The results of the present studies support theoretical approaches that hold deeper processing of tested vs. non-tested knowledge responsible for the effect, for example Carpenter's elaborative retrieval hypothesis (Carpenter, 2009). In Study 1, for example, this is evident in the longer processing time of test items compared to restudy items.

The second aim of this dissertation was to investigate the influence of individual learner characteristics on the effectiveness of testing. Study 1 in particular showed that individual characteristics, whether cognitive, emotional, or motivational, do not appear to have any influence on the effectiveness of the testing. The other two studies, in which sporadic individual characteristics were collected, confirmed this statement.

The exploration of individual characteristics to this extent is a new contribution to the literature on the testing effect, as they have only occasionally been addressed as potential moderators. That the selected characteristics have main effects on learning is already known

from previous research and could also be partially confirmed here. The fact that no interactions were found is new, but not completely surprising, as Jonsson has already shown. He claimed that the practice testing is a "learning method for all" (Jonsson et al., 2021), as he could not find any influence of cognitive abilities on the effectiveness of testing. Our study confirms this assumption and extends it by adding emotional, motivational, and non-cognitive individual factors.

The testing effect therefore does not appear to be linked to any prerequisites. Regardless of their characteristics, deeper processing in the sense of the desirable difficulties (Bjork & Bjork, 2011) is triggered in the same way in all people and they can benefit.

Concerning the second aim of this dissertation, it can be concluded that the testing effect is so strong that individual characteristics cannot have any influence on its effectiveness, neither positive nor negative.

The third aim of this dissertation was to examine the influence of situational characteristics or demands on the effectiveness of testing, specifically whether testing is also effective in the context of transfer tasks and whether metacognitive activation changes the effectiveness of testing.

Generally, the studies could show that testing is more effective than restudy, especially multiple testing in the case of wrong answers (see Study 2 and 3). This fits in with the theoretical approaches considering deeper processing to be responsible for the testing effect (e.g., Bjork & Bjork, 2011; Carpenter, 2009). Also, restudy with a delay of two days after the lesson is better than studying only in the lesson (see Study 2). This is an important finding, as many studies on the testing effect in the past have not provided any further learning opportunity in the control group (for a meta-analytic overview see Pan & Rickard, 2018). The

fact that the testing effect is so strong despite an effective control intervention shows how robust it is.

Somewhat contrary to the idea of transfer-oriented processing (Morris et al., 1977), it did not matter whether the question format was the same in the practice and criterial tests. However, the context of asking questions may already be similar enough that the question format may not make much difference.

In Study 2, which examined the effectiveness of testing in the context of transfer, no benefit of testing could be found. This shows that the testing effect also has its boundaries. But: as described in the mediator effectiveness hypothesis (Pyc & Ramwson, 2010), testing can lead to a connection between two contents. However, the prerequisite is that the semantic link is inherent. In the study conducted, this cannot always be assumed, which is why the desired effect cannot be shown.

In Study 3, examining the effect of metacognitive activation on testing, no benefit could be found. This could mean that the testing effect works per se, regardless of whether one is metacognitively active or not, comparable to the independence of individual characteristics (see Study 1). However, it could also be that the effect did not have a chance to show up due to the study design. This needs to be examined in further studies.

Concerning the third aim of this dissertation, it can be concluded that although the testing effect is robust against the influences of many situational circumstances, it also has limitations, especially in the area of knowledge transfer.

The fourth aim of this dissertation was to assess the feasibility of testing in classroom practice. The evaluation of the student feedback showed that the learning method is generally supported. However, in line with the findings of Bannert (2009) and Foerst and colleagues

(2017), this is only true if the effort involved is low and the goal of the intervention is clearly recognizable.

Practical Implications

The results of this dissertation contribute not only to the body of research, but also to educational practice with many new insights that can benefit both teachers and learners. A much confirmed but nevertheless noteworthy result is that the testing effect shows robust effects in the real educational context, even if, like in all presented studies, the intervention is minimalistic. Short units consisting of repetition questions with feedback are already sufficient to consolidate the knowledge in the long term. Especially in the digital context, the effort is very low and sustainability is high, as the units can be easily reused or adapted. Teachers should therefore take advantage of this opportunity and not only use tests for the exam or as an exercise shortly before it, but also incorporate them into lessons as a routinely learning opportunity.

Further, as we used different question types (multiple-choice or short-answer questions) across the three studies, we were able to show that the testing effect occurs equally in both cases. Teachers therefore have a free choice of what question type they want to use and can so adapt it to the time resources they or the students can spare or to the format of the final exam to fit.

In addition, by using a short-term (see Study 1 & 3) and a long-term design (see Study 2), we were able to show that the retention effects are visible both after one and after up to 6 weeks. It is therefore possible to use testing not only as a rapid, effective learning method, but also for longer-term exam preparation. It is thus recommended that teachers routinely incorporate practice tests into their lessons or their revision. As mentioned in Chapter V, learners prefer to use the learning strategy in the short term, so it should be used at least for

exam preparation. However, routine use can also shift this preference in favor of long-term use of effective learning strategies.

In addition to the testing effect itself, we were also able to show with the 3 studies that the moderators we assumed had no influence. On the one hand, we were unable to identify differences in the effectiveness of the testing effect for either cognitive, motivational or emotional characteristics of the learners. So, it seems that the testing effect is the perfect learning method for any learner, as it works highly effective regardless of their individual characteristics. Teachers can therefore use it without any further adaptations to provide learning benefits for all students. On the other hand, also metacognitive activation did not cause any differences in the effectiveness of the testing effect. Apparently, it is not important for the effective use of the practical test to be familiar with the benefits of the effect. Teachers can therefore also offer the learning strategy without detailed explanations and learners benefit from this. However, the latter is an important point in practice: if the strategy is offered to learners, it makes no difference whether it is explained to them or not. However, as other research teams have shown (e.g., Biwer et al., 2020; Schwonke et al., 2013), if students decide on their own learning strategies, it could be advantageous to have metacognitive background knowledge. It is therefore advisable, where possible in terms of resources, to also explain the benefits of the learning strategies used or to incorporate "learning to learn" into the lessons.

An important limitation of the testing effect that teachers must be aware of is that it does not work for tests with transfer tasks, at least in the sense of semantic transfer. However, since such tasks are often used in an educational context, it is necessary to combine practical testing with other learning strategies that better support transfer.

Finally, student feedback from the studies should also be taken into account for practical implications. It seems to be important that when providing questions, a corrective

and also informative feedback after every learning occasion should be given. The learning process also needs to be well organized, especially if it is decentralized and takes place asynchronously online, so that students do not have to deal with additional work due to malfunctioning technology or unclear instructions. Otherwise, the learning strategy will be used reluctantly, and the intervention will lose its usefulness.

To sum up: practice tests are strongly recommended for the use in the classroom as they are universal and easy to implement. But they should be considered as an additional action to teaching and other learning strategies, as they cannot replace it, especially in terms of understanding and consolidating knowledge.

Limitations and Directions for Future Research

In the following chapter, I discuss the restrictions on the generalizability of the studies in a synopsis.

In the general discussion, the three studies are compared with each other, and common conclusions are drawn. However, it must be noted here that the studies had a similar, but not the same structure. The study designs were slightly different each time and different participants took part, so any interpretation should be treated with caution. Future studies in which common conclusions are to be drawn should be conducted with an identical design in order not to jeopardize the generalizability of the results across the study series.

As with many psychological studies, the participants in this study always came from the same group of people (student teachers at the beginning of their studies) and were carried out in a psychological course. Although the variety of topics within the studies showed that the results can be generalized to a certain extent, it is not guaranteed that they can be transferred to topics from a completely different domain. Future research should therefore concentrate on including new topics and making the target group more diverse.

Despite many efforts to keep the test conditions the same for all participants, it is not guaranteed that this will work out in field experiments, as in these three studies. For example, it is not possible to check whether and how much the students have learned in addition to completing the practice tests, which could influence the results in the critical test. Although this was queried and either examined by comparing whether people who have studied additionally differ, or those affected were directly excluded, the query is based on trust and cannot be checked for accuracy. In addition, the students came from over 10 different seminar groups with different lecturers, all of whom have a different teaching style. Some lecturers also use practice questions as part of their lectures. Although it was also asked in advance how the sessions were designed in order to avoid confounding with the manipulation and to exclude affected seminar groups, it cannot be ruled out that individual practice opportunities were integrated into the sessions or that students received and used additional learning material from other courses. And learning tips given by some lecturers could also have influenced the learning behavior of students in the various groups differently. Future research should therefore consider paying more attention to the comparability of the courses to increase generalizability.

Finally, it should be borne in mind that although the studies in this dissertation were able to demonstrate few effects, there could also be effects of testing on transfer or influence by individual characteristics or metacognitive activation with a different group of participants, with different materials or topics, or under other experimental conditions. Future research should therefore consider more diverse groups of participants, different types of transfer and other ways of metacognitive instruction to see if there are specific conditions under which effects occur.

Conclusion

The aim of this dissertation was to investigate how practice testing as a learning strategy of self-directed learning can be adapted to individual characteristics or situational circumstances to generate best possible learning progress in a real-educational online setting. Therefore, various possible influencing factors on practical testing were examined in existing university courses. An overview of existing research has shown that the topic is relevant, but that there are still research gaps or inconsistencies in the results. The findings of this work should therefore help to reduce the gaps.

The results clearly show that the testing effect is a robust phenomenon that can be easily integrated digitally into common classroom practice. The effect is strong and benefits students for long-term retention, both over a short (one week) and longer period (up to 6 weeks).

Furthermore, the results show that the testing effect is resistant to potential influencing factors such as cognitive, emotional or motivational factors. Regardless of these individual characteristics, all learners can benefit equally from the advantages of testing, which is why the learning strategy can and should be widely used in the classroom context. In addition to the individual characteristics that the learners themselves bring along, the metacognitive activation that the teacher can trigger in the learners also appears to be of no significance for the effectiveness of the test. No differences were found as to whether a person was aware of the advantages of using the strategy or not. Testing can therefore be integrated into teaching practice without detailed instructions.

In fact, the results of the available studies also clearly show the limitations of testing. For example, the testing effect was only found for content that was tested identically in the practical test. Transfer questions in the sense of semantically linked information could not be answered better by testing related information beforehand. This does not diminish the broad

applicability of practice tests but shows that they should be supplemented by other learning methods, depending on the context.

To summarize, practice testing is a very effective, universal tool that should be integrated into the classroom wherever possible, ideally embedded in a system of different learning strategies that all promote the sustainable retention of information.

References

Bannert, Maria (2009): Promoting Self-Regulated Learning Through Prompts. *Zeitschrift für Pädagogische Psychologie 23*(2), 139-145. https://doi.org/10.1024/1010-0652.23.2.139

Biwer, F., Egbrink, M. G. A. o., Aalten, P., & de Bruin, A. B. H. (2020). Fostering effective learning strategies in higher education—A mixed-methods study.*Journal of Applied Research in Memory and Cognition, 9*(2), 186–203. https://doi.org/10.1016/j.jarmac.2020.03.004

Bjork, E. L., & Bjork, R. A. (2011). Making things hard on yourself, but in a good way: Creating desirable difficulties to enhance learning. In M. A. Gernsbacher, R. W. Pew, L. M. Hough, J. R. Pomerantz (Eds.) & FABBS Foundation, *Psychology and the real world: Essays illustrating fundamental contributions to society* (pp. 56–64). Worth Publishers.

Carpenter, S. K. (2009). Cue strength as a moderator of the testing effect: The benefits of elaborative retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35*(6), 1563–1569. https://doi.org/10.1037/a0017021

Foerst, N. M., Klug, J., Jöstl, G., Spiel, C., & Schober, B. (2017). Knowledge vs. Action: Discrepancies in University Students' Knowledge about and Self-Reported Use of Self-Regulated Learning Strategies. *Frontiers in Psychology, 8,* 1288. https://doi.org/10.3389/fpsyg.2017.01288

Jonsson, B., Wiklund-Hörnqvist, C., Stenlund, T., Andersson, M., & Nyberg, L. (2021). A learning method for all: The testing effect is independent of cognitive ability. *Journal of Educational Psychology*, *113*(5), 972–985. https://doi.org/10.1037/edu0000627

Pan, S. C., & Rickard, T. C. (2018). Transfer of test-enhanced learning: Meta-analytic review and synthesis. *Psychological Bulletin, 144*(7), 710–756. https://doi.org/10.1037/bul0000151

Pyc, M. A., & Rawson, K. A. (2010). Why testing improves memory: Mediator effectiveness hypothesis. *Science*, *330*(6002), 335. https://doi.org/10.1126/science.1191465

Schwonke, R., Ertelt, A., Otieno, C., Renkl, A., Aleven, V., & Salden, R. J. C. M. (2013). Metacognitive support promotes an effective use of instructional resources in intelligent tutoring. *Learning and Instruction, 23*, 136–150. https://doi.org/10.1016/j.learninstruc.2012.08.003

Acknowledgements

First of all, I would like to thank my supervisor, Prof. Dr. Tobias Richter, for the support on the way to the dissertation, the constructive feedback on my work and the opportunity to learn a lot during this time.

I would also like to thank my colleagues in the CoTeach project who, thanks to their diversity of views from different disciplines, triggered entirely new thoughts that have been incorporated here. In this place, special thanks go to my colleague from my workpackage, Dr. Moiken Jessen, for her mental and creative support, no matter whether a new study design needed to be discussed or we were digesting the daily business of research over a coffee.

Further thanks go to my colleagues from the department, who helped me with tips, constructive criticism and in useful discussions to many useful ideas.

Special thanks go to my auxiliary researcher Ann-Sophie Schriewer, who actively supported me in the project for two years.

A big thank you goes to my parents who have always supported me all the way till here, stood behind me and always believe in me.

Finally, I am grateful for all the people who - knowingly or unknowingly - accompanied me on my way, supported me, inspired me, gave me tips or were just there for me.

Appendix

Note: The enclosed DVD contains parts of the materials mentioned. In the appendix it is listed where the materials can be found.

Appendix A: Chapter II

Datafile & Syntax

The data file and syntax can be accessed both as an SPSS version and in Excel format on the enclosed DVD. In addition to the data files, there is also a list of the occurring variables in the analysis and a file with an explanation of these.

Data source: *Appendix A – Chapter II > Datafile & Syntax*

Materials

The questions of the practice and the criterial test as well as the ones for the prior knowledge test can be found in German below and on the DVD. The English translations can also be found on the DVD.

The DVD also contains a file with the open answers of the participants in raw and coded form for a better understanding of the evaluation.

Data source: *Appendix A – Chapter II > Materials*

Fragen für Übungs- und kriterialen Test
(Deutsche Version/German version – English version see DVD)

Anmerkung: Fragen sind als MC sowie als offene Fragen formuliert; entweder können die Fragen direkt so gestellt werden, oder eine „Alternativfrage" für das offene Format steht unten drunter.

Die markierten Lösungen sind zugleich auch die Musterlösung für die Auswertung offener Fragen. In manchen Fällen stehen ergänzende Antwortmöglichkeiten bei den Fragen dabei, die im Falle von offenen Fragen auch in die Bewertung einfließen können.

Im Übungstest ist das Format bei allen Fragen offen. Im kriterialen Test variiert es zwischen offen und Multiple-Choice.

Fragen zu Lese-Rechtschreib-Schwäche

Fragen Übungstest	1, 2, 3, 4, 5, 6, 8, 11, 13, 14, 15, 16, 17, 18, 24
Fragen kriterialer Test	1, 4, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 18, 19, 20, 21, 22, 23, 25
Art kriterialer Test	Offene Fragen

*die hier aufgeführte Reihenfolge der Fragen ergibt die Nummerierung im Datenblatt.

1. Welche Vorläuferfähigkeiten benötigen Kinder zum normalen Spracherwerb?
    a. **Bewusstheit um Wörter und Silben**
    b. **Bewusstheit um Sprachlaute**
    c. Bewusstheit um Wortklang
    d. Bewusstheit um Satzbau

2. Wodurch zeichnen sich Kinder mit chronischen Schriftsprachschwierigkeiten aus?
    a. **Langsame Entwicklung**
    b. Unsichere Aussprache
    c. Physische Änderungen im Gehirn
    d. **Höhere Fehlerrate**
3. Der Verlauf einer LRS…
    a. Ist stabil und nicht beeinflussbar.
    b. **Ist stabil, eine Therapie kann aber positiven Einfluss haben.**
    c. Ist instabil, je nach sonstiger Entwicklung des Kindes.
    d. Ist instabil und die Problematik verschwindet in manchen Fällen ohne Therapie.
    Alternativfrage offen: Welche Aussage lässt sich über die Stabilität der Entwicklung einer LRS tätigen?
4. Worauf zielen Präventionsmaßnahmen ab?
    a. **Phonologische Fähigkeiten**
    b. Semantische Fähigkeiten
    c. **Buchstabenwissen**
    d. Wortwissen
5. Interventionen bei einer LRS…
    a. Setzten am aktuellen Leistungsstand des Kindes an.
    b. **Bauen den Schriftspracherwerb systematisch neu auf.**
    c. Gehen solange im Schulstoff rückwärts, bis das Problem aufgedeckt wird.
    d. Begleiten den Unterricht und versuchen, den Stoff verständlich zu machen.
    Alternativfrage offen: Was wird mit Interventionen bei LRS versucht?
6. Ein Nachteilsausgleich… (MC)
    a. Beinhaltet auch einen Notenschutz.
    b. **Gilt nicht nur bei LRS, sondern auch bei anderen Beeinträchtigungen.**
    c. **Wird nur nach schulpsychologischer Stellungnahme gewährt.**
    d. Umfasst ausschließlich eine Zeitverlängerung in Prüfungssituationen.
    Alternativfrage offen: Gibt es einen Unterschied zwischen Nachteilsausgleich und Notenschutz? Erläutern sie in max. 2 Sätzen.
7. Typische Verhaltensweisen in der logographemischen Phase sind?
    a. **Malen von Wörtern**
    b. Lauttreue Verschriftungen
    c. Übergeneralisierungen
    d. Erfindung eigener Regeln
8. In welcher Reihenfolge durchlaufen Kinder die unterschiedlichen Phasen des Schriftspracherwerbs?
    a. **Logographemische Phase -> Alphabetische Phase -> Orthografische Phase**
    b. Alphabetische Phase -> logographemische Phase -> Orthografische Phase
    c. Orthografische Phase -> alphabetische Phase -> logographemische Phase
    d. Alphabetische Phase -> orthografische Phase -> logographemische Phase
    Alternativfrage offen: In welcher Reihenfolge laufen die Phasen des Schriftspracherwerbs ab? Nennen und ordnen Sie die Phasen?

9. Was unterscheidet die direkte von der indirekten Route im Dual Route Cascaded Model?
   a. **Direkte Route lässt sich auch als lexikalische Route beschreiben.**
   b. Indirekte Route lässt sich auch als lexikalische Route bezeichnen.
   c. Bei der indirekten Route werden die Worte direkt am Erscheinungsbild erkannt und mit einem im Kopf vorhandenen Lexikon abgeglichen.
   d. Bei der direkten Route werden Worte nicht als Ganzes, sondern nur in Einzelteilen ermittelt und dann zusammengefügt.

10. Welche kognitiven Fähigkeiten haben einen Einfluss auf das Leseverstehen? (Nennen Sie mind. 3)
    a. **Hintergrundwissen**
    b. **Schlussfolgerndes Denken**
    c. Aufmerksamkeit
    d. Wahrnehmung
       Weitere Fähigkeiten: Strategienutzung, Leseflüssigkeit, Wortschatz

11. Welche lese- und rechtschreibbezogenen Probleme müssen vorliegen, um von einer Lese-Rechtschreib-Störung zu sprechen (nach ICD11)? (Nennen Sie 3)
    a. **Probleme in der Lesegenauigkeit**
    b. Probleme in der Aussprache von Worten
    c. **Probleme in der Genauigkeit der Orthografie**
    d. Probleme beim Schreiben einzelner Buchstaben
       Weitere Probleme: Lesegeschwindigkeit, Leseverständnis, Genauigkeit der Syntax, fehlerhafte Zeichensetzung, mangelnde Kohärenz und Organisation der Ideen beim Verschriftlichen

12. Was sind Ausschlusskriterien in der Diagnostik einer LRS? (nennen Sie 3)
    a. Motorische Probleme
    b. **Sensorische Behinderung**
    c. Normale Beschulung
    d. **Unterdurchschnittliche Intelligenz (IQ < 70)**
       Weitere Kriterien: mangelnde Beschulung, mangelnde Sprachbeherrschung, psychosozial widrige Umstände

13. Welche Arten von Rechtschreibfehler gibt es und was beschreiben diese? (Nennen & beschreiben Sie 2)
    a. **Phonemfehler: Verstöße gegen lautgetreue Schreibung**
    b. **Speicherfehler: Verstöße gegen regelhafte Abweichungen, z. B. Doppelvokale**
    c. Graphemfehler: Phonem-Graphem-Zuordnungsprobleme
    d. **Orthographiefehler: Zeichensetzung, Großschreibung substantivierter Verben & Adjektive, …**
       Weitere Fehler: Regelfehler (Verstöße gegen regelhafte Abweichungen der lautgetreuen Schreibung), Restfehler (Zeichensetzung, Großschreibung substantivierter Verben & Adjektive, …)

14. Welche Ursachen kann eine LRS haben? (Nennen Sie 2)
    a. Intelligenzminderung
    b. **Genetische Ursachen**
    c. Schlechte Beschulung
    d. **Neuronale Ursachen**
       Weitere Ursachen: Sprachentwicklungsverzögerung

15. Welche Aussagen über die Sprachentwicklung sind wahr?
    a. Kinder sollten mit 1 Jahr bereits 50 Worte sprechen können.
    b. **Kinder, die unter die „Late talkers" fallen, haben eine 50% Wahrscheinlichkeit, noch zu einer normalen Sprachentwicklung zu gelangen.**

    c. Kinder, die mit zwei Jahren weniger als 50 Worte sprechen, werden „Late Talkers" genannt.

    d. Eine Legasthenie kann auch ohne verzögerte Sprachentwicklung im Kleinkindalter auftreten.

Alternativfrage offen: Beschreiben Sie kurz den alterstechnischen Verlauf einer Sprachentwicklungsverzögerung. Nutzen Sie dabei Fachbegriffe!

16. Was wird in einer Therapie bei einer LRS behandelt? (Nennen Sie jeweils 2 Beispiele zur Bekämpfung der Primär- & der Sekundärsymptomatik)

    a. Einführung der Rechtschreibregeln

    b. Abbau von leistungsbezogenen Ängsten

    c. Behandlung familienbezogener Problematiken

    d. Üben der Aussprache von Worten

    Weitere Therapiemaßnahmen

Primärsymptomatik

* Verbesserung der phonologischen Bewusstheit
* Systematisches Üben der Wortdurchgliederung
* Graphem-Phonem-Zuordnung
* Einführung der Rechtschreibregeln
* Aufbau eines Merkwortschatzes

Sekundärsymptomatik

* Abbau von leistungsbezogenen Ängsten / Aufbau von Lernmotivation
* Übungen zur Konzentration und Entspannung, die Erarbeitung von Selbsthilfemethoden, Techniken der Fehlerkontrolle und Selbstbestätigung
* Einübung von Bewältigungsstrategien: Verarbeiten von Fehlererfahrung und Versagenserlebnissen
* Behandlung spezifischer psychopathologischer Symptome wie z. B. Schulangst, Einnässen oder dissoziale Entwicklung.

17. Was muss man erfassen, um Lese-Rechtschreibschwierigkeiten für ein formelles Gutachten zu erfassen?

    a. Art der Rechtschreibfehler

    b. Anzahl der Rechtschreibfehler in normiertem Rechtschreibtest

    c. Lesegeschwindigkeit in einem für jüngere Jahrgangsstufen normierten Lesetest, um das Lesealter zu ermitteln

    d. Häufigkeit des Lesens im Privaten

18. Was gehört zur phonologischen Informationsverarbeitung?

    a. Sich vorgesprochene Unsinnswörter merken zu können

    b. Den ersten Laut eines vorgesprochenen Wortes zu finden

    c. Ein bildlich dargestelltes Objekt möglichst schnell zu benennen

    d. Das Verständnis eines gelesenen Textabschnittes selbstständig überprüfen zu können

19. Wie unterscheiden sich die Schwierigkeiten von Tom (schreibt Hunt statt Hund, Kwelle statt Quelle, länge statt Länge, Bal statt Ball, Schpiegel statt Spiegel) und Tim (schreibt Geraffe statt Giraffem Beg statt Berg)? Beschreiben Sie in max. 3 kurzen Sätzen!

    a. Tom hat größere Probleme im Rechtschreiben als Tim

    b. Tom steht auf der logografischen Stufe, Tim dagegen auf der orthografischen

    c. Tom hat Schwierigkeiten mit der orthografischen Strategie, Tim mit der alphabetischen

    d. Tom ist im Schriftspracherwerb eine Stufe weiter als Tim

20. Warum sollte man die Rechtschreibfehler (zumindest die von Fachbegriffen unabhängigen Rechtschreibfehler) in einer Biologie-Probearbeit nicht bei der Notengebung berücksichtigen?

    a. Die Frage ist schlecht formuliert – sollte man nur bei vorliegendem Attest nicht berücksichtigen.

    b. Weil SchülerInnen nichts für die Rechtschreibfehler können

      c. Weil in den Lehrplänen für Biologie die allgemeine Rechtschreibung nicht als zentral angesehen wird.

      d. Weil es die Validität der Messung beeinträchtigen würde

21. Förderung bei einer diagnostizierten Lese-Rechtschreibstörung – was trifft zu?

      a. Da die Ursachen von Leseschwierigkeiten v.a. im visuellen Bereich liegen, sollte bei einer isolierten Lesestörung in erster Linie die visuelle Wahrnehmung trainiert werden.

      b. Im Gegensatz zu Lese-Rechtschreibschwierigkeiten ist bei einer Störung eine Förderung wenig erfolgsversprechend – daher sollte der Notenschutz im Vordergrund stehen

      c. Es gibt ausgearbeitete Programme, die auch von Lehrkräften zur Förderung genutzt werden können

      d. Förderprogramme unterscheiden sich hinsichtlich ihrer Schwerpunkte (z.B. lautgetreue Schreibung, Rechtschreibregeln, Leseverständnis)

22. Wenn Tino folgende Worte folgendermaßen schreibt - zwerk, rante, fersperte, weg – spricht man von einem…

      a. Phonemfehler

      b. Regelfehler

      c. Speicherfehler

      d. Restfehler

23. Welche neurologischen Auffälligkeiten zeigen sich bei LRS?

      a. Änderung der Aktivierungsmuster in der linken Hemisphäre

      b. Änderung der Aktivierungsmuster in der rechten Hemisphäre

      c. Ineffektivere Reizweiterleitung und -verarbeitung visueller Reize

      d. Ineffektivere Reizweiterleitung und -verarbeitung auditiver Reize

24. Wie hoch ist die Inzidenz einer LRS in der Bevölkerung?

      a. Unter 1%

      b. Ca. 3,5%

      c. Ca. 5%

      d. Ca. 10%

25. Welche Folgen in der psychischen, beruflichen oder sozialen Entwicklung kann LRS bewirken?

      a. Selteneres Auftreten von Verhaltensstörungen

      b. Rate der Schulabbrüche höher

      c. Zusammenhang mit Delinquenz erkennbar

      d. Wenn Intelligenz durchschnittlich schaffen auch LRS-Betroffene das Abitur

Fragen zu ADHS

Fragen Übungstest	1, 2, 3, 4, 5, 8, 10, 13, 15, 16, 17, 21, 22, 23, 24
Fragen kriterialer Test	3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 24, 25
Art kriterialer Test	Multiple-Choice Fragen

*die hier aufgeführte Reihenfolge der Fragen ergibt die Nummerierung im Datenblatt.

1. Das Geschlechterverhältnis der Prävalenz bei ADHS… (MC)

      a. Ist für Jungen ungünstig

      b. Ist für Mädchen ungünstig

      c. Ist für beide Geschlechter gleich

Alternativfrage offen: Wie ist das Geschlechterverhältnis der Prävalenz bei ADHS (ohne Zahlangaben)?

2. Unterscheidet sich das Auftreten der verschiedenen Subtypen von ADHS geschlechtsabhängig? (offen)
   a. Ja: Aufmerksamkeit betrifft beide gleichermaßen, Impulsivität & Hyperaktivität mehr bei Jungen zu finden
3. Welche Subtypen von ADHS gibt es? (offen oder MC)
   a. Vorwiegend hyperaktiv-impulsiv
   b. Vorwiegend unaufmerksam
   c. Vorwiegend hyperaktiv-unaufmerksam
   d. Kombination
4. Welche Aussagen über die Diagnostik von ADHS treffen zu? (MC)
   a. Es sollte nur das Kind beobachtet werden
   b. Lehrkräfte sind die zuverlässigste Informationsquelle zur Diagnostik
   c. Neben Verhaltensauffälligkeiten müssen psychische oder körperliche Symptome betrachtet werden
   d. Die Verwechslungsgefahr mit anderen Störungen ist nicht hoch.
   Alternativfrage offen: Was muss neben ADHS-Tests noch betrachtet werden, um eine valide Diagnostik zu erhalten und Alternativerkrankungen auszuschließen? (max. 2 Sätze)
5. Komorbide Störungen bei ADHS… (MC)
   a. Sind eine Ausnahme
   b. Betreffen oft das Sozialverhalten
   c. Können Affekt- oder Angststörungen sein
   d. Hängen fast nie mit Lernstörungen zusammen
   Alternativfrage offen: Wird bei einem Kind ADHS diagnostiziert, sollten auch weitere psychische Störungen in Betracht gezogen werden. Stimmen Sie dieser Aussage zu? Nennen Sie EIN Beispiel für häufige Komorbiditäten oder bringen Sie EIN Gegenargument.
   BESSER: Was können komorbide Störungen bei ADHS sein? Nennen Sie 2 Bsp!
   Beispiele: Störung mit oppositionellem Trotzverhalten, Ticstörungen, Lernstörungen
6. Wie lässt sich der Verlauf von ADHS grob beschreiben? (offen (2 Sätze) oder MC)
   a. Aktivität nimmt bis zum Erwachsenenalter zu, Aufmerksamkeitsstörung persistiert
   b. Physische Aktivität bleibt bis zum Jugendalter hoch, danach abnehmend
   c. Aggressivität entwickelt sich erst im Jugendalter
   d. Alle Symptome bessern sich im Erwachsenenalter
7. Welche Faktoren gelten als Ursachen für ADHS? (offen oder MC)
   a. Alkohol und Nikotin in der Schwangerschaft
   b. Hoher Sozialstatus der Eltern
   c. Ungünstige Bedingungen in der Schule
   d. Wenig Bewegung im Kleinkindalter
8. Intervention bei ADHS setzt am besten an… (MC)
   a. Kind selbst
   b. Eltern
   c. Betreuungspersonen
   d. Ganz Klasse/Gruppe
   Alternativfrage offen: Welche Personen aus dem Umfeld des Betroffenen sollten in eine Intervention einbezogen werden?
9. Bei einer Intervention zur Steigerung der Aufmerksamkeit sollte vor allem Wert gelegt werden auf…?
   a. Selbstkontrolle
   b. Planvolles Herangehen an Probleme
   c. Bestrafung von Hyperaktivität
   d. Reizüberflutung zur Gewöhnung
   Weitere Antwortoption: Organisation des Lernprozesses

10. Welche Kriterien nach ICD 11 müssen für eine ADHS erfüllt sein? Nennen Sie 3
    a. **Situationsunabhängigkeit des Verhaltens**
    b. IQ durchschnittlich *(aber: IQ nicht lernbehindert wäre richtig)*
    c. **Symptome seit mind. 12 Monaten**
    d. Störung bereits vor 4. Lebensjahr begonnen

11. Was sind die Nachteile einer medikamentösen Therapie bei ADHS? Nennen Sie 2!
    a. **Nebenwirkungen möglich, z.B. Schlafstörungen, Magenbeschwerten, depressive Verstimmungen**
    b. Etwa 40% der Kinder erleidet schwere Nebenwirkungen
    c. **Gefahr des Missbrauchs im Umgang mit Stimulanzien**
    d. **Keine Heilung von ADHS**
       Weitere Nachteile: Vermittelt betreffenden Personen das Gefühl, ohne Stimulanzien nicht korrekt zu funktionieren, Gefahr des Drogenhandels

12. Welche Möglichkeiten bestehen als Lehrkraft, mit Kindern mit ADHS umzugehen? Nennen Sie 3!
    a. **Für ruhige Umgebung mit möglichst wenig Ablenkungspotential sorgen**
    b. Drastische Strafen, wenn Regeln verletzt werden
    c. **Häufiger Pausen einlegen**
    d. Möglichst viele verschiedene Medien nutzen

    Weitere Möglichkeiten:

    - Für eine ruhige Umgebung mit möglichst wenig Ablenkungspotential sorgen (ruhige, reizarme Räumlichkeiten, usw.)
    - Eine inhaltlich geeignete Gliederung und Gestaltung der Lern-und Spielmaterialien schaffen
    - Strukturierung unter Zuhilfenahme von Routinen und Ritualen (z.B. jede Stunde mit ein paar Minuten Entspannung beginnen)
    - Häufiger Pausen einlegen (z.B. 15 min arbeiten, dann 5 min Pause -> Konzentrationsdauer nicht überlasten)
    - Klare Vereinbarungen mit dem Kind treffen und ggf. gemeinsam mit dem Kind (!) Regeln aufstellen
    - Auf strikte Einhaltung der Regeln achten
    - Häufiges Loben der Kinder anstatt sie auf Defizite hinzuweisen; besonders bei hyperaktiven Kindern ist die Verstärkung für positives Verhalten wichtig (in Form von Lob, einem Lächeln, etc.)

13. Welche Komponenten bilden die Aufmerksamkeit laut Attention-Network-Theory?
    a. **Orientierung**
    b. **Wachsamkeit**
    c. Vernetzung
    d. Aktivität
       Weitere: Exekutive

14. Welche Faktoren (aka Ebenen) tragen laut biopsychosozialem Modell zur Entstehung von Aufmerksamkeitsstörungen bei? Nennen Sie 3
    a. **Neuropsychologie**
    b. **Interaktionen**
    c. Interferenzen
    d. **Symptome**
       Weitere: Biochemie * Neurologie, Interaktionen, Komorbide Störungen

15. Wie häufig tritt ADHS in Deutschland auf und was lässt sich im Vergleich mit anderen Ändern sagen? Beschreiben Sie in max. 2 Sätzen
    a. **Die inzidenzrate liegt in Deutschland bei ungefähr 5%, was bedeutet, dass 1-2 Schüler in einer regulären Klasse (30 Kinder) betroffen sind.**
    b. Daten aus den 80er Jahren zeigen, dass es sich bei ADHS um eine Erkrankung handelt, sie erst seit 10 Jahren so hohe Inzidenzraten aufweist.
    c. ADHS ist ein Problem der westlichen Industrienationen. Studien mit weltweiten Daten kommen zu dem Schluss, dass die Inzidenzraten weltweit erheblich von denen aus Deutschland abweichen.
    d. In Deutschland wurde in der Vergangenheit zu häufig ADHS diagnostiziert. Seit 10 Jahren sinkt die Diagnoserate aber und bewegt sich im Bereich der weltweiten Inzidenzraten.
16. Welche verschiedenen Ebenen bieten sich für Interventionen bei ADHS an? Nennen Sie 3
    a. **Gestaltung der Lernumwelt**
    b. Veränderte Bewertungen im Schulkontext
    c. Motorische Trainings
    d. **Erziehungsberatung**
    Weitere: Pharmakotherapie, a. Selbstregulationskompetenzen des Kindes
17. Welche Störungen können mit ADHS leicht verwechselt werden? Nennen Sie 3 Beispiele
    a. **Störungen des Sozialverhaltens**
    b. Beziehungsstörung mit Hemmung
    c. **Persönlichkeitsstörungen**
    d. Hormonstörungen

    Weitere: Störungen des Sozialverhaltens und oppositionelles Trotzverhalten, stereotype Bewegungsstörungen, Autismus-Spektrum-Störungen, Tic-und Tourette-Störungen, umschriebene Entwicklungsstörungen und Lernstörungen, Intelligenzminderung & Überforderung, Autismus-Spektrum-Störungen, Beziehungs-/ Bindungsstörung mit Enthemmung, Angststörungen, depressive Störungen & bipolare Störungen, disruptive Affektregulationsstörungen, Substanzkonsumstörungen, Persönlichkeitsstörungen, psychotische Störungen, Medikamenteninduzierte Störungen (z. B. Bronchospasmolytika), Müdigkeit und Unaufmerksamkeit bei Schlafstörungen, Hyperarousalbei Posttraumatischer Belastungsstörung

18. Welche organischen Probleme können ADHS-Symptome auslösen?
    a. **Sehstörungen**
    b. **Schilddrüsenfunktionsstörungen**
    c. Herzerkrankungen
    d. Lungenfunktionsstörungen
    Weitere: Hörstörungen, Anfallsleiden, Epilepsie, Fragiles-X-Syndrom, 22q11 Mikrodeletionssyndrom, NeurofibromatoseTyp 1, Fetale Alkohol Spektrum Störung (FASD)
19. Was können Folgen einer ADHS-Symptomatik sein? Nennen Sie 3 Beispiele!
    a. Erhöhte Frustrationsintoleranz
    b. **Motivationsschwäche**
    c. Mangelnde Verzögerungsaversion
    d. **Erhöhte Unfallgefahr**
    Weitere: niedrige Frustrationstoleranz, Soziale Konflikte, Bullying, Leistungsschwäche in Schule und Beruf, Substanzmissbrauch, Angst, Psychische Folgeerkrankungen, Teilleistungsstörungen, Affektive Störungen

20. Wessen Urteil ist am zuverlässigsten bei der Erkennung einer ADHS?
    a. Urteil der Eltern
    b. Urteil der Freunde
    c. Urteil des Kindes
    d. Urteil der Lehrkräfte
21. Welche Probleme gibt es bei der Diagnostik einer ADHS? (offen; MC möglich)
    ➔ Grenzen zur Norm sind nicht eindeutig, kontinuierlicher Übergang von altersgemäßem Bewegungsdrang zu ADHS
    ➔ „Vielfalt" der Symptomatik (Verhalten kann von Kind zu Kind oder auch von Situation zu Situation stark variieren)
    ➔ Unterschiedliche Diagnoserichtlinien (ICD; DSM5)
    ➔ Lehrkräfte vermuten bei jüngeren Kinder innerhalb von Grundschulklassen häufiger ADHS, vermutlich da deren Verhalten weniger kontrolliert ist
22. Ist der IQ von ADHS-Betroffenen durch die Erkrankung betroffen? Falls ja, in welchem Ausmaß?
    ➔ IQ-leistungen durchschnittlich 7-15 Punkte vermindert
23. Welche AHDS-typischen Merkmale weisen betroffene Kinder im Säuglings- und Kleinkindalter auf? Nennen Sie 2! (offen; MC möglich)
    ➔ hohes Aktivitätsniveau
    ➔ ungünstige Temperamentsmerkmale und negative Eltern-Kind-Interaktion
    ➔ Schwieriges Ess- und Schlafverhalten
24. Welche AHDS-typischen Merkmale weisen betroffene Kinder im Grundschulalter auf? Nennen Sie 2! (offen; MC möglich)
    ➔ Unruhe / Ablenkbarkeit im Unterricht
    ➔ Lernschwierigkeiten / Teilleistungsschwächen
    ➔ Umschulungen / Klassenwiederholungen
    ➔ aggressives Verhalten
    ➔ Ablehnung durch Gleichaltrige
    ➔ Leistungsunsicherheit / Selbstwertprobleme
25. Welche AHDS-typischen Merkmale weisen Betroffene im Erwachsenenalter auf? Nenne 2! (offen; MC möglich)
    ➔ Persistenz hyperkinetischer Symptome bei 30-60%
    ➔ ausgeprägte Symptomatik bei ca. 30%
    ➔ geringere Schulbildung, berufliche Probleme
    ➔ Delinquenz und dissoziale Persönlichkeitsstörung bei etwa 15-30%
    ➔ Erhöhte Unfallgefahr (50% mehr Fahrradunfälle; 2-4x mehr Autounfälle)
    ➔ Frauen mit ADHS werden deutlich häufiger ungewollt schwanger (44% versus 10%; Owens et al., 2017)

Fragen zur Entwicklung des Denkens

Fragen Übungstest	1, 2, 4, 5, 8, 9, 10, 11, 12, 13, 15, 18, 22, 23, 25
Fragen kriterialer Test	1, 2, 3, 5, 6, 7, 9, 10, 11, 14, 15, 16, 17, 18, 19, 20, 21, 22, 24, 25
Art kriterialer Test	Offene Fragen

*die hier aufgeführte Reihenfolge der Fragen ergibt die Nummerierung im Datenblatt.

Die hier aufgeführten Stichpunkte sind – sofern nichts grün hervorgehoben ist - als komplette Lösung, nicht als MC-Frage zu verstehen!

1. Welche theoretischen Ansätze zur Beschreibung und Erklärung der (kindlichen) Entwicklung des Denkens und Problemlösens gibt es?
    a. Konstruktivistischer Ansatz (Piaget)
    b. Informationsverarbeitungstheorie

c. soziokulturelle Theorien

d. Theorien dynamischer Systeme

2. *Welche Rolle spielt das Kind als aktive(r) Konstrukteur(in) von Wissen? -> Konstruktivistischer Ansatz: Wie lernen Kinder laut des konstruktivistischen Ansatzes?*

➔ Kinder konstruieren in Reaktion auf Erfahrungen und Umwelt Wissen

- o Wissenschaftliches Experimentieren (Hypothesen, Experimente, Folgerungen)
- o Ohne fremde Hilfe
- o Selbst motiviert

3. Beschreiben Sie die kontinuierlichen Prozesse der Theorie von Piaget in jeweils einem Satz!

a. (Assimiliation -> Integration neuer Informationen in bereits bestehende Konzepte)

b. Akkomodation -> vorhandene Wissensstrukturen werden in Reaktion auf neue Erfahrungen angepasst

c. Äquilibration -> Assimilation und Akkommodation werden ausbalanciert, um ein stabiles Verstehen zu ermöglichen

4. Assimilation nach der Theorie der kognitiven Entwicklung von Piaget bezeichnet den Prozess…? Beschreiben Sie in einem Satz!

a. Eintreffende Informationen in eine Form überführen, die man versteht.

b. Vorhandene Wissensstrukturen als Reaktion auf neue Erfahrungen anpassen.

c. Durch Akkomodation und Äquilibration ein stabiles Verstehen schaffen.

5. Was sind die Hauptmerkmale des konkret-operationalem Stadiums nach Piaget? Beschreiben Sie stichpunktartig!

a. logische Denkprozesse in konkreten Situationen

b. Es können mehrere Dimensionen simultan berücksichtigt werden (z.B. Invarianzkonzept)

c. aber kein systematisches Schlussfolgern oder Nachdenken über hypothetische Situationen (z.B. Pendelproblem)

6. Welchem Stadium der Entwicklung nach Piaget würden Sie die folgende Aussage zuordnen und warum? Erklären Sie in 1-2 Sätzen!

Es gibt 3 Kekse. Tom nimmt sich 2. Für Hannes bleibt noch einer übrig, den er in der Mitte durchteilt. Nach dem Verständnis von Hannes und Tom ist das eine faire Aufteilung.

a. Präoperationales Stadium -> Zentrierung -> Fokus nur auf einer Dimension => Anzahl

7. Welche praktischen Implikationen ergeben sich aus Piagets Theorie für die pädagogische Praxis?

a. Rolle des Experimentierens in kindlicher Entwicklung, eigenständige Erfahrungen mit Umwelt gezielt fördern

b. Denkanstöße geben, Widersprüche/Ambivalenzen aufbauen -> Äquilibration

c. kindzentrierter Ansatz: Berücksichtigung der Art des Denkens der Kinder in pädagogischen Kontexten

8. Welche Kritik kann an Piagets Theorie geübt werden? Nennen Sie 2 Punkte

a. Stufen können in unterschiedlichen Bereichen unterschiedlich schnell erreicht werden

b. Fähigkeiten von Säuglingen und Kleinkindern werden überschätzt

c. Unterschätzt soziokulturelle Einflüsse auf kindliche Entwicklung

d. Nicht festgelegt, welche kognitiven Prozesse den Entwicklungsstufen zugrunde liegen

Weitere: kein Kriterium, wann Entwicklung abgeschlossen ist; Fähigkeit von Säuglingen und Kleinkindern werden unterschätzt

9. Inwiefern lässt sich ein Kind laut informationstechnischem Ansatz mit einem Computer vergleichen? Beschreiben Sie in 1-2 Sätzen!
    a. Kognitive Strukturen: Komponenten und ihre Organisation = Hardware (Gedächtniskapazität und Leistungsfähigkeit)
    b. Prozesse: Regeln und Strategien des Gedächtnisses und Problemlösens = Software

10. Wie wandeln sich laut informationstechnischem Ansatz die kognitiven Strukturen und Prozesse im Verlauf der Entwicklung? Beschreiben Sie stichpunktartig!
    a. Kognitive Entwicklung verläuft stetig (kontinuierlich) und nicht in Stufen
    b. Kognitives Wachstum ist gekennzeichnet durch
        i. Erweiterung der Gedächtniskapazität und Leistungsfähigkeit (Informationsumfang, der verarbeitet werden kann)
        ii. Steigerung der Effizienz kognitiver (Denk-)Prozesse
        iii. Erwerb neuen Wissens und neuer Strategien zum Problemlösen

11. Nennen Sie die einzelnen Teile des Gedächtnissystems!
    a. Sensorisches Gedächtnis
    b. Arbeitsgedächtnis
    c. Langzeitgedächtnis

12. Welche Funktion erfüllt das sensorische Gedächtnis? Beschreiben Sie in 1 Satz.
    a. Sensorisches Gedächtnis -> kurze, flüchtige Sinneseindrücke -> konstante Kapazität

13. Was zeichnet das Arbeitsgedächtnis aus (Aufgabe, Kapazität, Entwicklung)? Beschreiben Sie in Stichpunkten!
    a. Arbeitsgedächtnis/KZG -> Infoverarbeitung & -integration -> Kapazität wächst mit Alter

14. Welche Rolle spielen die Exekutiven Funktionen mi Zusammenhang mit dem Arbeitsgedächtnis?
    a. Kontrollsystem zur Steuerung und Verbesserung von AG-Prozessen (z.B. Planung, Aufmerksamkeit, Strategieeinsatz)
    b. Hemmung von Handlungsimpulsen (ab Grundschulalter)
    c. Kognitive Flexibilität (ab ca. 5 Jahre)

15. Was sagt die Theorie der überlappenden Wellen aus? -> 1 Satz
    a. Ansicht, dass das Denken des Kindes variabel ist, nicht durch Alter festgelegt -> Vielzahl von Herangehensweisen wird genutzt, um Probleme zu lösen -> keine starre Entwicklung zu erkennen

16. Welche Rolle spielt das Inhaltswissen bei der Informationsverarbeitung?
    a. Grundlage für Aufnehmen und Speichern neuer Informationen
    b. verbessert Gedächtnis, indem neues Wissen mit bereits vorhandenem integriert wird (z.B. Merken von Schachpositionen)
    c. stellt ok nützliche Assoziationen für Verarbeitung neuer Informationen bereit
    d. lenkt die Aufmerksamkeit & sorgt damit für eine bessere & effizientere Enkodierung

17. Welche praktischen Implikationen ergeben sich aus Informationsverarbeitungstheorien für die pädagogische Praxis?
    a. Übungsgelegenheiten schaffen (verschiedene Strategien aufzeigen, z.B., Zahlen-Brettspiel)
    b. Üben des Anwendens von Strategien in immer unähnlicher werdenden Kontexten -> fördert Transfer
    c. Inhaltswissen vermitteln, das notwendig für Problemlöseprozesse ist oder Problemlösen unterstützt

18. Wie beschreiben soziokulturelle Theorien die Entwicklung des Denkens? Beschreiben Sie stichpunktartig!
    a. Lernen & Entwicklung im zwischenmenschlichen & kulturellen Kontext

b. Kulturwerkzeuge -> z.B. Geschichten, Werte… -> beeinflussen Denken & Problemwesen

c. Entwicklung kontinuierlich & durch quantitative Veränderungen gekennzeichnet

19. Welche Phasen kann man unterscheiden, betrachtet man das Denken als Internalisierung von Sprache? Beschreiben Sie stichpunktartig!
   a. Phase (1): Verhalten des Kindes durch Kommunikation mit anderen Menschen gesteuert
   b. Phase (2): Selbststeuerung des Kindes durch lautes Selbstgespräch (v.a. bei 4- bis 6-Jährigen)
   c. Phase (3): Selbststeuerung durch inneres Sprechen = Denken

20. Was unterscheidet die Theorien dynamischer Systeme von anderen Modellen? /spezifische Annahmen des Modells
   a. Zentrale Rolle von Handlungen für die kognitive Entwicklung
      i. Denken und Handeln in ständigem Wandel begriffen (kontinuierliche Veränderungen)
      ii. Beeinflussen sich gegenseitig (z.B. nach Objekten greifen ermöglicht gezieltes Untersuchen und Manipulieren)

21. Durch was ist das voroperatorische Denken nach Piaget gekennzeichnet? Nennen Sie 2 Punkte
   a. A-/nicht-B-Fehler
   b. Egozentrismus
   c. Zentrierung des Denkens auf eine Dimension
   d. Fehlende Objektpermanenz

22. Was wird mit dem Begriff „Gelenkte Partizipation" beschrieben?
   a. Form der sozialen Interaktion, bei der informierte Menschen (Experten) Handlungsgelegenheiten so organisieren, dass uninformierte Menschen (Novizen) etwas lernen (z.B. gemeinsames Anlegen eines Beets)
   b. Erreichung eines praktischen Ziels wird von Lernprozessen begleitet
   c. Kinder übernehmen immer aktivere und verantwortlichere Rollen

23. Was wird mit dem Begriff Intersubjektivität beschrieben?
   a. wechselseitiges Verstehen in der zwischenmenschlichen Kommunikation (Beziehen auf gemeinsame Inhalte und aufeinander bezogenes Reagieren -> fördert z.B. Konzeptbildung)
   b. 2-3 Monate: Säuglinge lebhafter & interessierter, wenn auf ihre Aktionen reagiert wird
   c. Modelllernen und joint attention (Aufmerksamkeit auf gemeinsamem Bezugspunkt)
   d. Perspektivenübernahme -> z.B. wichtig, um sich gegenseitig etwas beizubringen

24. Welche praktischen Implikationen ergeben sich aus soziokulturellen Theorien für die pädagogische Praxis?
   a. Rolle der Lehrpersonen und individueller Anpassung von Lehre
   b. Berücksichtigung kulturspezifischen Wissens beim Lernen
   c. Frühes Schaffes interaktiver Erfahrungskontexte
   d. Einsatz von Lerngemeinschaften im Unterricht (fördert Verstehen und Lernen, Zusammenhalt und Motivation)

25. Wie wird Entwicklung im Rahmen der Theorien dynamischer Systeme definiert?
   a. Entwicklung als sich selbst organisierender Prozess
   b. Integration verschiedener Komponenten zur Anpassung an Umwelt
   c. Kognitive Entwicklung nur aus sich heraus, ohne Einfluss anderer Menschen
   d. Denken und Handeln als zwei unabhängige Prozesse

Fragen zur Entwicklung der Intelligenz

Fragen Übungstest	1, 2, 3, 4, 5, 6, 7, 8, 9, 11, 12, 13, 14, 15, 24
Fragen kriterialer Test	2, 4, 5, 6, 8, 9, 10, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25
Art kriterialer Test	Multiple Choice Fragen

*die hier aufgeführte Reihenfolge der Fragen ergibt die Nummerierung im Datenblatt.

Die hier aufgeführten Stichpunkte sind – sofern nichts grün hervorgehoben ist - als komplette Lösung, nicht als MC-Frage zu verstehen! Manche Stichpunkte sind ausführlicher formuliert, als die MC-Fragen für die Probanden.

1. Welche Fähigkeiten beschreibt die Intelligenz? (offene Frage)
   a. Anpassungsfähigkeit an neue Gegebenheiten
   b. Flexibles Reagieren auf herausfordernde Situationen
2. Wozu wird ein Intelligenzkonstrukt benötigt?
   a. Diagnose von Hochbegabung
   b. Diagnose von Minderbegabung
   c. Vorhersage beruflicher Leistungen
   d. Vorhersage Schulischer Leistungen
3. Wie ist der Generalfaktor der Intelligenz bei Einfaktorenmodellen beschrieben?
   a. g = bereichsunspezifische Fähigkeit, Aufgaben zu lösen -> hat Einfluss auf alle spezifischen aufgabenbezogenen Konpetenzen
4. Welche zwei Faktoren nimmt Cattell in seinem Modell an? Beschreiben Sie in wenigen Stichpunkten die Hauptmerkmale dieser Faktoren!
   o Fluide Intelligenz
     ▪ non-verbal
     ▪ angeboren
     ▪ neuen Situationen anpassen
     ▪ induktives Denken
   o Kristalline Intelligenz
     ▪ Verbal
     ▪ erworben
     ▪ Faktenwissen über die Welt
     ▪ Bereichsspezifisch
   a. Fluide und kristalline Intelligenz
   b. Verbale und nonverbale Intelligenz
   c. Logisches Denken und Informationsverarbeitung
   d. Deklarative und non-deklarative Intelligenz
5. Wie viele Faktoren gibt es in Thurstones Primärfaktorenmodell?
   a. 5
   b. 6
   c. 7
   d. 8
6. Welche zwei Ebenen werden im kognitionspsychologischen Modell von Campione und Brown unterschieden? Beschreiben Sie diese in wenigen Stichpunkten!
   a. Architekturebene
     i. Speicherkapazität
     ii. Effizienz
     iii. Verarbeitungs-/Abrufgeschwindigkeit
   b. Exekutive Ebene
     i. Deklaratives & prozedurales Wissen
     ii. Metakognitives Wissen

     iii. Metakognitive Kontrollprozesse
  c. Deklarative Ebene
  d. Prozedurale Ebene

7. Friedrich ist ein aufgewecktes Kind und seinen Klassenkameraden in vielen Leistungsbereichen voraus. Auf Anraten der Lehrerin lassen seine Eltern ihn bei einer Psychologin testen. Es wird ein IQ von 120 ermittelt. Wie ist dieser Wert einzuschätzen?
  a. Überdurchschnittlicher Intelligenzquotient

8. Welcher Teil der Intelligenz ist ein besonders guter Prädiktor für schulische Leistungen?
  a. Kristalline Intelligenz
  b. Kreative Intelligenz
  c. Fluide Intelligenz
  d. Generelle Intelligenz

9. Wie entwickelt sich die Intelligenz über die Lebensspanne? Beschreiben Sie in 1-2 Sätzen!
  a. Fluide Intelligenz nimmt bis 25. Lebensjahr zu -> Maximum -> danach Abnahme
  b. Kristalline Intelligenz nimmt bis zum Erwachsenenalter zu und bleibt dann weitestgehend stabil
  c. Kristalline Intelligenz nimmt bis 25. Lebensjahr zu -> Maximum -> danach Abnahme
  d. Fluide Intelligenz nimmt bis zum Erwachsenenalter zu und bleibt dann weitestgehend stabil

10. Wie stabil sind die individuellen Intelligenztest-Leistungen?
  a. Sind ab dem Grundschulalter stabil
  b. Sind erst im Erwachsenenalter stabil
  c. Unterscheiden sich immer völlig
  d. Sind schon im Kindergartenalter stabil

11. Wie wirkt sich der pädagogische Kontext auf die Entwicklung der (non-)verbalen Intelligenz aus (Schneider & Stefanek, 2004)?
  a. Früh auftretende Unterschiede bleiben bestehen
  b. Werden durch pädagogischen Kontext = verschiedene Schulformen noch verstärkt

12. Wie nennt sich der Effekt, dass der IQ um ca. 3 Punkte pro Jahrzehnt steigt und was sind mögliche Erklärungen? Nennen Sie 3 in Form von Stichpunkten!
  a. Flynn-Effekt

13. Was sind mögliche Erklärungen für den Flynn-Effekt?
  a. verbesserte Ernährung und Gesundheitsversorgung
  b. Verbesserungen im Bildungssystem und Erhöhung der Schuldauer
  c. wachsende Vertrautheit mit dem Aufgabenmaterial und komplexen Denkaufgaben
  d. Verbreitung von Lernspielzeug und Übungsmaterial

14. Welche Kritik an der Intelligenzmessung, wie sie heute betrieben wird, könnte man äußern? Nennen Sie 2 Kritikpunkte.
  a. Die meisten Intelligenztests erfassen die zugrundeliegenden Fähigkeiten nicht umfassend genug
  b. Konzentrieren sich zu wenig auf schulisch relevante Fähigkeiten
    i. Richtig wäre: zu viel
  c. Viele Tests sind kulturell verzerrt
  d. Intelligenztests nicht überall anerkannt
   Weitere: Ethisch fragwürdig -> Reduktion so komplexen Konstrukts auf eine Zahl (IQ); Verwendung von Intelligenztests kann weitreichende und mitunter unerwünschte Konsequenzen haben -> IQ als willkürliches Cut-Off-Kriterium

15. Welchen Einfluss haben Anlage und Umwelt auf die Intelligenz? Welche Wechselwirkungen existieren? Beschreiben Sie in 1-2 Sätzen!
    a. Beide haben Einfluss auf IQ, der bedeutsam ist
    b. Wechselwirkung zwischen Anlage und Umwelt:
        i. Bestimmte Anlagen führen dazu, dass man sich bestimmte Umwelten sucht
        ii. Intelligenzunterschiede in Unterschicht eher durch Umwelteinflüsse bedingt, in Oberschicht eher genetisch bedingt
        iii. Intelligenz nimmt mit steigender sozialer Schicht zu -> genetisches Potenzial zu höherer Intelligenz kann sich nur dann förderlich auswirken, wenn sie in entsprechenden Umwelten gefördert wird
    ALS MC
    a. Umwelt und Anlage sind unabhängig voneinander, es gibt keine Wechselwirkungen
    b. Umwelt und Anlage haben beide bedeutsamen Einfluss auf den IQ
    c. Bestimmte Anlage führt dazu, dass man sich bestimmte, passende Umwelt sucht
    d. Intelligenzunterschiede in Oberschicht eher durch Umwelt, in Unterschicht eher durch Genetik bestimmt

16. Bei welchen Intelligenzbereichen konnten Unterschiede sich zwischen Hauptschülern und Gymnasiasten beobachtet werden?
    a. Verbale Intelligenz
    b. Nonverbale Intelligenz
    c. Allgemeine Intelligenz
    d. Alle drei

17. Wie viele und welche Schichten nimmt Carroll in seinem Modell der Intelligenz an? Nennen Sie diese
    a. Faktoren mittlerer Allgemeinheit
    b. Spezifische Prozesse/Aufgaben/Maße
    c. Verbale & Kristalline Intelligenz
    d. Prozedurale Intelligenz
    Weitere: Allgemeine Intelligenz

18. Was trifft auf den Zusammenhang zwischen Intelligenz und Schulleistung zu?
    a. In der Grundschule spielt Intelligenz eine größere Rolle als in der weiterführenden Schule.
    b. Nonverbale Intelligenz ist der bessere Prädiktor für Leistungen in der Grundschule als verbale Intelligenz.
    c. Zwischen Intelligenz und Schulleistung gibt es nur einen verschwindend kleinen Zusammenhang.
    d. Intelligenz bestimmt während der gesamten Schulzeit in entscheidendem Maß, wie die Leistungen ausfallen.

19. Was bedeutet es, wenn der Heritabilitätskoeffizient von Intelligenz bei 50% liegt?
    a. Dass bei 50% aller Menschen die Gene, bei den anderen 50% die Umwelt wichtiger ist für die Intelligenzentwicklung.
    b. Dass die Intelligenz eines Menschen zu 50% durch seine Gene beeinflusst wird.
    c. Dass 50% der Unterschiede in der Intelligenzleistung in der Stichprobe durch Gene zu erklären sind.
    d. Dass maximal 50% der Intelligenz der Stichprobe durch die Umwelt beeinflusst werden.

20. Wie beschrieb David Wechsler 1939 den IQ?
    a. Abweichungsquotient
    b. Ableitungsquotient
    c. Altersquotient
    d. Alternativquotient

21. Was gibt der Labilitätskennwert an?
    a. **wie konstant der IQ eines Individuums über verschiedene Messzeitpunkte hinweg ist**
    b. wie konstant der IQ einer Bevölkerungsgruppe über verschiedene Messzeitpunkte hinweg ist
    c. wie konstant der IQ einer Altersgruppe über verschiedene Messzeitpunkte hinweg ist
    d. dass der IQ keine Konstante, sondern ein sehr variables Konstrukt ist, dass um einen bestimmten Wert im Leben schwankt

22. Wie viele verschiedene Intelligenztypen gibt es in Gardners Theorie der multiplen Intelligenzen?
    a. 6
    b. 7
    c. **8**
    d. 9

23. Welche Aspekte der Intelligenz nutzt Sternberg zur Fundierung seiner triarchischen Begabungstheorie?
    a. **Praktische Alltagsintelligenz**
    b. **Kreative Intelligenz**
    c. **Analytische Intelligenz**
    d. Künstlerische Intelligenz

24. Welche Merkmale weisen kreative Aufgaben nach Guilford auf? Nennen Sie 2
    a. **Ziel des Problemlöseprozesses nicht genau vorgegeben**
    b. Eingeschränkte Zahl an Lösungsmöglichkeiten -> Gegenteil wäre richtig
    c. **Neuartige Kombination vorhandenen Wissens gefordert**
    d. Werden hauptsächlich durch kristalline Intelligenz gelöst

25. Was versteht man unter passiver Genom-Umwelt-Korrelation?
    a. Menschen wählen sich Umwelten, die zu ihren genetisch beeinflussten Merkmalen passen
    b. Soziale Umwelt reagiert auf genetisch bedingte Merkmale
    c. **Biologische Eltern schaffen Umwelt, die zu ihrer eigenen genetischen Disposition passt und somit vermutlich auch zu der ihrer Kinder**
    d. Soziale Umwelt versucht ohne Rücksicht auf genetisch bedingte Merkmale zu handeln

Fragen zum Vermeidungslernen & Erlernte Hilflosigkeit

Fragen Übungstest	1, 2, 3, 4, 6, 7, 8, 9, 11, 12, 13, 14, 17, 18, 19
Fragen kriterialer Test	1, 2, 3, 5, 6, 7, 8, 10, 12, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25
Art kriterialer Test	Multiple Choice Fragen

*die hier aufgeführte Reihenfolge der Fragen ergibt die Nummerierung im Datenblatt.

1. Unter welche Art von Verstärkung fällt das Vermeidungslernen? (offen oder MC)
    a. *Negative Verstärkung*
    b. Positive Verstärkung
    c. Präsentationsbestrafung
    d. Entzugsbestrafung

2. Wie sähe der klassische Experimentalaufbau aus, um Vermeidungslernen zu demonstrieren? Beschreiben Sie diesen knapp!
    Zweiteiliger Raum wird abgedunkelt, danach erhält Tier in bestimmter Hälfte Elektroschock; in anderer Hälfte nicht; Hund bemerkt, dass er Schocks entgehen kann; erkennt Dunkelheit als konditionierten Stimulus für Schocks -> springt nach wenigen Durchgängen bereits vor dem Schock weg, sobald es dunkel wird.
    Oder als offene Frage:

3. Welcher Befund in einem Experimentalaufbau wie bei Solomon et al. Würde für Vermeidungslernen sprechen?
    a. Der Hund springt im Dunklen immer nach dem Elektroschock weg.
    b. Der Hund springt nach wenigen Durchgängen bereits vor dem Schock weg, sobald es dunkel wird.
    c. Der Hund springt, egal welche Lichtverhältnisse herrschen, immer hin und her, um möglichen Schocks zu entgehen.
    d. Der Hund bleibt regungslos und winselnd liegen und erträgt die Schocks.

4. Welche Vorgänge sind nach der Zwei-Faktoren-Theorie von Mowrer an Vermeidungsreaktionen beteiligt? (offen oder MC)
    a. Klassisches und operantes Konditionieren
    b. Nur das klassische Konditionieren
    c. Nur das operante Konditionieren
    d. Es handelt sich nicht um Konditionierung

5. Was sind die Probleme an der Zwei-Faktoren-Theorie im Kontext Vermeidungslernen? Nennen Sie 2!
    a. Vermeidung ohne erkennbare Zeichen von Angst
    b. Fluchtreaktion verhindert Paarung -> Löschung sollte erfolgen
    c. Löschung erfolgt nur, wenn Vermeidungsverhalten verhindert wird

6. Welcher Mechanismus ist der einzig existierende in der Ein-Faktoren-Theorie?
    a. Klassisches Konditionieren
    b. Operantes Konditionieren
    c. Erwartungen
    d. Motivation

7. Weshalb ist durch Vermeidungslernen gelerntes Verhalten besonders löschungsresistent laut der kognitiven Theorie?
    Aufgrund gebildeter Erwartungen, die sich nicht entkräften lassen
    Oder als offene Frage:
    Weshalb ist durch Vermeidungslernen gelerntes Verhalten besonders löschungsresistent laut der kognitiven Theorie?
    a. Motivation
    b. Emotion
    c. Erwartungen
    d. Persönlichkeit

8. Wann verändert sich Verhalten in einer Vermeidungsaufgabe?
    a. Wenn Übereistimmung zwischen Erwartung und Realität herrscht
    b. Wenn Diskrepanz zwischen Erwartung und Realität herrscht
    c. Wenn Übereinstimmung zwischen Emotion und Motivation herrscht
    d. Wenn Diskrepanz zwischen Emotion und Motivation herrscht

9. Wodurch kann erlerntes Vermeidungsverhalten gelöscht werden? (offen oder MC)
    a. Flooding
    b. Blocking
    c. Shocking
    d. Smoothing

10. Was sind die Prinzipien hinter Flooding laut 2-Faktoren Theorie und kognitiver Theorie? (OFFENE FRAGE)
    a. 2-Faktoren-Theorie: Löschung einer klassisch konditionierten Furchtreaktion
    b. Kogn. Theorie: Korrektur unzutreffender Kontingenzerwartungen

11. Welche Aussage zur systematischen Desensibilisierung ist korrekt?
    a. **Wird auch Schwellentechnik genannt**
    b. Wird auch Methode des vorsichtigen Experimentierens genannt
    c. **Stimulus wird so schwach präsentiert, dass die Reaktion noch nicht ausgelöst wird**
    d. Stimulus wird präsentiert, wenn Reaktion gerade nicht auftreten kann
12. Welche Anzeichen sprechen für erlernte Hilflosigkeit?
    a. *Die Situation ist aversiv, ein Ausweg wäre durch eine Reaktion aber möglich.*
    b. *Das gezeigte Verhalten ist resignativ.*
    c. *Das gezeigte Verhalten ist passiv.*
    Oder als offene Frage:
    a. Die Situation ist aversiv, ein Ausweg wäre durch eine Reaktion aber möglich.
    b. Das gezeigte Verhalten ist resignativ.
    c. Das gezeigte Verhalten ist passiv.
    d. **Die Situation ist aversiv und bietet keinen Ausweg.**
13. Welche Reaktion zeigen Menschen, die unlösbare Aufgaben erhalten hatten und nun mit erlernter Hilflosigkeit reagieren?
    a. Spätere lösbare Aufgaben gingen sie mit doppelter Motivation an
    b. **Spätere lösbare Aufgaben trauten sie sich weniger zu**
    c. Spätere lösbare Aufgaben ließen sie länger durchhalten
    d. **Spätere lösbare Aufgaben bearbeiteten sie nicht immer bis zum Ende**
14. Welche Konsequenzen hat erlernte Hilflosigkeit auf motivationaler, kognitiver und affektiver Ebene? (OFFENE FRAGE)
    a. Unkontrollierbarkeitserfahrung
    b. Erwartung nicht kontrollierbarer Umwelt
    c. Affektive Störung/Depressivität
15. Was beschreibt der Begriff Attribution?
    a. **Attribution bedeutet Ursachen zuschreiben**
    b. Attribution bedeutet Folgen vorhersagen
    c. **Ein Beispiel für Attribution ist, eine gute Note dem Zufall zuzuschreiben**
    d. Ein Beispiel für Attribution ist, sich zu überlegen, welche Folgen eine gute Note wohl

16. Welche Attributionsstile existieren wirklich?
    a. **global-internal-stabil**
    b. global-external-spezifisch
    c. spezifisch-stabil-variabel
    d. **spezifisch-external-variabel**
17. Welche Dimension beschreiben die Begriffe „global & spezifisch" in der Attributionstheorie?
    a. Stabilitätsdimesion
    b. Lokationsdimension
    c. **Kontrollierbarkeitsdimension**
    d. Sicherheitsdimension
18. Sie haben eine gute Note in einer Klausur erhalten. Sie vermuten, die Klausur war sehr leicht. Welchem Attributionsstil entspricht dies?
    a. global-internal-stabil
    b. global-internal-variabel
    c. **spezifisch-external-variabel**
    d. spezifisch-external-stabil
19. Woran lässt sich erlernte Hilflosigkeit im Unterricht erkennen?
    a. Misserfolge werden external stabil attribuiert
    b. **Fehlende Unterrichtsbeteiligung**
    c. Hohe Motivation, aber trotzdem schlechte Leistungen

       d. Geringe Ausdauer bei Schwierigkeiten

20. Mit welchem klinischen Krankheitsbild ist das Verhalten unter erlernter Hilflosigkeit gut zu vergleichen? (offen oder MC)

    *a. **Major Depression***
    b. Angststörung
    c. Autismus
    d. Posttraumatische Belastugnsstörung

21. Was kann helfen, um gegen erlernte Hilflosigkeit vorzugehen? (offen oder MC)

    a. *Attributionsmuster und Kontrollüberzeugungen verändern*
    b. Expositionstherapie
    c. Vor Misserfolgen schützen, z.B. durch einfache Tests
    d. Leider gar nichts

22. Einer Ihrer Schüler hat sichtbar Angst davor, in der Klasse vorzutragen oder auf im Unterricht gestellte Fragen antworten zu müssen. Sie vermuten eine soziale Angst bei dem Schüler und möchten ihm zusammen mit einem befreundeten Therapeuten mit der Methode des Floodings helfen. Was müssen Sie beim Flooding beachten? (am besten MC)

    a. Sie müssen eine Hierarchie der angstauslösenden Reize erstellen und diese schrittweise abarbeiten.
    b. Sie sollten die Sitzung unterbrechen, sobald der Schüler sichtbare Angst zeigt.
    c. Auch wenn es für beide Seiten unangenehm sein kann, sollte sich der Schüler beim Flooding so lange einem angstauslösenden Reiz aussetzen, bis die Angst nachgelassen hat.
    d. Sie sollten dem Schüler ein Referat nur vor Ihnen (der Lehrkraft) halten lassen.

23. Franz hat beim Sportunterricht ein Eigentor geschossen, wurde von seinen Mitschülern ausgelacht und ist aus der Sporthalle gerannt. In den Folgewochen hat Franz jedes Mal, wenn der Sportlehrer die Fußbälle holt, neue Ausreden, warum er nicht am Sportunterricht teilnehmen kann. Welche 2 Prozesse spielen bei diesem Vermeidungsverhalten laut **Zwei-Faktoren-Theorie** eine Rolle?

    a. Franz hat die Erwartung entwickelt, dass er beim Fußball wieder einen Fehler macht und die anderen Schüler ihn wieder auslachen werden. Zusätzlich erwartet er, dass er diesem entgehen kann, wenn er dem Sportunterricht fernbleibt. Wegen dieser zwei Erwartungen vermeidet er es, beim Sportunterricht mit Fußball zu spielen.
    b. Das Schießen des Eigentors und das Lachen der anderen Schüler hat bei Franz Angst ausgelöst, die nun in neuen Situationen durch den Anblick der Fußbälle wieder ausgelöst wird. Das Verlassen der Turnhalle ist die Flucht vor dem angstauslösenden Anblick der Fußbälle.
    c. Das Verlassen der Sporthalle dient Franz als Verstärker. In der Sporthalle zu bleiben, wäre eine Bestrafung.
    d. Franz ist nicht mehr motiviert, am Sportunterricht teilzunehmen, da die Mitschüler ihn schlecht behandelt haben. Die negativen Emotionen vermeidet er, indem er einfach wegläuft.

24. Franz hat beim Sportunterricht ein Eigentor geschossen, wurde von seinen Mitschülern ausgelacht und ist aus der Sporthalle gerannt. In den Folgewochen hat Franz jedes Mal, wenn der Sportlehrer die Fußbälle holt, neue Ausreden, warum er nicht am Sportunterricht teilnehmen kann. Welche 2 Prozesse spielen bei diesem Vermeidungsverhalten laut **kognitiver Theorie** eine Rolle?

    a. Franz hat die Erwartung entwickelt, dass er beim Fußball wieder einen Fehler macht und die anderen Schüler ihn wieder auslachen werden. Zusätzlich erwartet er, dass er diesem entgehen kann, wenn er dem Sportunterricht fernbleibt. Wegen dieser zwei Erwartungen vermeidet er es, beim Sportunterricht mit Fußball zu spielen.

b. Das Schießen des Eigentors und das Lachen der anderen Schüler hat bei Franz Angst ausgelöst, die nun in neuen Situationen durch den Anblick der Fußbälle wieder ausgelöst wird. Das Verlassen der Turnhalle ist die Flucht vor dem angstauslösenden Anblick der Fußbälle.

c. Das Verlassen der Sporthalle dient Franz als Verstärker. In der Sporthalle zu bleiben, wäre eine Bestrafung.

d. Franz ist nicht mehr motiviert, am Sportunterricht teilzunehmen, da die Mitschüler ihn schlecht behandelt haben. Die negativen Emotionen vermeidet er, indem er einfach wegläuft.

25. Sie passen auf eine Katze und einen Kanarienvogel Ihrer Nachbarin auf. Während Sie in der Küche Ihr Lieblingslied hören und das Futter für die Tiere vorbereiten, fällt Ihnen eine Metallschüssel mit lautem Knall runter. Als sie das nächste Mal Ihr Lieblingslied summend zum Füttern kommen, finden Sie den Vogel reglos im Käfig, die Katze fauchend unter dem Bett vor. Erklären Sie das so unterschiedliche Verhalten der Tiere!

a. Die beiden Tiere verhalten sich unterschiedlich, weil spezies-spezifische Abwehrreaktionen ihr Vermeidungsverhalten beeinflussen.

b. Die beiden Tiere verhalten sich unterschiedlich, weil Katzen, im Gegensatz zu Vögeln, kein Vermeidungsverhalten erlernen können, da dieses in ihrem Verhaltensrepertoire biologisch nicht angelegt ist.

c. Preparedness kann das unterschiedliche Verhalten von Katze und Vogel verursachen.

d. Blocking könnte dafür sorgen, dass die Katze das Verhalten nicht erlernen kann.

26. Jessica ist eine mittelmäßige Schülerin, die leider seit einigen Wochen jegliche Mitarbeit im Unterricht verweigert. Sie bearbeitet die Hausaufgaben nicht und beteiligt sich auch nicht mehr am Unterricht. Es geht Jessica damit sichtlich schlecht. Was könnten Gründe für ihr Verhalten sein?

a. Jessica hat im letzten Schuljahr mehrheitlich die Erfahrung gemacht, dass ihr eigenes Handeln und ihre Anstrengungen nicht zu positiven Konsequenzen führen. Durch den scharfen Wettbewerb in der Klasse ist sie unter den Schülern, die immer wieder schulischen Misserfolg haben. Daher glaubt sie nicht mehr daran, dass ihr Verhalten positive Folgen nach sich ziehen wird.

b. Jessica hat eine Reihe an neuen Freundinnen aus einer Nachbarschule, die Schule doof und unnötig finden. Um ihren Freundinnen zu imponieren, macht auch Jessica im Unterricht nicht mehr mit.

c. Jessica hat ein motivationales Defizit. Ihr fehlt der Antrieb, Ereignisse durch ihr eigenes Verhalten zu beeinflussen.

d. Jessica hat ein kognitives Defizit, wodurch sie nicht in der Lage ist, länger andauernde und stärkere Gedächtnisspuren für die vermittelten Sachverhalte zu entwickeln. So vergisst sie eben Gehörtes und die Hausaufgaben.

Vorwissenstests (Deutsche Version / German version – for English see DVD)

Vorwissenstest Lese-Rechtschreib-Schwierigkeiten

-   Was sind Lernstörungen?
    o   **Sowohl Lesestörungen, als auch Konzentrationsstörungen zählen zu den Lernstörungen**
    o   **Erwerb einer Fähigkeit von frühen Entwicklungsstadien an beeinträchtigt**
    o   Lernstörungen sind selten neurologisch bedingt
    o   Lernstörungen lösen sich auch ohne therapeutische Maßnahmen im Jugendalter meist auf

Appendix

- Wann ist etwas eine Lernstörung?
  - o **Leistungsrückstände betragen mehrere Schuljahre**
  - o Können Folge schlechten Unterrichts oder unzureichenden Lernangebots sein
  - o **Es darf keine Intelligenzminderung vorliegen (IQ < 70)**
  - o **Leistung ist geringer, als aufgrund des IQs vermutet würde**
- Was bezeichnet der Begriff Vorläuferfähigkeiten?
  - o Fähigkeiten, die eine gewisse Vorlaufzeit brauchen, bis sie erworben bzw. gut ausgeführt werden können
  - o **Fähigkeiten, die vor einer anderen Fähigkeit erworben werden müssen**
  - o **Ein Beispiel für Vorläuferfähigkeiten zum Schriftspracherwerb könnte das Kritzeln von Buchstaben sein.**
  - o Ein Beispiel für eine Vorläuferfähigkeit könnte das Schreiben sein, da es etwas dauert, bis es flüssig ausgeführt werden kann.
- Welche Fähigkeiten müssen vorliegen, bevor der Schriftspracherwerb stattfinden kann?
  - o **Reimwörter finden**
  - o **Erkennen, wo Worte aufhören und anfangen**
  - o Buchstaben schreiben können
  - o Grammatikregeln im Mündlichen korrekt umsetzen
- Woran ist eine Lese-Rechtschreib-Schwäche zu erkennen?
  - o **Lese-Rechtschreibleistung ist weit unter den erwarteten Durchschnittsleistungen der Altersgruppe**
  - o Intelligenz kann vermindert sein
  - o Auf neurologischer Ebene sind keine Unterschiede zu Gesunden zu erkennen
  - o **Der Verlauf einer LRS ist stabil, wenn nicht mit therapeutischen Maßnahmen interveniert wird**
- Was ist ein Phonem?
  - o **Fachbegriff für Laut**
  - o Fachbegriff für Buchstabe
  - o Fachbegriff für Silbe
  - o Fachbegriff für Bedeutungseinheit
- Was ist das ICD 11?
  - o **Ein Nachschlagewerk für alle Krankheiten**
  - o Ein Nachschlagewerk ausschließlich für Lernstörungen
  - o Ein Nachschlagewerk ausschließlich für psychische Erkrankungen
  - o Ein Nachschlagewerk für Therapiemöglichkeiten bei bestimmten Krankheiten
- Was bedarf es, um einen gelesenen Text zu verstehen?
  - o **Einen gewissen Wortschatz mit Basis- und Fachbegriffen**
  - o Vorwissen wird nicht benötigt
  - o Transferleistungen sind bei den meisten Texten zum Verständnis nicht nötig
  - o **Zusammenhang der Informationen innerhalb des Textes sollte erkannt werden**
- Wie ist die Methode bei Zwillingsstudien?
  - o **Untersuchung getrennt aufgewachsener eineiiger Zwillinge hinsichtlich bestimmter Merkmale**
  - o **Vergleich der Ähnlichkeiten gemeinsam aufgewachsener eineiiger und zweieiiger Zwillinge**
  - o Vergleich verschiedener Zwillingspaare miteinander hinsichtlich bestimmter Merkmale
  - o Vergleich von Menschen, die sich in Situationen ähnlich verhalten würden, also „Persönlichkeits-Zwillinge" sind

- Was soll mit Zwillingsstudien untersucht werden?
  - o Durch die Zwillingsforschung versucht man herauszufinden, ob die Geburtszeit Auswirkungen auf den Charakter hat. Da Zwillinge eng miteinander verwandt sind, aber nicht gleichzeitig geboren werden, eignen sie sich besonders gut.
  - o Durch die Zwillingsforschung versucht man zu überprüfen, ob Zwillinge häufiger männlich oder weiblich sind.
  - o **Durch die Zwillingsforschung versucht man herauszufinden, ob die Gene oder eher die Umwelteinflüsse den Charakter und den Gesundheitszustand eines Menschen beeinflussen.**
  - o **Durch die Zwillingsforschung versucht man zu klären, inwiefern die Erziehung und das soziale Umfeld einen Menschen formen.**
- Was ist mit dem Begriff Kortex gemeint?
  - o Das gesamte Gehirn
  - o Das Bewusstsein des Menschen
  - o **Die Hirnrinde**
  - o Der Hirnstamm
- In welchem Alter sollte Kinder die sog. 50-Wort-Grenze erreicht haben?
  - o 12 Monate
  - o **18 Monate**
  - o 24 Monate
  - o 30 Monate
- Was bedeutet der Begriff Epidemiologie?
  - o **Wissenschaft der Entstehung, Verbreitung und Bekämpfung von Krankheiten**
  - o Wissenschaft der körperlichen Auswirkungen von psychischen Krankheiten
  - o Wissenschaft der psychischen Auswirkungen von körperlichen Krankheiten
  - o Wissenschaft der Geschichte von Krankheiten
- Welche Kriterien muss eine Verhaltensstörung erfüllen?
  - o **Inadäquates Verhalten**
  - o Muss nicht zeitlich überdauernd sein
  - o **Können durch Erziehung ausgelöst werden**
  - o Entwickeln sich immer im Kindesalter
- Was ist eine Komorbidität?
  - o Das gleichzeitige Auftreten mehrerer Krankheiten bei einem Patienten.
  - o **Das Auftreten zusätzlicher Erkrankungen im Rahmen einer definierten Grunderkrankung.**
  - o Gesundheitsschäden, welche durch ärztliche Therapiemaßnahmen entstehen.
  - o Die Wahrscheinlichkeit einer Spontanheilung einer Krankheit.

Vorwissenstest ADHS

- Was ist Aufmerksamkeit?
  - o **Selektion**
  - o **Orientierung**
  - o **Wachsamkeit**
  - o Aktivität
- Was zeichnet eine ADHS aus?
  - o **Aufmerksamkeitsstörung**
  - o Hypersensibilitätsstörung
  - o **Impulsivität**
  - o **Hyperaktivität**

Appendix

- ADHS wird ausgelöst durch…?
  - o Erziehungsprobleme
  - o **Neurologische Ursachen**
  - o **Genetik**
  - o Bewegungsmangel im Kleinkindalter
- ADHS…
  - o …entwickelt sich erst im Grundschulalter
  - o …verschwindet meist bis zum Erwachsenenalter von allein
  - o **…macht sich bereits im Kindergarten bemerkbar**
  - o **…verlagert sich im Jugendalter eher nach innen (innere statt motorische Unruhe)**
- Welche Behauptung/en über ADHS stimmt/stimmen?
  - o **Jungs sind häufiger von ADHS betroffen als Mädchen.**
  - o Linkshänder sind häufiger betroffen als Rechtshänder.
  - o ADHS kann durch zu viel Smartphonekonsum ausgelöst werden.
  - o **ADHS gibt es auch im Erwachsenenalter.**
- Dopamin, dessen Regulation bei ADHS gestört ist, …
  - o **… ist ein Neurotransmitter im Gehirn**
  - o … ist ein Hormon im Körper
  - o … kann medikamentös nicht beeinflusst werden
  - o **… wird zu schnell abgebaut**
- Was ist das ICD 11?
  - o **Ein Nachschlagewerk für alle Krankheiten**
  - o Ein Nachschlagewerk ausschließlich für Lernstörungen
  - o Ein Nachschlagewerk ausschließlich für psychische Erkrankungen
  - o Ein Nachschlagewerk für Therapiemöglichkeiten bei bestimmten Krankheiten
- Wie ist die Methode bei Zwillingsstudien?
  - o **Untersuchung getrennt aufgewachsener eineiiger Zwillinge hinsichtlich bestimmter Merkmale**
  - o **Vergleich der Ähnlichkeiten gemeinsam aufgewachsener eineiiger und zweieiiger Zwillinge**
  - o Vergleich verschiedener Zwillingspaare miteinander hinsichtlich bestimmter Merkmale
  - o Vergleich von Menschen, die sich in Situationen ähnlich verhalten würden, also „Persönlichkeits-Zwillinge" sind
- Was soll mit Zwillingsstudien untersucht werden?
  - o Durch die Zwillingsforschung versucht man herauszufinden, ob die Geburtszeit Auswirkungen auf den Charakter hat. Da Zwillinge eng miteinander verwandt sind, aber nicht gleichzeitig geboren werden, eignen sie sich besonders gut.
  - o Durch die Zwillingsforschung versucht man zu überprüfen, ob Zwillinge häufiger männlich oder weiblich sind.
  - o **Durch die Zwillingsforschung versucht man herauszufinden, ob die Gene oder eher die Umwelteinflüsse den Charakter und den Gesundheitszustand eines Menschen beeinflussen.**
  - o **Durch die Zwillingsforschung versucht man zu klären, inwiefern die Erziehung und das soziale Umfeld einen Menschen formen.**
- Was ist mit dem Begriff Kortex gemeint?
  - o Das gesamte Gehirn
  - o Das Bewusstsein des Menschen
  - o **Die Hirnrinde**
  - o Der Hirnstamm

- Was sind psychische Störungen?
  - o Kurzfristige Veränderungen in Erleben und Verhalten
  - o **umfassen oft verminderte Selbstregulationskompetenz**
  - o **müssen Krankheitswert haben**
  - o sind selten mit einer Beeinträchtigung der Lebensführung assoziiert
- Was ist die Begriff Prävalenz einer Krankheit?
  - o **Anteil der Erkrankten zu einem bestimmten Zeitpunkt**
  - o Anteil der Neuerkrankten innerhalb eines bestimmten Zeitraums
  - o Anteil der Erkrankten innerhalb einer Lebensspanne
  - o Anteil der zu erwartenden Neuerkrankungen im Folgejahr
- Was ist eine Komorbidität?
  - o Das gleichzeitige Auftreten mehrerer Krankheiten bei einem Patienten.
  - o **Das Auftreten zusätzlicher Erkrankungen im Rahmen einer definierten Grunderkrankung.**
  - o Gesundheitsschäden, welche durch ärztliche Therapiemaßnahmen entstehen.
  - o Die Wahrscheinlichkeit einer Spontanheilung einer Krankheit.
- Welche Kriterien muss eine Verhaltensstörung erfüllen?
  - o **Inadäquates Verhalten**
  - o Muss nicht zeitlich überdauernd sein
  - o **Können durch Erziehung ausgelöst werden**
  - o Entwickeln sich immer im Kindesalter
- Das Exekutive System im Gehirn ist zuständig für…
  - o Reaktionssteigerung
  - o **Geistige Flexibilität**
  - o **Arbeitsgedächtnis**
  - o **Inhibition**

Vorwissenstest Entwicklung des Denkens

- Was versteht man unter einem konstruktivistischen Ansatz in Bezug auf das Denken?
  - o Es wird davon ausgegangen, dass die Realität nur konstruiert, nur Einbildung ist.
  - o **Es wird davon ausgegangen, dass jegliches Wissen durch Erfahrung selbst konstruiert wird.**
  - o Es wird davon ausgegangen, dass das Denken hauptsächlich in Form von abstrakten Konstrukten abläuft.
  - o Es wird davon ausgegangen, dass die konstruierte Weltsicht der Gesellschaft von den Individuen im Laufe ihrer Entwicklung angenommen wird.
- Was bedeutet der Begriff Akkommodation?
  - o **Anpassung**
  - o Entfremdung
  - o Ausbalancierung
  - o Integration
- Was ist der Unterschied zwischen diskontinuierlichen und kontinuierlichen Prozessen?
  - o **Kontinuierliche Prozesse laufen übergangslos ineinander über**
  - o **Bei diskontinuierlichen Prozessen lassen sich einzelne Entwicklungsstufen eindeutig abgrenzen**
  - o Bei kontinuierlichen Prozessen ist stetig Verbesserung erkennbar, meist in Form von großen Sprüngen
  - o **Das Kochen von Wasser auf dem Herd ist ein Beispiel für einen kontinuierlichen Prozess**

- Was bedeutet ein qualitativer Unterschied?
  - o Etwas verändert sich hinsichtlich seiner absoluten Menge
  - o **Etwas verändert sich hinsichtlich der Art und Weise**
  - o Etwas verändert sich hinsichtlich seiner Güte
  - o Etwas verändert sich hinsichtlich seiner Häufigkeit
- Was bedeutet ein quantitativer Unterschied?
  - o **Etwas verändert sich hinsichtlich seiner absoluten Menge**
  - o Etwas verändert sich hinsichtlich der Art und Weise
  - o Etwas verändert sich hinsichtlich seiner Güte
  - o **Etwas verändert sich hinsichtlich seiner Häufigkeit**
- Welche Komponenten gehören zum Mehrspeichermodell von Atkinson & Shiffrin?
  - o **Sensorisches Gedächtnis**
  - o **Arbeitsgedächtnis**
  - o Erinnerungsgedächtnis
  - o Prozedurales Gedächtnis
- Das Exekutive System im Gehirn ist zuständig für…
  - o Reaktionssteigerung
  - o **Geistige Flexibilität**
  - o **Arbeitsgedächtnis**
  - o **Inhibition**
- Der Begriff Enkodieren meint…?
  - o **Das Speichern von Infos im Gehirn**
  - o Das Verpacken von Infos in kurzen Abschnitten, ähnlich Codes beim Programmieren
  - o Den Abruf von verschlüsselten Infos aus dem Gehirn
  - o Eine Merkstrategie, mit der sich mehr Infos als üblich gemerkt werden können
- Was passiert bei der sogenannten Myelinisierung im Gehirn?
  - o Es bilden sich neue Synapsen, um mehr Information zu speichern.
  - o **Nervenfasern werden von Membranen umkleidet, um Informationen schneller zu leiten.**
  - o Strukturen des Gehirns in tieferen Hirnregionen werden aktiv, um Wissen langfristiger zu speichern.
  - o Alte, ungenutzte Verbindungen zwischen Gehirnregionen werden aufgelöst, um Platz für neue Verbindungen zu schaffen.
- Welcher Aussage würden Sie zustimmen?
  - o **Wie wir denken ist in großem Ausmaß auch von unserer Kultur abhängig.**
  - o Wie wir denken wird kaum durch geschichtliche oder kulturelle Gegebenheiten beeinflusst.
  - o Wie wir denken wird hauptsächlich durch genetische Eigenschaften gesteuert.
  - o **Wie wir denken hängt auch mit dem uns verfügbaren Wortschatz zusammen.**
- Was ist Modelllernen?
  - o Personen lernen anhand von (Spielzeug)Modellen bestimmte Verhaltensweisen.
  - o **Personen lernen anhand von Beobachtung anderer bestimmte Verhaltensweisen.**
  - o Personen lernen, indem sie über das Verhalten, das erlernt werden soll, einen Text lesen.
  - o Personen lernen, indem sie das geforderte Verhalten aus sich heraus ausprobieren, also modellieren.
- Welche der hier aufgeführten Stufen sind Teil der vier Entwicklungsstufen In Piagets Theorie zur Entwicklung des Denkens?
  - o **Sensomotorisches Stadium**
  - o Postoperationales Stadium
  - o Abstrakt-operationales Stadium

- o **Formal-operationales Stadium**
- Welche Aussagen zur sogenannten Objektpermanenz sind wahr?
  - o Beschreibung eines Objektes, das permanent in der Nähe des Kindes ist
  - o **Erkenntnis, dass ein Objekt nach Abdecken nicht aufhört zu existieren**
  - o Objektpermanenz bedeutet so viel wie „aus den Augen, aus dem Sinn"
  - o **Objektpermanenz muss von Kindern erst erworben werden**
- Wie könnte man ein salientes Merkmal anders beschreiben?
  - o **Als auffälliges Merkmal**
  - o Als verstecktes Merkmal
  - o Als individuelles Merkmal
  - o Als übergreifendes Merkmal
- Was ist Aufmerksamkeit?
  - o **Selektion**
  - o **Orientierung**
  - o **Wachsamkeit**
  - o Aktivität

Vorwissenstest Entwicklung der Intelligenz

- Was sind standardisierte Tests?
  - o Bei standardisierten Tests erfolgt die Auswertung nach dem Ermessen des Korrektors.
  - o **Standardisierte Tests sind objektiv, zuverlässig und valide.**
  - o **Für standardisierte Tests können Normen erhoben werden, die als Richtwerte dienen.**
  - o Eine Schulnote kann als standardisiertes Maß für dieses Leistungsfeld verwendet werden.
- Was ist Korrelation?
  - o Korrelation impliziert meist auch Kausalität, d.h. aus Korrelation kann eine Ursache abgeleitet werden.
  - o **Korrelation beschreibt eine Beziehung zwischen mehreren Variablen.**
  - o Ein Synonym für Korrelation ist Unabhängigkeit.
  - o **Korrelation spielt eine wichtige Rolle bei der Betrachtung von psychologischen Modellen.**
- Was ist ein Prädiktor?
  - o **Vorhersagevariable**
  - o Rückblicksvariable
  - o **Gibt Auskunft, inwiefern ein Test ein Merkmal beeinflusst**
  - o Ist für die Praxis kaum relevant
- Welche Aussagen über Intelligenzmodelle sind zutreffend?
  - o **Es gibt eine Vielzahl verschiedener Modelle, die keine gemeinsame Basis haben.**
  - o Es gibt eine Vielzahl verschiedener Modelle, die auf einer gemeinsamen Basis beruhen.
  - o Es gibt ein Modell, das als allgemeingültig anerkannt wurde.
- Welche Aussagen über Intelligenz sind zutreffend?
  - o **Intelligenz ist die Fähigkeit, Probleme zu lösen.**
  - o Intelligenz ist vor allem in bekannten Situationen hilfreich.
  - o Intelligenz bringt in neuartigen Situationen keinen Vorteil.
  - o **Intelligenz ist nur durch Intelligenztests messbar.**
- Was kennzeichnet Hochbegabung?
  - o Überdurchschnittliche Leistungen in nur einem ganz speziellen Bereich genügt, um von Hochbegabung zu sprechen.

- o **Motivation und Interesse werden als Bestandteile der Hochbegabung angesehen.**
  - o **Kreatives Denken und Flexibilität zeichnen eine hochbegabte Person aus.**
  - o Der IQ einer hochbegabten Person kann auch im Durchschnittsbereich liegen.
- Welche Aussagen über Intelligenz treffen zu?
  - o **Der IQ kann aus dem errechneten Intelligenzalter in Beziehung zum Lebensalter berechnet werden.**
  - o **Der IQ kann nur über Intelligenztests ermittelt werden.**
  - o Der IQ ist ein standardisiertes Maß mit dem Mittelwert 0.
  - o Ein IQ über 130 ist nicht möglich.
- Welche Aussagen über den IQ treffen zu?
  - o **Bei einem IQ unter 70 spricht man von Intelligenzminderung.**
  - o Bei einem IQ ab 150 spricht man von Hochbegabung.
  - o **Der Großteil der Bevölkerung hat einen IQ um die 100 Punkte.**
  - o **Der IQ ist ein relativer Wert, der einen Vergleich innerhalb einer Altersgruppe erlaubt.**
- Welche Aussagen über Intelligenz treffen zu?
  - o **Intelligenz ist altersabhängig**
  - o Intelligenz ist im Laufe des Lebens relativ stabil
  - o Intelligenz allein bestimmt die Leistung in Tests
  - o **Intelligenz wird mit zunehmendem Alter immer unwichtiger für schulische/berufliche Leistungen**
- Welche Aussagen sind zutreffend?
  - o **Verbale Intelligenz ist durch Erfahrung erworbenes Wissen**
  - o Non-verbale Intelligenz entwickelt sich im Kindergartenalter
  - o Verbale Intelligenz hilft beim schlussfolgernden Denken
  - o **Non-verbale Intelligenz nimmt ab dem frühen Erwachsenenalter wieder ab**
- Was ist ein Konstrukt?
  - o Direkt beobachtbare Eigenschaft
  - o **Muss durch einen Test erschlossen werden**
  - o Selten theoriegestützt
  - o **Nicht unmittelbar fassbare Eigenschaft**
- Was bedeutet der Begriff Heritabilität?
  - o **Maß für die Erblichkeit von Eigenschaften**
  - o Maß für Nützlichkeit von Eigenschaften
  - o **Heritabilität von Intelligenz ist hoch**
  - o Heritabilität von Intelligenz ist gering
- Was ist induktives Denken?
  - o Bereichsspezifisches Denken
  - o Analytisches Denken
  - o **Schlussfolgerndes Denken**
  - o Denken in Kategorien
- Das Exekutive System im Gehirn ist zuständig für…
  - o Reaktionssteigerung
  - o **Geistige Flexibilität**
  - o **Arbeitsgedächtnis**
  - o **Inhibition**
- Was ist ein Index im psychologisch-pädagogischen Kontext?
  - o Stichwortverzeichnis
  - o **Zusammenfassender Wert komplexer Einzelwerte**
  - o Durchschnittswert einer Gruppe bei einem Test
  - o Größe, die Entwicklung in einem Test misst

Vorwissenstest Vermeidungslernen

- Welche Aussagen zu Verstärkung treffen zu?
  - **Wenn ich ein Eis für eine gute Note bekomme, ist das ein Beispiel für positive Verstärkung.**
  - Wird eine Belohnung entzogen, ist das ein Beispiel für negative Verstärkung.
  - **Das Entziehen einer Strafe, zum Beispiel lauter Alarmtöne, wird als negative Verstärkung bezeichnet.**
  - **Bekomme ich einen Stromschlag, wenn ich einen Elektrozaun berühre, ist das eine positive Bestrafung.**
- Schokolade als Belohnung ist ein Beispiel für einen…
  - **Primären Verstärker**
  - Sekundären Verstärker
  - Negativen Verstärker
  - Generalisierten Verstärker
- Eine gute Note in einer Klausur ist ein Beispiel für einen…
  - Primären Verstärker
  - **Sekundären Verstärker**
  - Negativen Verstärker
  - Generalisierten Verstärker
- Was passiert bei der klassischen Konditionierung?
  - Bewegungen werden verknüpft
  - Bewegungen werden erlernt
  - **Neue Reize lösen kontrollierte Reaktionen aus**
  - **Zwei Reize werden verknüpft**
- Was ist operantes Konditionieren?
  - **Die Wahrscheinlichkeit von Verhalten verändert sich durch Belohnung und Bestrafung**
  - Die eigene Reaktion auf Reize kann die Situation bzw. Umwelt nicht verändern.
  - Ziel der operanten Konditionierung ist das Erlenen von neuem Verhalten und Fertigkeiten.
  - **Menschen prüfen, ob das eigene Verhalten eine erwünschte Auswirkung auf die Situation hat.**
- Was ist Löschung im Kontext der Konditionierung?
  - **Abschwächung oder Aufhebung einer Koppelung**
  - Abschwächung oder Aufhebung des Auslösereizes
  - Abschwächung oder Aufhebung der Reaktion
  - Abschwächung oder Aufhebung des Lernprozesses
- Was ist Vermeidungslernen?
  - **Man lernt, wie man eine bestimmte aversive Konsequenz vermeidet.**
  - Man lernt, wie man ein bestimmtes Verhalten vermeidet.
  - Man vermeidet es, etwas Neues in einer bestimmten Situation zu lernen.
  - Man vermeidet es, eine bestimmte Situation zu erzeugen, in der man Lernen müsste.
- Was ist ein konditionierter Stimulus?
  - Stimulus, der zuverlässig bestimmte Reaktion auslöst
  - **Stimulus, der ursprünglich keine Reaktion auslöst, kann erst nach der Konditionierung Reaktion auslösen**
  - **Beispiel für einen konditionierten Stimulus ist die Oma, die mit Eis verbunden wird**
  - Beispiel für einen konditionierten Stimulus ist Speichel, der beim Anblick von Eis im Mund zusammenläuft

- Was ist das sogenannte Flooding?
  - **Stimulus wird solange präsentiert, bis der Organismus die Reaktion nicht mehr durchführen kann**
  - Stimulus wird so schwach präsentiert, dass die Reaktion noch nicht ausgelöst wird
  - Stimulus wird präsentiert, wenn Reaktion gerade nicht auftreten kann
  - Stimulus wird präsentiert – folgt eine Vermeidung, wird bestraft
- Welche Aussage zur systematischen Desensibilisierung ist korrekt?
  - **Wird auch Schwellentechnik genannt**
  - Wird auch Methode des vorsichtigen Experimentierens genannt
  - **Stimulus wird so schwach präsentiert, dass die Reaktion noch nicht ausgelöst wird**
  - Stimulus wird präsentiert, wenn Reaktion gerade nicht auftreten kann
- Welche Punkte treffen auf Attributionen zu?
  - **Attribution bedeutet Ursachen zuschreiben**
  - Attribution bedeutet Folgen vorhersagen
  - **Ein Beispiel für Attribution ist, eine gute Note dem Zufall zuzuschreiben**
  - Ein Beispiel für Attribution ist, sich zu überlegen, welche Folgen eine gute Note wohl haben wird.
- Was sind typische Merkmale einer Depression?
  - **Müdigkeit**
  - **Niedergeschlagenheit**
  - Rededrang
  - Übermäßiger Appetit
- Was bedeutet Kontiguität?
  - **räumlich-zeitliche Nähe von unkonditioniertem Stimulus und konditioniertem Stimulus**
  - räumlich-zeitliche Nähe von konditioniertem Stimulus und konditionierter Reaktion
  - Vorhersagbarkeit des unkonditionierten Stimulus aus dem konditionierten Stimulus
  - Vorhersagbarkeit der konditionierten Reaktion aus dem konditionierten Stimulus
- Was bedeutet Kontingenz?
  - räumlich-zeitliche Nähe von unkonditioniertem Stimulus und konditioniertem Stimulus
  - räumlich-zeitliche Nähe von konditioniertem Stimulus und konditionierter Reaktion
  - **Vorhersagbarkeit des unkonditionierten Stimulus aus dem konditionierten Stimulus**
  - Vorhersagbarkeit der konditionierten Reaktion aus dem konditionierten Stimulus
- Was ist eine konditionierte Reaktion?
  - **Gelernte Antwort auf einen Stimulus**
  - Natürliche Reaktion auf einen Stimulus
  - **Zum Beispiel hält sich ein Kind die Ohren zu, wenn es einen Luftballon sieht, aus Angst, er könnte platzen.**
  - Zum Beispiel hält sich ein Kind die Ohren zu, weil es ein lautes Geräusch hört.

Online Data Source

All mentioned data and materials can also be accessed online via the OSF:

https://osf.io/9hw2z/?view_only=e56cf9fcb7844469a1f05e7b3cbcc2d8

Appendix B: Chapter III

Datafile & Syntax

The data file and syntax can be accessed both as an SPSS version and in Excel format on the enclosed DVD. In addition to the data files, there is also a list of the occurring variables in the analysis and a file with an explanation of these.

Data source: *Appendix B – Chapter III > Datafile & Syntax*

Materials

The questions of the practice and the criterial test can be found in German below and on the DVD. The English translations can also be found on the DVD.

The DVD also contains a file with the open answers of the participants in raw and coded form for a better understanding of the evaluation.

Data source: *Appendix B – Chapter III > Materials*

Fragen für Übungs- und kriterialen Test (Deutsche Version/ German version – English version see DVD)

Anmerkung: Fragen sind als MC sowie als offene Fragen formuliert; entweder können die Fragen direkt so gestellt werden, oder eine „Alternativfrage" für das offene Format steht unten drunter.

Die markierten Lösungen sind zugleich auch die Musterlösung für die Auswertung offener Fragen. In manchen Fällen stehen ergänzende Antwortmöglichkeiten bei den Fragen dabei, die im Falle von offenen Fragen auch in die Bewertung einfließen können.

Im Übungstest ist das Format bei allen Fragen offen. Im kriterialen Test variiert es zwischen offen und Multiple-Choice.

Fragen zu Legasthenie

1. Der Verlauf einer LRS…
   a. Ist stabil und nicht beeinflussbar.
   b. **Ist stabil, eine Therapie kann aber positiven Einfluss haben.**
   c. Ist instabil, je nach sonstiger Entwicklung des Kindes.
   d. Ist instabil und die Problematik verschwindet in manchen Fällen ohne Therapie.
   Alternativfrage offen: Welche Aussage lässt sich über die Stabilität der Entwicklung einer LRS tätigen, insbesondere in Bezug auf eine Therapie?

2. Worauf zielen Präventionsmaßnahmen einer LRS ab?
   a. **Phonologische Fähigkeiten**
   b. Semantische Fähigkeiten
   c. **Buchstabenwissen**
   d. Wortwissen

3. Was unterscheidet die direkte von der indirekten Route im Dual Route Cascaded Model?
   a. **Direkte Route lässt sich auch als lexikalische Route beschreiben.**
   b. Indirekte Route lässt sich auch als lexikalische Route bezeichnen.
   c. Bei der indirekten Route werden die Worte direkt am Erscheinungsbild erkannt und mit einem im Kopf vorhandenen Lexikon abgeglichen.
   d. Bei der direkten Route werden Worte nicht als Ganzes, sondern nur in Einzelteilen ermittelt und dann zusammengefügt.

4. Welche kognitiven Fähigkeiten haben einen Einfluss auf das Leseverstehen? (Nennen Sie mind. 3)
   a. **Hintergrundwissen**
   b. **Schlussfolgerndes Denken**
   c. Aufmerksamkeit
   d. Wahrnehmung
      Weitere Fähigkeiten: Strategienutzung, Leseflüssigkeit, Wortschatz

5. Was sind Ausschlusskriterien in der Diagnostik einer LRS? (nennen Sie 3)
   a. Motorische Probleme
   b. **Sensorische Behinderung**
   c. Normale Beschulung
   d. **Unterdurchschnittliche Intelligenz (IQ < 70)**
      Weitere Kriterien: mangelnde Beschulung, mangelnde Sprachbeherrschung, psychosozial widrige Umstände

6. Welche Ursachen kann eine LRS haben? (Nennen Sie 2)
   a. Intelligenzminderung
   b. **Genetische Ursachen**
   c. Schlechte Beschulung
   d. **Neuronale Ursachen**
      Weitere Ursachen: Sprachentwicklungsverzögerung

7. Welche Aussagen über die Sprachentwicklung sind wahr?
   a. Kinder sollten mit 1 Jahr bereits 50 Worte sprechen können.
   b. **Kinder, die unter die „Late talkers" fallen, haben eine 50% Wahrscheinlichkeit, noch zu einer normalen Sprachentwicklung zu gelangen.**
   c. **Kinder, die mit zwei Jahren weniger als 50 Worte sprechen, werden „Late Talkers" genannt.**
   d. Eine Legasthenie kann auch ohne verzögerte Sprachentwicklung im Kleinkindalter auftreten.

   Alternativfrage offen: Beschreiben Sie kurz den alterstechnischen Verlauf einer Sprachentwicklungsverzögerung. Nutzen Sie dabei Fachbegriffe!

8. Wie unterscheiden sich die Schwierigkeiten von Tom (schreibt Hunt statt Hund, Kwelle statt Quelle, länge statt Länge, Bal statt Ball, Schpiegel statt Spiegel) und Tim (schreibt Geraffe statt Giraffem Beg statt Berg)? Beschreiben Sie in max. 3 kurzen Sätzen!
   a. Tom hat größere Probleme im Rechtschreiben als Tim
   b. Tom steht auf der logografischen Stufe, Tim dagegen auf der orthografischen

    c. **Tom hat Schwierigkeiten mit der orthografischen Strategie, Tim mit der alphabetischen**

    d. **Tom ist im Schriftspracherwerb eine Stufe weiter als Tim**

9. Warum sollte man die Rechtschreibfehler (zumindest die von Fachbegriffen unabhängigen Rechtschreibfehler) in einer Biologie-Probearbeit nicht bei der Notengebung berücksichtigen?

    a. Die Frage ist schlecht formuliert – sollte man nur bei vorliegendem Attest nicht berücksichtigen.

    b. Weil SchülerInnen nichts für die Rechtschreibfehler können

    c. **Weil in den Lehrplänen für Biologie die allgemeine Rechtschreibung nicht als zentral angesehen wird.**

    d. **Weil es die Validität der Messung beeinträchtigen würde**

10. Wenn Tino folgende Worte folgendermaßen schreibt - zwerk, rante, fersperte, weg – spricht man von einem…

    a. Phonemfehler

    b. **Regelfehler**

    c. Speicherfehler

    d. Restfehler

Fragen zu Dyskalkulie

1. Welche Kompetenzen beinhaltet das Mengenverständnis?

    a. **Mächtigkeit vergleichen**

    b. **Änderung der Mächtigkeit vergleichen**

    c. **Teil-Ganzes-Schema**

    d. Zählen in der richtigen Reihenfolge

2. Welche Eigenschaften von Zahlen kennen Kinder auf der letzten Entwicklungsstufe des Zahlwissens und wie heißt diese Stufe?

    a. **Echtes nummerisches Zählen**

    b. Entwicklung eines Teil-Ganzes-Schemas

    c. Noch kein kardinales Verständnis

    d. **Jedes Wort einer Zahlenreihe = eigene Einheit, eigenen Platz in der Reihe, alle vorausgegangenen Wörter umfasst, Menge steigert sich mit jedem Wort**

3. Mit welchen Fähigkeiten haben Kinder, die eine Dyskalkulie haben, Probleme? Nennen Sie mind. 3

    a. **Schwierigkeiten beim Umkodieren von Zahlen (z.B. eins = 1)**

    b. **Erwerb mathematischen Faktenwissens**

    c. Mangelnde Sprachbeherrschung

    d. **Mathematisches Problemlösen**

       Weitere: Flüssiges Rechnen, Genaues Rechnen, Number sense = Verständnis der Zahl

4. Wie unterscheidet sich die isolierte Rechenschwäche vom kombinierten Auftreten mit LRS?

    a. **Bei isolierter Rechenschwäche Fokus auf Zahlensemantik**

    b. Bei isolierter Rechenschwäche Sprache, Symbolcharakter, Arbeitsgedächtnis betroffen

    c. Bei Kombination Fokus auf Zahlensemantik

    d. **Bei Kombination: Sprache, Symbolcharakter, Arbeitsgedächtnis betroffen**

5. Welche Subtypen von Rechenstörungen lassen sich aus dem Triple Code Modell ableiten?
   a. **Sprachlicher Subtyp**
   b. **Arabischer Subtyp**
   c. Symbolischer Subtyp
   d. **Tiefgreifender Subtyp**

6. Welche mathematischen Probleme sind typisch bei Rechenstörungen? Nennen Sie 2.
   a. **Probleme beim Regelerwerb**
   b. **Probleme beim Speichern mathematischen Faktenwissens**
   c. Probleme beim Aussprechen von Ziffern
   d. **Probleme beim Verständnis von Rechenoperatoren**
      Weitere:
      - Probleme beim Erwerb des hindu-arabischen Stellenwertsystems & seiner syntaktischen Regeln (Einer, Zehner, Hunderter...; Zehner-übergang...) sowie der hierauf aufbauenden Rechenprozeduren
      - Zahlensemantik:
         i. Mangelndes Verständnis für Rechenoperationen und den ihnen zugrundeliegenden Konzepten (z.B. mehr-weniger, Vielfaches, Teil-Ganzes)
         ii. Probleme bei der Erfassung und dem Vergleich der Größe von Mengen
         iii. Der Aufbau gegliederter Zahlenstrahl-oder Zahlenraumvorstellungen und damit die Fähigkeit des Überschlagens und Schätzens von Mengen und Rechenergebnissen ist erschwert.
         iv. Probleme bei der sprachlichen Zahlenverarbeitung und dem Erwerb der Zahlwortsequenz
      - Speichern von Faktenwissen („Einmaleins") erschwert
      - Schwierigkeiten beim Übertragen von Zahlen aus einer Kodierung in eine andere (Zahlwort -arabische Ziffer -analoge Mengenrepräsentation)

7. Welche Folgeprobleme können aufgrund einer Rechenstörung auftreten? Nennen Sie 3 Beispiele!
   a. **Probleme in anderen Fächern, die math. Sachverhalte aufgreifen**
   b. **Probleme mit Geld**
   c. **Probleme mit Terminen, Jahresverlauf, Tagen, Wochen, Zeit…**
   d. **Probleme mit Landkarten und Tabellen**
   e. **Probleme mit geschichtlichen Ereignissen**

8. Was können Ursachen für die Entwicklung einer Rechenschwäche sein?
   a. **Familiäre Häufung**
   b. **Angeborene Fähigkeiten – genetische Ursachen**
   c. Mangelnde Intelligenz
   d. Mangelnde Förderung
      Weitere:
      - Defizite im vorschulisch erworbenen Mengen- & Zahlwissen
      - Mathematikdidaktik in der Schule

9. Wie wird, im Vergleich zu LRS, Dyskalkulie in der Schule beachtet?
   a. **Für LRS kann ein Notenschutz beantragt werden, für Dyskalkulie nicht**
   b. Für Dyskalkulie kann, wie für LRS, ein Nachteilsausgleich eingetragen werden
   c. Rechenfähigkeiten werden nur im Mathematikunterricht bewertet, in den anderen Fächern nicht
   d. LRS und Dyskalkulie werden genau gleichbehandelt

10. Was kann präventiv gegen eine Rechenstörung getan werden?
    a. Mit Kindern Rechenaufgaben leicht über ihrem Niveau lösen, um so zu fordern
    b. **Mit Kindern Zählspiele spielen oder Alltagsgegenstände abzählen**
    c. **Mit Kindern das Mengenverständnis im Alltag trainieren**
    d. Nichts. Eine Rechenstörung prägt sich aus, wenn sie genetisch veranlagt ist.
       Weitere:
       - Schulung und Verknüpfung der Basisfertigkeiten Mengenbegriff, Zählfertigkeiten und Zahlenkenntnis
       - Bewusstheit, dass hinter Zahlen Anzahlen (Mengen) stehen (-> Zunahme um Eins)
       - Bewusstheit, dass von einer zur nächsten Zahl immer Eins dazukommt und dass sich Anzahlen in kleinere Anzahlen zerlegen lassen (Teil-Ganzes-Schema)

Fragen zu ADHS

1. Welche Subtypen von ADHS gibt es? (offen oder MC)
   a. **Vorwiegend hyperaktiv-impulsiv**
   b. **Vorwiegend unaufmerksam**
   c. Vorwiegend hyperaktiv-unaufmerksam
   d. **Kombination**

2. Wie lässt sich der Verlauf von ADHS grob beschreiben? (offen (2 Sätze) oder MC)
   a. Aktivität nimmt bis zum Erwachsenenalter zu, Aufmerksamkeitsstörung persistiert
   b. **Physische Aktivität bleibt bis zum Jugendalter hoch, danach abnehmend**
   c. Aggressivität entwickelt sich erst im Jugendalter
   d. Alle Symptome bessern sich im Erwachsenenalter

3. Welche Faktoren gelten als Ursachen für ADHS? (offen oder MC)
   a. **Alkohol und Nikotin in der Schwangerschaft**
   b. Hoher Sozialstatus der Eltern
   c. **Ungünstige Bedingungen in der Schule**
   d. Wenig Bewegung im Kleinkindalter

4. Welche Kriterien nach ICD 11 müssen für eine ADHS erfüllt sein? Nennen Sie 3
   a. **Situationsunabhängigkeit des Verhaltens**
   b. IQ durchschnittlich *(aber: IQ nicht lernbehindert wäre richtig)*
   c. **Symptome seit mind. 12 Monaten**
   d. Störung bereits vor 4. Lebensjahr begonnen

5. Welche Möglichkeiten bestehen als Lehrkraft, mit Kindern mit ADHS umzugehen? Nennen Sie 3!
   a. **Für ruhige Umgebung mit möglichst wenig Ablenkungspotential sorgen**
   b. Drastische Strafen, wenn Regeln verletzt werden
   c. **Häufiger Pausen einlegen**
   d. Möglichst viele verschiedene Medien nutzen
      Weitere Möglichkeiten:
      - Für eine ruhige Umgebung mit möglichst wenig Ablenkungspotential sorgen (ruhige, reizarme Räumlichkeiten, usw.)
      - Eine inhaltlich geeignete Gliederung und Gestaltung der Lern-und Spielmaterialien schaffen
      - Strukturierung unter Zuhilfenahme von Routinen und Ritualen (z.B. jede Stunde mit ein paar Minuten Entspannung beginnen)

- Häufiger Pausen einlegen (z.B. 15 min arbeiten, dann 5 min Pause -> Konzentrationsdauer nicht überlasten)
- Klare Vereinbarungen mit dem Kind treffen und ggf. gemeinsam mit dem Kind (!) Regeln aufstellen
- Auf strikte Einhaltung der Regeln achten
- Häufiges Loben der Kinder, anstatt sie auf Defizite hinzuweisen; besonders bei hyperaktiven Kindern ist die Verstärkung für positives Verhalten wichtig (in Form von Lob, einem Lächeln, etc.)

6. An welchen Punkten/Aspekten kann eine Intervention bei ADHS ansetzen? Nennen Sie 3
   a. **Gestaltung der Lernumwelt**
   b. Veränderte Bewertungen im Schulkontext
   c. Motorische Trainings
   d. **Erziehungsberatung**
   Weitere: Pharmakotherapie, Selbstregulationskompetenzen des Kindes

7. Welche Störungen können mit ADHS leicht verwechselt werden? Nennen Sie 3 Beispiele
   a. **Störungen des Sozialverhaltens**
   b. Beziehungsstörung mit Hemmung
   c. **Persönlichkeitsstörungen**
   d. Hormonstörungen
   Weitere: Störungen des Sozialverhaltens und oppositionelles Trotzverhalten, stereotype Bewegungsstörungen, Autismus-Spektrum-Störungen, Tic-und Tourette-Störungen, umschriebene Entwicklungsstörungen und Lernstörungen, Intelligenzminderung & Überforderung, Autismus-Spektrum-Störungen, Beziehungs-/ Bindungsstörung mit Enthemmung, Angststörungen, depressive Störungen & bipolare Störungen, disruptive Affektregulationsstörungen, Substanzkonsumstörungen, Persönlichkeitsstörungen, psychotische Störungen, Medikamenteninduzierte Störungen (z. B. Bronchospasmolytika), Müdigkeit und Unaufmerksamkeit bei Schlafstörungen, Hyperarousal bei Posttraumatischer Belastungsstörung

8. Was können Folgen einer ADHS-Symptomatik sein? Nennen Sie 3 Beispiele!
   a. Erhöhte Frustrationsintoleranz
   b. **Motivationsschwäche**
   c. Mangelnde Verzögerungsaversion
   d. **Erhöhte Unfallgefahr**
   Weitere: niedrige Frustrationstoleranz, Soziale Konflikte, Bullying, Leistungsschwäche in Schule und Beruf, Substanzmissbrauch, Angst, Psychische Folgeerkrankungen, Teilleistungsstörungen, Affektive Störungen

9. Wessen Urteil ist am zuverlässigsten bei der Erkennung einer ADHS?
   a. Urteil der Eltern
   b. Urteil der Freunde
   c. Urteil des Kindes
   d. **Urteil der Lehrkräfte**

10. Welche AHDS-typischen Merkmale weisen betroffene Kinder im Grundschulalter auf? Nennen Sie 2!
   a. Unruhe / Ablenkbarkeit im Unterricht
   b. Lernschwierigkeiten / Teilleistungsschwächen
   c. Umschulungen / Klassenwiederholungen
   d. aggressives Verhalten
   e. Ablehnung durch Gleichaltrige
   f. Leistungsunsicherheit / Selbstwertprobleme

Fragen zu aggressivem Verhalten

1. Wie unterscheiden sich die Begriffe Aggressives Verhalten, Aggression und Aggressivität?
   a. Aggression ist die Bereitschaft zu aggressiven Handlungen.
   b. Aggressivität ist die Bereitschaft zu aggressiven Handlungen.
   c. Aggressives Verhalten zielt auf die Schädigung von Personen, Tieren oder Objekten.
   d. Aggressivität mündet immer in Aggression.

2. Welche Beschreibungen treffen auf das folgende aggressive Verhalten zu?
   Jonas schlägt Mia, um ihre Hausaufgaben zu bekommen.
   a. Das Verhalten ist instrumentell
   b. Das Verhalten ist reaktiv
   c. Das Verhalten ist feindselig
   d. Das Verhalten ist offen

3. Gibt es Effekte von Gewalt in Computerspielen und falls ja, was könnte der Haupteinflussfaktor sein?
   a. Ja, es sind kleine bis mittlere Effekte sichtbar
   b. Nein, es gibt keine Effekte
   c. Der Haupteinflussfaktor könnte die Sozialisation sein.
   d. Der Haupteinflussfaktor könnte die Selektion sein.

4. Was können Gründe für aggressives Verhalten sein? Nennen Sie 3!
   a. Reaktion auf vermeintliche Bedrohung
   b. Durchsetzen eigener Interessen
   c. Ungeschickte Kontaktaufnahme
   d. Hilflosigkeit
   - Antwort auf Provokation
   - Identität oder Selbstbewusstsein gewinnen
   - Appell für mehr Zuwendung

5. Wie ist stehen die Störung des Sozialverhaltens, die Störung mit oppositionellem Trotzverhalten und die intermittierende explosible Störung in Beziehung zueinander?
   a. Es lässt sich kein Zusammenhang zwischen den Störungen feststellen.
   b. Die Störung des Sozialverhaltens kann als Mischform aus Störung mit oppositionellem Trotzverhalten und intermittierender explosibler Störung gesehen werden.
   c. Die Störung mit oppositionellem Trotzverhalten kann als Mischform aus Störung des Sozialverhaltens und intermittierender explosibler Störung gesehen werden.
   d. Die intermittierende explosible Störung kann als Mischform aus Störung mit oppositionellem Trotzverhalten und Störung des Sozialverhaltens gesehen werden.

6. Emils Verhalten im Kindergarten ist seit über einem halben Jahr auffällig. Er ist oft beleidigt, reagiert auf Fehler anderer nachtragend und schiebt seine eigenen Fehler auf andere. Welcher Störung würden Sie das zuordnen?
    a. **Störung mit oppositionellem Trotzverhalten**
    b. Störung des Sozialverhaltens
    c. Intermittend explosive disorder
    d. Disruptive mood dysregulation disorder

7. Was sind Einflussfaktoren auf die Ausübung von Aggression?
    a. Alter: je jünger, desto mehr Aggression
    b. **Umfeld: aggressive Rollenbilder in Familien- & Freundeskreis**
    c. **Psychische Auffälligkeiten, z.B. ADHS**
    d. **Persönliche Gewalterfahrung**

8. Wie können Depressionen und aggressives Verhalten im Kindesalter zusammenhängen?
    a. **Eine Depression kann die Folge aggressiven Verhaltens sein.**
    b. **Aggressives Verhalten kann als Ausdruck einer depressiven Störung auftreten.**
    c. Eine Depression verdrängt aggressive Verhaltensweisen und führt zu Lethargie.
    d. Jedes Kind mit Depressionen zeigt aggressive Züge.

9. Wie unterscheiden sich früh vs. spät begonnene Störungen in ihrem Verlauf?
    a. **Early onset: massivere Probleme im späteren Leben, kaum Besserung des Verhaltens**
    b. **Late onset: Therapie aussichtsreich, Verhalten normalisiert sich bis Erwachsenenalter**
    c. Early onset: Therapie aussichtsreich, da früh entdeckt; Verhalten normalisiert sich bis Erwachsenenalter
    d. Late onset: massivere Probleme im späteren Leben, da spät entdeckt; kaum Besserung des Verhaltens

10. Was sind Ziele von Interventionsmaßnahmen bei aggressiven Störungen?
    a. Symptome direkt unterbinden
    b. **Zunahme prosozialen Verhaltens**
    c. **Empathiefähigkeit steigern**
    d. **Bindung zu Eltern verbessern**

Fragen zu schulischer Gewalt

1. Welche Aspekte treffen auf den Begriff Mobbing/Bullying zu?
    a. **Über einen längeren Zeitraum**
    b. Mehrmals oder auch nur einmalig
    c. **Ungleichgewicht der Kräfte**
    d. Gibt es nur im Jugendalter

2. Wie unterscheidet sich das Gewaltverhalten von Jungen und Mädchen typischerweise?
    a. **Jungen setzen sich eher physisch auseinander**
    b. Es gibt kaum Geschlechterunterschiede
    c. **Bei Mädchen überwiegen indirekte Methoden**
    d. **Dass Jungen Mädchen mobben und umgekehrt ist eher selten**

3. Welche situationalen Faktoren können zu Gewalt in der Schule beitragen? Nennen Sie 2 Beispiele!
    a. Strafende Lehrkräfte
    b. **Kompetitiver Unterricht**

  c. **Verstärkung der Bullies durch Nicht-Reagieren der Lehrkräfte**
  d. **Inkonsequentes Handeln der Lehrkräfte**

4. Wie beschreibt das Bedingungsmodell des Bullying, weshalb ein Täter immer wieder zuschlägt?
  a. Erstes Bullying wird meist vom Opfer noch abgewehrt.
  b. **Opfer reagiert ängstlich oder passiv**
  c. Mitschüler schenken meist keine Beachtung
  d. **Bully fühlt sich bestärkt, da er nicht zurückgehalten wird**

5. Was können typische psychische Folgen des Bullying für Opfer im Erwachsenenalter sein?
  a. **Angststörungen**
  b. Antisoziale Persönlichkeitsstörung
  c. Suizid
  d. Depression

6. Was können typische psychische Folgen des Bullying für Täter im Erwachsenenalter sein?
  a. Angststörungen
  b. **Antisoziale Persönlichkeitsstörung**
  c. Suizid
  d. Depression

7. Was können typische psychische Folgen des Bullying für Täter-Opfer im Erwachsenenalter sein?
  a. **Angststörungen**
  b. Antisoziale Persönlichkeitsstörung
  c. **Suizid**
  d. **Depression**

8. Welche Merkmale weist Cybermobbing auf?
  a. Häufig aus Spaß, daher an zufällig erwählten Personen im Netz
  b. **Besonders belastend, da schnell öffentlich verbreitet**
  c. Opfer kennen Täter meist nicht
  d. **Zeigt sich meist in Textnachrichten oder Fake-Profilen**

9. Welche Ziele sollten Interventionen gegen Gewalt in der Schule haben? Nennen Sie mind. 3!
  a. **Reduzierung mittelbarer und unmittelbarer Gewalt**
  b. **Verbesserung der Beziehungen unter den Schülerinnen und Schülern**
  c. **Bedingungen schaffen, die sowohl Opfern als auch Tätern ein besseres Auskommen miteinander, innerhalb und außerhalb der Schule, ermöglichen**
  d. **Förderung und Erweiterung der sozialen Kompetenz**
  e. **Entwicklung des Schulklimas**

10. Auf welchen Ebenen sollten Interventionen ansetzen?
  a. **Individuum**
  b. **Klasse**
  c. **Schule**
  d. Eltern

Fragen zu Angst

1. Wie kann man Angst von Furcht unterscheiden?
  a. Angst hat einen Bezug auf eine spezifische Gefahr

b. **Furcht hat einen Bezug auf eine spezifische Gefahr**

c. **Angst blockiert, Furcht führt zu Angriff oder Flucht**

d. Furcht blockiert, Angst führt zu Angriff oder Flucht

2. Auf welchen Ebenen zeigt sich Angst? Beschreiben Sie 2 davon genauer!

a. **Physiologie: Herzklopfen, erhöhter Puls & Blutdruck, beschleunigte Atmung, Harndrang, Schweißausbruch, Muskeltonus**

b. **Kognition: Bewertung einer Situation als bedrohlich**

c. **Affekt: Unwohlsein, Gefühl von Bedrohung, depressive Verstimmung, Angespanntheit**

d. **Verhalten: Unruhe, Zittern, unkontrollierte Bewegungen, Flucht, Aggression, Artikulationsstörung**

3. Wann ist eine Angst klinisch bedeutsam? Nennen Sie 3 Kriterien!

a. **Passt nicht zur aktuellen Entwicklungsphase**

b. Bereits kurzfristige Ängste können klinisch bedeutsam sein

c. Folgen auf Individuum beschränkt

d. **Schränkt die Entwicklung ein**

4. Was sind die häufigsten Komorbiditäten von Angststörungen? Nennen Sie 2!

a. **Depressive Störungen**

b. **Substanzmissbrauch**

c. Lernstörungen

d. ADHS

5. Beschreiben Sie 2 Ansatzpunkte, um Ihren Unterricht so zu gestalten, dass es möglichst wenige angstauslösende Faktoren gibt.

a. **Wettbewerbe im Unterricht vermeiden**

b. Viele Gruppenarbeiten

c. **Offen über Leistungserwartungen sprechen**

d. **Gemischte Unterrichtsform**

6. Wie erhält sich laut Behavioristen die Angst aufrecht?

a. Angst erhält sich aufrecht, indem sie immer wieder ausgelöst und bestätigt wird.

b. **Angst erhält sich aufrecht, indem sie vermieden und dadurch verringert wird.**

c. Angst kann gelöscht werden, indem sie bewusst herbeigeführt wird.

d. **Angst kann gelöscht werden, indem falsche Erwartungen gelöscht werden.**

7. Wie wird mittels der Methode der systematischen Desensibilisierung versucht, Angst abzubauen?

a. **Klassische Methode der Angstbehandlung im Kindesalter**

b. Methode, die erst im Erwachsenenalter angewendet werden kann

c. **Konfrontation mit Angstreizen erfolgt nach vorher festgelegter Angsthierarchie**

d. Es muss in der Realität stattfinden, reine Vorstellung wirkungslos

8. Wie wird mittels der Methode der Reizkonfrontation versucht, Angst abzubauen?

a. **Massive Konfrontation mit angstauslösendem Reiz**

b. **Keine Möglichkeit zur Flucht**

c. Zeitgleich aber Anwendung von Entspannungsmaßnahmen

d. **Angst geht von selbst zurück und wird so gelöscht**

9. Anhand welcher Symptome zeigt sich Leistungsangst deutlich? Nennen Sie 2
   a. Attributionsstil bei Misserfolg vorwiegend internal und variabel
   b. **Antizipieren Versagen, für das sie sich selbst verantwortlich machen**
   c. **Selbstabwertende Kommentare nach Versagen**
   d. Attributionsstil bei Erfolg internal und variabel

10. Welche Ansatzpunkte gibt es, um soziale Ängste zu mindern?
    a. **Aufbau von Selbstvertrauen**
    b. **Rollenspiele zu angstauslösenden Situationen**
    c. **Techniken zur Selbstinstruktion in Angstsituationen**
    d. **Beratung von Bezugspersonen**

Fragen zu Denken

1. Welche theoretischen Ansätze zur Beschreibung und Erklärung der (kindlichen) Entwicklung des Denkens und Problemlösens gibt es?
   a. **Konstruktivistischer Ansatz (Piaget)**
   b. **Informationsverarbeitungstheorie**
   c. soziokonstruktivistische Theorien
   d. Theorien systematischer Dynamiken

2. Beschreiben Sie die drei kontinuierlichen Prozesse der Theorie von Piaget in jeweils einem Satz! // Wie lassen sich die drei kontinuierlichen Prozesse der Theorie von Piaget beschreiben?
   a. Assimiliation bedeutet, dass vorhandene Wissensstrukturen in Reaktion auf neue Erfahrungen angepasst werden.
   b. Akkomodation bedeutet, dass neue Informationen in bereits bestehende Konzepte integriert werden.
   c. **Assimilation bedeutet, dass eintreffende Informationen in eine Form überführt, die man versteht.**
   d. **Äquilibration ist das Ausbalancieren von Assimilation und Akkommodation, um ein stabiles Verstehen zu ermöglichen.**

3. Welchem Stadium der Entwicklung nach Piaget würden Sie die folgende Aussage zuordnen und warum? Erklären Sie in 1-2 Sätzen!
   Es gibt 3 Kekse. Tom nimmt sich 2. Für Hannes bleibt noch einer übrig, den er in der Mitte durchteilt. Nach dem Verständnis von Hannes und Tom ist das eine faire Aufteilung.
   a. Sensomotorisches Stadium
   b. **Präoperationales Stadium**
   c. Konkret-operationales Stadium
   d. Formal-operationales Stadium

4. Welche praktischen Implikationen ergeben sich aus Piagets Theorie für die pädagogische Praxis? Nennen Sie 3!
   a. Man sollten Kindern Inhaltswissen vermitteln, um den Problemlöseprozess zu unterstützen.
   b. **Man sollten Kindern Denkanstöße geben und Widersprüche aufbauen.**
   c. **Man sollte nach dem kindzentrierten Ansatz handeln, d.h. die Art des Denkens der Kinder in pädagogischen Kontexten berücksichtigen.**
   d. Man sollte mit Kindern üben, Strategien anzuwenden, v.a. in unterschiedlichen Kontexten.

5. Inwiefern lässt sich ein Kind laut Theorie der Informationsverarbeitung mit einem Computer vergleichen? Beschreiben Sie in 1-2 Sätzen!
   a. **Die „Hardware" der Kinder sind die kognitiven Strukturen.**
   b. Gedächtniskapazität und Leistungsfähigkeit werden durch die „Software" vermittelt.
   c. Die „Software" der Kinder umfasst z. B. das Arbeitsgedächtnis und das Kurzzeitgedächtnis.
   d. **Regeln und Strategien des Gedächtnisses und Problemlösens können als „Software" bezeichnet werden.**

6. Welche Funktion erfüllt das sensorische Gedächtnis? Beschreiben Sie in 1 Satz.
   a. **Kurze, flüchtige Sinneseindrücke wahrnehmen.**
   b. Kurze Sinneseindrücke durch Wiederholung im Gedächtnis halten.
   c. Kurze Sinneseindrücke wahrnehmungsgetreu im Langzeitgedächtnis abspeichern.
   d. Nur die wichtigen Sinneseindrücke aussortieren und abspeichern.

7. Welche praktischen Implikationen ergeben sich aus Informationsverarbeitungstheorien für die pädagogische Praxis? Nennen Sie 3!
   a. **Man sollte Kindern Übungsgelegenheiten schaffen.**
   b. **Strategien sollten in immer unähnlicher werdenden Kontexten geübt werden.**
   c. Es sollten Lerngemeinschaften eingesetzt werden.
   d. Man sollten Kindern Denkanstöße geben und Widersprüche aufbauen.

8. Wie beschreiben soziokulturelle Theorien die Entwicklung des Denkens? Beschreiben Sie stichpunktartig!
   a. **Lernen & Entwicklung findet im zwischenmenschlichen & kulturellen Kontext statt.**
   b. Entwicklung ist stufenartig durch qualitative Veränderungen gekennzeichnet.
   c. Kulturwerkzeuge, die zum Lernen notwendig sind, unterscheiden sich grundsätzlich kaum in verschiedenen Kulturen.
   d. Erwachsene sollten Kindern nur einen Anstoß geben, sie den restlichen Prozess aber allein durchlaufen lassen.

9. Welche praktischen Implikationen ergeben sich aus soziokulturellen Theorien für die pädagogische Praxis?
   a. **Die Lehrpersonen sowie die individuelle Anpassung von Lehre an die Kinder sind zentral.**
   b. Strategien sollten in immer unähnlicher werdenden Kontexten geübt werden.
   c. **Kindern sollten möglichst früh interaktive Erfahrungskontexte geboten werden.**
   d. **Der Einsatz von Lerngruppen im Unterricht ist wichtig.**

10. Wie wird Entwicklung im Rahmen der Theorien dynamischer Systeme definiert?
    a. **Entwicklung als sich selbst organisierender Prozess**
    b. **Integration verschiedener Komponenten zur Anpassung an Umwelt**
    c. Kognitive Entwicklung nur aus sich heraus, ohne Einfluss anderer Menschen
    d. Denken und Handeln als zwei unabhängige Prozesse

Fragen zu Intelligenz

1. Welche Fähigkeiten beschreibt die Intelligenz?
   a. **Anpassungsfähigkeit an neue Gegebenheiten**
   b. Sich stark verändernde kognitive Leistung
   c. **Flexibles Reagieren auf herausfordernde Situationen**
   d. Schnell Handlungen automatisieren können

2. Was bringt die Intelligenzmessung im schulischen Kontext? Nennen Sie 3 Punkte!
    a. **Diagnose von Hochbegabung**
    b. **Diagnose von Minderbegabung**
    c. **Vorhersage beruflicher Leistungen**
    d. **Vorhersage Schulischer Leistungen**

3. Welche zwei Faktoren nimmt Cattell in seinem Modell an? Beschreiben Sie in wenigen Stichpunkten die Hauptmerkmale dieser Faktoren!
    a. **Fluide und kristalline Intelligenz**
        - Fluide Intelligenz
            i.   non-verbal
            ii. (angeboren)
            iii. neuen Situationen anpassen
            iv. induktives Denken
        - Kristalline Intelligenz
            i.   Verbal
            ii. erworben
            iii. Faktenwissen über die Welt
            iv. Bereichsspezifisch
    b. Verbale und nonverbale Intelligenz
    c. Logisches Denken und Informationsverarbeitung
    d. Deklarative und non-deklarative Intelligenz

4. Friedrich ist ein aufgewecktes Kind und seinen Klassenkameraden in vielen Leistungsbereichen voraus. Auf Anraten der Lehrerin lassen seine Eltern ihn bei einer Psychologin testen. Es wird ein IQ von 120 ermittelt. Wie ist dieser Wert einzuschätzen?
    a. **Friedrich hat einen überdurchschnittlichen Intelligenzquotienten.**
    b. Friedrich ist hochbegabt.
    c. **Der IQ-Wert bedeutet, dass Friedrich im Vergleich zu seiner Altersgruppe überdurchschnittlich intelligent ist.**
    d. **Der IQ-Wert könnte durch situationale Faktoren beeinflusst sein und der eigentliche Wert etwas höher/niedriger liegen.**

5. Wie entwickelt sich die Intelligenz über die Lebensspanne? Beschreiben Sie in 1-2 Sätzen!
    a. **Fluide Intelligenz nimmt bis 25. Lebensjahr zu -> Maximum -> danach Abnahme**
    b. **Kristalline Intelligenz nimmt bis zum Erwachsenenalter zu und bleibt dann weitestgehend stabil**
    c. Kristalline Intelligenz nimmt bis 25. Lebensjahr zu -> Maximum -> danach Abnahme
    d. Fluide Intelligenz nimmt bis zum Erwachsenenalter zu und bleibt dann weitestgehend stabil

6. Wie wirkt sich der pädagogische Kontext auf die Entwicklung der (non-)verbalen Intelligenz aus (Schneider & Stefanek, 2004)?
    a. **Früh auftretende Unterschiede in der Entwicklung bleiben bestehen.**
    b. Früh auftretende Unterschiede in der Entwicklung können angeglichen werden.
    c. **Unterschiede in der verbalen Intelligenz werden oft durch unterschiedliche Schulformen verstärkt.**
    d. Schüler verschiedener Schulformen unterscheiden sich kaum hinsichtlich ihrer verbalen Intelligenz.

7. Welche Kritik an der Intelligenzmessung, wie sie heute betrieben wird, könnte man äußern? Nennen Sie 2 Kritikpunkte.
   a. **Die meisten Intelligenztests erfassen die zugrundeliegenden Fähigkeiten nicht umfassend genug**
   b. Konzentrieren sich zu wenig auf schulisch relevante Fähigkeiten
   c. **Viele Tests sind kulturell verzerrt**
   d. Intelligenztests nicht überall anerkannt
   Weitere: Ethisch fragwürdig -> Reduktion so komplexen Konstrukts auf eine Zahl (IQ); Verwendung von Intelligenztests kann weitreichende und mitunter unerwünschte Konsequenzen haben -> IQ als willkürliches Cut-Off-Kriterium

8. Bei welchen Intelligenzbereichen konnten Unterschiede sich zwischen Hauptschülern und Gymnasiasten beobachtet werden?
   a. Verbale Intelligenz
   b. Nonverbale Intelligenz
   c. Allgemeine Intelligenz
   d. **Alle drei**

9. Welche Merkmale weisen kreative Aufgaben nach Guilford auf? Nennen Sie 2
   a. **Ziel des Problemlöseprozesses nicht genau vorgegeben**
   b. Eingeschränkte Zahl an Lösungsmöglichkeiten -> Gegenteil wäre richtig
   c. **Neuartige Kombination vorhandenen Wissens gefordert**
   d. Werden hauptsächlich durch kristalline Intelligenz gelöst

10. Was versteht man unter passiver Genom-Umwelt-Korrelation?
    a. Menschen wählen sich Umwelten, die zu ihren genetisch beeinflussten Merkmalen passen
    b. Soziale Umwelt reagiert auf genetisch bedingte Merkmale
    c. **Biologische Eltern schaffen Umwelt, die zu ihrer eigenen genetischen Disposition passt und somit vermutlich auch zu der ihrer Kinder**
    d. Soziale Umwelt versucht ohne Rücksicht auf genetisch bedingte Merkmale zu handeln

Fragen zu Gedächtnis I

1. Welche Komponenten enthält das Drei-Speicher-Modell (Mehrspeichermodell) des Gedächtnisses nach Atkinson & Schiffrin?
   a. **Sensorischer Speicher**
   b. **Arbeitsgedächtnis**
   c. Ultralangzeitgedächtnis
   d. **Langzeitgedächtnis**

2. Welche Funktionen erfüllt das Arbeitsgedächtnis?
   a. Verarbeiten aller Sinneseindrücke
   b. **Serielle Verarbeitung kleiner Informationsmengen**
   c. Speicherung wichtiger Informationen für später
   d. **Memorieren und Abruf von Informationen aus dem LZG**

3. Was unterscheidet das episodische vom semantischen Gedächtnis?
   a. **Beim episodischen Gedächtnis wird der Erwerbskontext mitgespeichert.**
   b. Beim semantischen Gedächtnis wird der Erwerbskontext mitgespeichert.
   c. **Informationen durchlaufen meist erst das episodische Gedächtnis, bevor sie im semantischen gespeichert werden.**
   d. Im semantischen Gedächtnis finden sich persönliche Erinnerungen wieder.

4. Wann werden die frühesten Erinnerungen ins Langzeitgedächtnis gespeichert?
   a. **Bereits vor und kurz nach der Geburt werden erste Eindrücke gespeichert.**
   b. Als Säugling, ca. 2 Wochen nach der Geburt, beginnt das Langzeitgedächtnis, Erinnerungen zu speichern.
   c. Erst im Kleinkindalter ist das Langzeitgedächtnis weit genug entwickelt, um Erinnerungen zu speichern.
   d. Da die Sprachentwicklung erst mit 2 Jahren einsetzt, werden auch hier erst Erinnerungen im Langzeitgedächtnis produziert.

5. Welchen Vorteil bringen die Speicherung bzw. der Abruf von Skripts?
   a. Bringt eigentlich keinen Vorteil, ist aber trotzdem so
   b. **Ermöglichen eine ökonomische Enkodierung.**
   c. **Ermöglichen das Antizipieren und Kontrollieren von Ereignissen.**
   d. **Automatisiert Handlungen.**

6. Was beschreibt der Begriff der Infantilen Amnesie und was sind mögliche Gründe dafür?
   a. **Bis zum Alter von 3 Jahren gibt es kaum autobiografische Erinnerungen.**
   b. Bis zum Alter von 3 Jahren werden überhaupt keine Erinnerungen im Langzeitgedächtnis abgelegt.
   c. Ein Grund für die infantile Amnesie könnte sein, dass sich die Form der Erinnerungen ungefähr in diesem Alter verändert.
   d. **Ein Grund für die infantile Amnesie könnte sein, dass Kinder vor diesem Alter noch kein kognitives Selbst besitzen.**

7. Frieda kann sich nicht daran erinnern, wie sie im Alter von 5 Jahren im Sommer von einem Baum gefallen ist – ihr ist dabei nicht viel passiert. Georg kann sich noch gut daran erinnern, wie er im Alter von 5 Jahren vom Fahrrad gefallen ist und sich dabei die Hand gebrochen hat.
   Warum gibt es diesen Unterschied, obwohl beide Kinder gleich alt waren?
   a. Georg ist intelligenter als Frieda und erinnert sich daher besser.
   b. **Für Frieda war das Ereignis weniger wichtig, da ihr nicht schlimmes passiert ist.**
   c. Beide Kinder waren eigentlich noch zu jung, um Erinnerungen zu bilden, daher ist es seltsam, dass Georg sich erinnert.
   d. **Für Georg war das Ereignis schmerzhafter als für Frieda.**

8. Nennen Sie 3 Möglichkeiten, die explizite Erinnerungsleistung von Kindern zu verbessern!
   a. Man sollte Kindern Suggestivfragen stellen, um sie auf die richtige Spur zu bringen.
   b. Man sollte Kindern nur ja/nein Fragen stellen, um ihnen eine einfache Antwort zu ermöglichen.
   c. **Man sollte mit Kindern regelmäßig in den Dialog über Ereignisse treten, um zu reflektieren.**
   d. **Man sollte eine warme, interpersonelle Atmosphäre im Gespräch schaffen.**

9. Wie entwickelt sich das explizite Gedächtnis über das gesamte Leben hinweg?
   a. Den stärksten Leistungszuwachs gibt es in der späten Jugend.
   b. **Ab dem frühen Erwachsenenalter nimmt die Gedächtnisleistung wieder ab.**
   c. **Von Ende des Kindergartens bis Ende der Grundschule stärkster Leistungszuwachs.**
   d. **Bestimmte Teile des Gedächtnisses entwickeln sich mit dem Alter zurück, andere nicht.**

10. Welche Erklärungen könnte es für die Altersunterschiede der Gedächtnisleistung geben?
    a. **Entwicklung des Kurzzeitgedächtnisses**
    b. **Nutzung von Gedächtnisstrategien**
    c. **Metakognitives Wissen über Gedächtnisprozesse**
    d. **Bereichsspezifisches Vorwissen**

Fragen zu Gedächtnis II

1. Welche biologischen Faktoren könnte es für die Entwicklung des Gedächtnisses mit dem Alter geben?
    a. Es können mehr elektrische Signale produziert werden, was zu einer höheren Leistung führt.
    b. **Die Nervenzellen des Gehirns werden myelinisiert, was zu einer schnelleren Übertragung führt.**
    c. **Gehirnareale werden ausdifferenziert, was zu einer besseren Leistung führt.**
    d. Das Gehirn wächst innerlich noch bis zum Jugendalter, was zu einer Geschwindigkeitszunahme führt.

2. Wie beschreibt die Theorie nach Case die Entwicklung des Kurzzeitgedächtnisses und seiner Komponenten?
    a. Die Gesamtverarbeitungskapazität des Gedächtnisses nimmt mit dem Alter zu.
    b. **Die Gesamtverarbeitungskapazität des Gedächtnisses wird aufgeteilt in den Arbeitsspeicher und den Kurzzeitspeicher.**
    c. Mit steigendem Alter werden Gedächtnisprozesse effizienter, daher steht mehr Arbeitsspeicher zur Verfügung.
    d. **Bei kleineren Kindern sind beide Komponenten stark ausgelastet, weshalb das Denken langsamer ist als bei Älteren.**

3. Welche Beispiele für Wiederholungs-, Organisations- und Elaborationsstrategien gibt es?
    a. **Auswendiglernen ist ein gutes Beispiel für eine Wiederholungsstrategie.**
    b. Eselsbrücken sind ein gutes Beispiel für Wiederholungsstrategien.
    c. Sich Gliederungen erstellen oder Concept-Maps zeichnen ist ein gutes Beispiel für Elaborationsstrategien.
    d. **Lückentexte oder Fragen zu bearbeiten ist ein gutes Beispiel für Elaborationsstrategien.**

4. Was ist das sog. Mediationsdefizit, woher kommt es und wen betrifft es?
    a. **Das Mediationsdefizit bedeutet, dass Strategien nicht genutzt werden, selbst wenn sie gezielt beigebracht wurden.**
    b. Das Mediationsdefizit bedeutet, dass Strategien erst benutzt werden, wenn sie automatisiert sind.
    c. **Vermutlich kommt es wegen mangelnder Ausreifung des präfrontalen Kortex zustande.**
    d. Das Mediationsdefizit betrifft vor allem Schulkinder.

5. Wann beherrschen Kinder welche Art des Strategieeinsatzes?
    a. Alle Arten von Strategien werden bereits im frühen Schulalter beherrscht, nur noch nicht angewandt.
    b. **Wiederholungsstrategien werden mit steigendem Alter auch vermehrt spontan genutzt.**
    c. **Organisationsstrategien werden von älteren Schulkindern genutzt, jüngere müssen noch angewiesen werden.**
    d. **Elaborationsstrategien werden erst im frühen Jugendalter erworben.**

6. Welche Vorläuferfähigkeiten metakognitiven Wissens gibt es?
    a. **Theory of Mind**
    b. Intelligenz
    c. Vorwissen
    d. **Basale Sprachfähigkeiten**

7. Was unterscheidet das deklarative vom prozeduralen Metagedächtnis?
    a. **Im deklarativen Metagedächtnis findet sich bewusst abrufbares Wissen über Gedächtnisvorgänge**.
    b. Die Fähigkeit zur Überwachung von Denkprozessen findet sich im deklarativen Metagedächtnis.
    c. **Gedächtnisprozesse werden vom prozeduralen Teil des Metagedächtnisses kontrolliert.**
    d. **Das prozedurale Metagedächtnis enthält non-verbalisierbares Wissen über Gedächtnisvorgänge.**

8. Wie entwickeln sich die metakognitiven Überwachungsprozesse mit dem Alter?
    a. **Kindergartenkinder überschätzen ihre Gedächtnisleistung deutlich.**
    b. Kindergartenkinder unterschätzen ihre Gedächtnisleistung deutlich.
    c. Einschätzungen der eigenen Gedächtnisleistung werden erst im Jugendalter genauer.
    d. **Einschätzungen der eigenen Gedächtnisleistung werden bereits in der Grundschule genauer, verändern sich bis zum Jugendalter noch weiter.**

9. Was wären Beispiele für metakognitive Strategien beim Lernen?
    a. Ein Mindmap erstellen
    b. **Sich eine To-Do-Liste schreiben**
    c. **Die richtigen Rahmenbedingungen schaffen**
    d. Eine inhaltliche Gliederung des Lernstoffs erstellen

10. Gibt es einen Zusammenhang zwischen dem Gedächtnis und dem Metagedächtnis?
    a. **Mit zunehmendem Alter nimmt Zusammenhang zwischen deklarativem Metagedächtnis über Strategiegebrauch und Leistungen in strategischen Gedächtnisaufgaben zu.**
    b. Mit zunehmendem Alter nimmt Zusammenhang zw. Deklarativem Metagedächtnis über Strategiegebrauch und Leistungen in strategischen Gedächtnisaufgaben ab.
    c. Wie stark der Zusammenhang ist, hängt von der Intelligenz der jeweiligen Person ab.
    d. **Wie stark der Zusammenhang ist, hängt vom Alter der jeweiligen Person und der Aufgabenschwierigkeit ab.**

Fragen zur Sprachentwicklung

1. Was sprachbezogenen Fähigkeiten müssen Kinder auf dem Weg zum kompetenten Sprecher erwerben? Nennen Sie 3 Beispiele!
    a. **Syntax**
    b. **Pragmatik**
    c. Prodastik
    d. **Semantik**

2. In welcher zeitlichen Abfolge gehen vorläufige sprachlichen Äußerungen dem ersten Wort voran?
    a. Zuerst lallen und brabbeln Kinder mit ca. 2 Monaten.
    b. **Lachen und Nachahmung von Vokalen setzt ca. im 2.-4. Monat ein.**
    c. **Die Reihenfolge der Vorläufer ist: Gurren, Lachen, Nachahmen, Brabbeln.**
    d. **Mit 10 -14 Monaten sprechen Kinder dann erste Wörter.**

3. Welche verschiedenen Quellen nutzen Kinder, um beim Erwerb von Wortbedeutungen klarzukommen?
   a. **Kinder nutzen den Gesichtsausdruck oder die Blickrichtung des Sprechers.**
   b. **Kinder nutzen die Wortart oder die Syntax des Gehörten.**
   c. **Kinder nutzen bestimmte Faustregeln (constraints) zur Orientierung.**
   d. **Kinder nutzen Fehler als Informationsquelle.**

4. Ab wann werden komplexere sprachliche Ausdrücke z.B. Widersprüche oder Ironie, erkannt?
   a. Bereits ab 5 Jahren ist ausreichend Sprachwissen vorhanden, um solche Äußerungen zu erkennen.
   b. **Erst ab dem Schulalter (frühestens 6 Jahre) können komplexe Ausdrücke richtig gedeutet werden.**
   c. Erst mit dem Wechsel auf eine weiterführende Schule ist die Sprache weit genug entwickelt, um komplexe Sprachäußerungen zu erkennen.
   d. Bei jedem Kind ist das unterschiedlich, man kann keinen Richtwert angeben

5. Welche Einflussfaktoren bestimmen die Sprachentwicklung im Verlauf der Grundschulzeit?
   a. **Die Intelligenz des Kindes hat einen wesentlichen Einfluss auf die Sprachentwicklung.**
   b. **Wenn es ein sprachliches Vorbild in der Familie gibt, bestimmt dies maßgeblich die Entwicklung.**
   c. Schulische Förderung spielt keine große Rolle in der Entwicklung, da sie nur einen kleinen Teil des Alltags darstellt.
   d. Die individuelle Sprachbegabung spielt nur bis zum Kindergarten eine Rolle, danach sind es andere Faktoren.

6. Was macht Sprache humanspezifisch und damit so besonders?
   a. Sprache gibt es nur bei Menschen, etwas vergleichbar Komplexes gibt es nirgendwo sonst auf der Erde.
   b. **Menschen können sich mit Sprache auch auf Dinge beziehen, die außerhalb ihres Sicht- oder Verstandhorizonts liegen.**
   c. **Allein durch Sprache werden Dinge wie Gesetzte oder Staaten erst möglich.**
   d. **Durch Sprache lassen sich ständig neue Dinge erschaffen, indem kreativ sprachliche Wendungen kombiniert werden.**

7. Welche Voraussetzungen müssen für den Spracherwerb gegeben sein? Nennen Sie mindestens 3!
   a. **Ein gut entwickeltes Arbeitsgedächtnis erleichtert den Spracherwerb.**
   b. Visuelle Fähigkeiten sind nicht nötig, da Spracherwerb zunächst rein auditiv funktioniert.
   c. **Die Steuerung der Aufmerksamkeit durch Bezugspersonen begünstigt den Spracherwerb.**
   d. Es spielt keine Rolle, ob Sprache rein verbal der gestengestützt erworben wird.

8. Mit welchen intuitiven Sprachstilen unterstützen Eltern ihre Kinder?
   a. Verniedlichungen zur Anpassung an die Größe der Kinder.
   b. **Höhere Tonlage beim Sprechen zur Anpassung an die Hörfähigkeit von Säuglingen.**
   c. **Scaffolding – der Sprache ein Gerüst geben, um Worterwerb zu unterstützen.**
   d. **Bestätigende und korrektive Rückmeldung zu Aussagen von Kindern zum Aufbau von Grammatik.**

9. Was sind Late-Talker, und welche Prognose gibt es für ihre Entwicklung?
   a. Late-Talker sind Kinder, die die 50-Wort-Grenze erst mit 2 Jahren erreichen.
   b. **Late-Talker sind Kinder, die die 50-Wort-Grenze bis zum Alter von 2 Jahren nicht erreicht haben.**
   c. **Late-Talker können aufholen und sich zu Late-Bloomer entwickeln, wenn sie innerhalb eines Jahres die 50 Wörter erreichen.**
   d. **50% aller Late-Talker entwickeln eine Sprachentwicklungsstörung.**

10. Was könnten Ursachen einer (allgemeinen) Sprachentwicklungsstörung sein? Nennen Sie 3!
    a. **Zuviel TV-, PC-und Internetkonsum statt echter verbaler Kommunikation**
    b. **"Über-Förderung" des Kindes**
    c. **Körperliche Einschränkungen**
    d. Zu hohes Selbstkonzept

Fragen zu Lernmotivation

1. Wie würde eine selbstwertdienliche, wie eine nicht selbstwertdienliche Attribution aussehen?
   a. **Bei selbstwertdienlichen Attributionen werden Erfolg auf internale und Misserfolge auf externale, oder internal variable Ursachen zurückgeführt.**
   b. Bei selbstwertdienlichen Attributionen werden Erfolge auf externale Ursachen, Misserfolg auf internal-stabile Ursachen attribuiert.
   c. **Bei nicht selbstwertdienlichen Attributionen werden Erfolge auf externale Ursachen, Misserfolg auf internal-stabile Ursachen attribuiert.**
   d. Bei nicht selbstwertdienlichen Attributionen werden Erfolg auf internale und Misserfolge auf externale, oder internal variable Ursachen zurückgeführt.

2. Welche Kausalattribution zeigt Peter (Note 2): „Das Mathe-Talent habe ich von Papa geerbt, ich brauche gar nicht viel üben, und trotzdem funktioniert es immer."?
   a. **Internal**
   b. **Stabil**
   c. Variabel
   d. External

3. Welche Art von Aufgaben sind am sinnvollsten, um die Motivation zu erhalten und warum (bei Erfolgszuversichtlichen)?
   a. Leichte Aufgaben halten die Motivation immer hoch, da man sie leicht bewältigen und dennoch einen Erfolg daraus ziehen kann.
   b. **Mittelschwere Aufgaben halten die Motivation hoch, da sie den höchsten Informationsgehalt über die eigene Tüchtigkeit haben.**
   c. Schwere Aufgaben halten die Motivation hoch, da man sich anstrengen muss, um einen Erfolg zu haben.
   d. Die Aufgabenschwierigkeit spielt nur eine geringe Rolle – wichtig ist das eigene Durchhaltevermögen.

4. Was ist ein Beispiel für self-handicapping im Kontext der Kausalattribution?
   a. **Petra behauptet, sie kann kein Deutsch und lernt nicht für die Schulaufgabe, da sie darin keinen Sinn sieht. Als sie die Noten bekommen, hat sie eine 5.**
   b. Rolf ist nicht besonders gut in Deutsch. Er lernt viel, ist in der Schulaufgabe aber aufgeregt und macht sich zu viele Gedanken. Als er die Note bekommt, hat er eine 5.
   c. **Anna hat eine fünf in Deutsch bekommen. Sie redet sich jetzt ein, dass sie niemals eine bessere Note haben wird und dass sich alle Anstrengung ja überhaupt nicht lohnt und sie es deshalb auch gleich lassen kann.**

    d. Magnus hat eine strenge Lehrerin, die selbst bei sehr guten Leistungen nur Zweier verteilt. Er weiß, dass er ihren Ansprüchen nicht gerecht werden kann, gibt sich aber trotzdem Mühe. Als er in der Schulaufgabe, in der er sehr ängstlich war, eine 5 zurückbekommt, ist er enttäuscht.

5. Was sind Beispiele für Bezugsnormen und wie wirken sie?
   a. Wenn ich mich mit meinen früheren Leistungen vergleiche, ist das eine kriteriale Bezugsnorm.
   b. **Bei dem Vergleich mit einer individuellen Bezugsnorm mache ich die Erfahrung, dass Anstrengung kausal für Erfolg ist.**
   c. **Im sozialen Vergleich können Emotionen wie Stolz und Scham aufkommen.**
   d. **Die kriteriale Bezugsnorm sagt etwas über inhaltliche Kriterien aus.**

6. Wann ist das Messen von Leistungen an sozialen Bezugsnormen u.U. problematisch?
   a. **Soziale Bezugsnormen begünstigen Entwicklung leistungsstärkerer Schüler: innen, aber nicht leistungsschwächerer Schüler: innen.**
   b. Normalerweise leistungsstärkere Schüler:innen führen Misserfolge irgendwann auf mangelnde Fähigkeiten und Erfolge eher auf zufällige externale Ursachen (z.B. Glück) zurück, wenn sie im sozialen Vergleich immer wieder schlecht abschneiden.
   c. Individuelle Leistungszuwächse werden immer wichtiger, je mehr der soziale Vergleich ungünstig ausfällt, um negative Emotionen zu kompensieren.
   d. **Schüler: innen, die Leistungen ungünstig attribuieren, können durch soziale Vergleiche noch stärker in ihre Attributionsmuster verfallen.**

7. Woran kann es liegen, dass Lernmotivation ab der 6. Klasse absinkt, v.a. in naturwissenschaftlichen Fächern? Nennen Sie 2 mögliche Gründe.
   a. **Die Interessen der Jugendlichen verlagern sich meist auf andere Gebiete und die Motivation sinkt.**
   b. Das Elternhaus hat keinen Einfluss darauf, ob sich die Motivation ändert.
   c. **Der Korrumptionseffekt macht die Motivation aufgrund des Leistungsdrucks zunichte.**
   d. **Während der Pubertät wird das Selbstbild angreifbar und Schüler fokussieren sich auf Dinge, die ihnen liegen – daher verschiebt sich auch die Motivation**

8. Wie wirkt sich das Elternhaus auf die Leistungsmotivation aus?
   a. **Leistungsanforderungen sollten hoch, aber realistisch sein.**
   b. Eltern sollten bei Kindern die Autonomie einschränken, da sie das sonst ausnutzen und nichts lernen.
   c. **Ein warmherziges und unterstützendes Umfeld fördert die Entstehung von Leistungsmotivation**.
   d. Eltern sollten ihre Kinder möglichst viel loben, um ihnen zu zeigen, dass sie sie wertschätzen.

9. Wie wirkt sich die Schule auf die Leistungsmotivation aus?
   a. Die Schule hat wenig Einfluss – wichtiger ist, was die Eltern zuhause vorgeben.
   b. **Intrinsische Motivation kann durch schulische Anforderungen untergraben werden.**
   c. Wenn Schüler: innen Autonomie gegeben wird, verlieren sie ihre Motivation, da sie sich nicht anstrengen wollen.
   d. **Lob und Tadel in angemessenem Einsatz schadet der Motivation kaum**.

10. Was macht es schwer, Unterricht motivationsförderlich zu gestalten? Nennen Sie 3 Punkte!
    a. Wenn die Klassen zu klein sind, ist es schwierig den Unterricht motivationsförderlich zu gestalten.
    b. **Schüler: innen machen nicht immer gleich gut im Unterricht mit.**
    c. **Individuelle Beurteilung von Schülerleistungen lässt sich z.B. bei Tests nicht vermeiden.**
    d. **Kinder vergleichen sich auch untereinander, selbst wenn es Lehrende nicht tun.**

Tables & Further Results

The tables already integrated in the text are also available in one pdf-document on the DVD. Further, a analysis of the dropouts was made and data as well as syntax for this is to be found on the DVD in SPSS- as well as in Excel-format.

Some further results of the study, not mentioned in the text above, are stated below and can be also found on the DVD:

Perceived Comprehensibility, Usefulness and Information Density of the Study

Participants also provided feedback on the comprehensibility of the tasks, the usefulness of the study, that is, how useful participants judged the study for their personal learning, and the density of information learned in the study. Most participants (73.1%) perceived the study as very comprehensible, 23.9% as acceptable, and 3.0% did not judge the comprehensibility ($N = 67$). Most participants also found the study to be useful (37.3% useful, 31.3% very useful), and 26.9% rated them as medium useful for their own learning. Only 4.5% of participants found the study not very useful. The information density was judged as appropriate by 49.3%, 35.8% thought it to be rather appropriate, 11.9% did not judge it and 3% found it to be not appropriate.

Data source: *Appendix B – Chapter III > Tables & Further Results*

Online Data Source

All mentioned data and materials can also be accessed online via the OSF:

https://osf.io/wc3kh/?view_only=23bc6a041d4a467b87788fbe2cc3bd9b

Appendix C: Chapter IV

Datafile & Syntax

The data file and syntax can be accessed both as an SPSS version and in Excel format on the enclosed DVD. In addition to the data files, there is also a list of the occurring variables in the analysis and a file with an explanation of these.

Data source: *Appendix C – Chapter IV > Datafile & Syntax*

Materials

The questions of the practice and the criterial test are partly the same as in Appendix B (see Fragen zu Legasthenie, Fragen zu Dyskalkulie, Fragen zu ADHS, Fragen zu Denken, Fragen zu Intelligenz, Fragen zu Gedächtnis I) and can be additionally found on the DVD. The English translations can also be found on the DVD.

The DVD also contains a file with the open answers of the participants in raw and coded form for a better understanding of the evaluation.

Further, the instruction, already depicted in the text, is available on the DVD.

Data source: *Appendix C – Chapter IV > Materials*

Further Results

As further results, an analysis of the potentially confounding variables was implemented. The syntax, which can be run with the data file, can be found on the DVD.

Data source: *Appendix C – Chapter IV > Further Results*

Online Data Source

All mentioned data and materials can also be accessed online via the OSF:

https://osf.io/ncvy9/?view_only=83c6e1cf2c4c442bb16f4ec8de4a9a7b

Appendix D: Chapter V

Datafile & Syntax

The data file and syntax can be accessed both as an SPSS version on the enclosed DVD. In addition to the data file, there is also the SPSS output with the results.

Data source: *Appendix D – Chapter V > Datafile & Syntax*

Materials

The feedback questions asked are listed below and can be additionally found on the DVD. The English translations can also available on the DVD.

Data source: *Appendix D – Chapter IV > Materials*

Fragen zum Feedback (Deutsche Version/ German version – English version see DVD)

Wie bewerten Sie die Studie hinsichtlich ihrer ...?

	angemessen	eher angemessen	neutral	eher unangemessen	unangemessen
Länge der Studie	○	○	○	○	○
Informationsgehalt der Studie	○	○	○	○	○
Nützlichkeit für Ihr Lernen	○	○	○	○	○
Verständlichkeit in den Anweisungen	○	○	○	○	○
Organisation (Anmeldung, Informationen...)	○	○	○	○	○
Transparenz über das Forschungsvorhaben	○	○	○	○	○

Möchten Sie uns noch etwas für weitere Studien mitgben?