universität wien

MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master's Thesis

„Fishing for Competence – Validation of a self-report questionnaire assessing student's competencies for self-regulated learning"

verfasst von / submitted by

Armin Schmolmüller, BSc

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Master of Science (MSc)

Wien, 2016 / Vienna 2016

Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet:	A 066 840
Studienrichtung lt. Studienblatt / degree programme as it appears on the student record sheet:	Psychologie / Psychology
Betreut von / Supervisor:	Dipl.-Psych. Dr. Julia Klug

Acknowledgement

First of all, I want to thank my friend Gregor for bringing me into the family of PRO-SRL researchers and thus enabling me to be part of innumerable rich discussions. It is impossible to overestimate the true impact these amicable meetings had and will have on me. It is, yet again, impossible to overestimate how grateful I am for our encounter and what followed hitherto. Moreover, I want to thank you for your moral and professional support during the very writing of this thesis.

I also want to thank my supervisor Julia for having so much trust in my doing and for being available when needed. I consider myself very fortunate for having had you as my supervisor, not least because of your obligingness and good advice. Thank you very much!

Moreover, I want to express my deep gratitude to my colleague Andy who was an unexhaustable source of technical and methodical know-how. It truly is a delight to witness you juggling with figures and quaternions, and coming up with new solutions again and again. If only I had a hat, I would doff it!

Additionally, I want to thank Barbara Schober for confiding in me. I don't think I wouldn't have learnt as much if it wasn't for you trust. Thank you!

I also feel priviledged to have had my friend Ingo as a source of consult, also concerning the minor parts of the larger picture. Thank you for your moral and professional support during these last couple of weeks.

Furthermore, I want thank Laura, as well as my flatmates Alina, Chrissi, Jenny, Mathias, and Philipp for being understanding towards the minor or major excentricities I showed these days. Thank you for getting me out of one continuus loop or another.

Last but not least, I thank my family for being such a great financial and moral support, not only during the last couple of weeks but also during the time of my studies. Thank you for everything, I am extraordinarily grateful for having you in my life.

Table of content

Introduction

University students are challenged to autonomously acquire large bodies of knowledge (Wild, 2000) in complex learning situations without receiving many opportunities for external feedback (Peverly, Brobst, Graham, & Shaw, 2003). It is assumed that "self-determined and active efforts (i.e. actions, thoughts, and feelings) to initiate activities targeted towards learning goals, to perform them effectively, to monitor progress, and to adapt them if necessary" (Dresel et al., 2015, p. 455), also referred to as self-regulated learning (henceforth: SRL), are crucial for both developing high expertise and successfully mastering a degree course in higher education (e.g. Robbins et al., 2004; Wirth & Leutner, 2008). However, there is a comprehensive body of research on SRL that is both manifold and heterogeneous, and yet not always conclusive (Puustinen & Pulkkinen, 2001; Sitzmann & Ely, 2011). Thus, reasons for the heterogeneity of the findings in research on SRL are discussed to rather lie in the diversity of conceptualizations of SRL (Schmitz, 2003), in the variety of measurement approaches to SRL (Baumert & Köller, 1996; Jamieson-Noel & Winne, 2003) or in different contexts which research on SRL is focusing on (Boekaerts, 2007). By taking the heterogeneity in SRL research into consideration, many scholars (e.g. Azevedo, 2009; Schober et al., 2015; Wirth & Leutner, 2008) have suggested that in order to provide a more comprehensive understanding of SRL, the integration of different vantage points in terms of conceptualizations, measurement methods and contextual frameworks is indispensable.

The objective of this master thesis is to present first empirical evidence for the validity argument of a self-report questionnaire approach to assess context-specific competence components in higher education (Jöstl, Klug, Pichler, Schober, & Spiel, 2014; Jöstl, Klug, Schober, & Spiel, 2015). The instrument is derived from an integrative structural model of competencies for successful SRL in higher education (Dresel et al., 2015) and is complemented by other measurement approaches (see Engelschalk, Steuer, & Dresel, 2014; Roth & Schmitz, 2013), to constitute a multi-method-multi informant approach (Azevedo, 2009) for assessing SRL as a competence. However, although self-report questionnaires are most commonly used to assessed SRL in higher education (Roth, Ogrin, & Schmitz, 2015), considerable validity concerns have been expressed

towards these approaches (e.g. Samuelstuen & Bråten, 2007; Spörer & Brunstein, 2006; Veenman, 2011). Considering the limited scope of the present master thesis, this validation study is particularly concerned with the most critical and weakest assumptions of the instrument's design. In order to take into account that the instrument is still in the developmental phase, the conception of validation underlying this master thesis is derived from modern approaches introduced by scholars like Kane (1992, 2013a, 2013b) or Messick (1995) which regard validation as a crucial part in instrument development and validity as a property of proposed test score interpretations rather than of the instrument. Anyhow, it must be acknowledged that this master thesis cannot pursue an exhaustive validation as proposed by Kane or Messick – not least because that would require, amongst others, assessment triangulation with the other instruments derived from the structural model- and is therefore limited to the question of whether the instrument measures what it claims to measure (e.g. Borsboom, Mellenbergh, & van Heerden, 2004).

The structure of this master thesis embodies the following sections: First, the conceptual framework provided by the structural model of competencies for successful SRL in higher education (Dresel et al., 2015) will be presented. Second, theoretical and practical considerations proposed in literature about the assessment of SRL competencies will be discussed, particularly in regards to self-report questionnaire approaches. Third, the self-report measurement approach (Jöstl et al., 2014, 2015) will be introduced and located in the underlying conceptual framework (Dresel et al., 2015). Then, the conception of validation (Kane, 1992, 2013a, 2013b, 2013c) on which this master thesis is based will be reviewed and applied by stating the proposed interpretation of the test scores (i.e. that the instrument measures context-specific competence components in SRL) and then the weakest assumptions and inferences about the scoring, the generalization and the extrapolation. Based on a pilot study conducted by the developers of the instrument (Jöstl et al., 2014, 2015), as well as an expert rating, empirical evidence about the adequacy of the scoring inference (e.g. from expert rating and item analysis), the generalization inference (e.g. coefficients of internal consistency) and the extrapolation inference (e.g. correlation analyses) will be presented. On the basis of the empirical and methodological, further validity measures will be proposed and possible implications for the revision of the instrument will be discussed.

Conceptual framework: A structural model of SRL competencies

As a conceptual framework for defining SRL, for the construction of the scales, and for validating the instrument, the structural model of SRL competencies proposed by Dresel and his colleagues (2015) was used. The model provides an integrative approach for both measuring and analyzing various facets of successful SRL in higher education and aims to contribute to a more comprehensive understanding of SRL. The structural model proposes an integrative approach that builds up on Wirth and Leutner's (2008) definition of SRL as a competence and combines predominant conceptualizations of SRL, namely component-oriented models (e.g. Boekaerts, 1997, 1999; Boekaerts, Pintrich, & Zeidner, 2000; Pintrich, 1999, 2000) and process oriented models process-oriented models (e.g. Schmitz & Wiese, 2006; Wirth, 2004; Zimmerman, 2000; Zimmerman & Schunk, 2001). In addition, the structural model acknowledges that competencies in SRL can be defined by means of different types of knowledge (Paris, Lipson, & Wixson, 1983). Thus, the structural model offers a comprehensive framework for SRL by integrating these different conceptualizations and by acknowledging the dynamic and complex nature of the competencies underlying a successful regulatory mastery of learning challenges.

The concept of SRL

Although SRL has been subject of educational research for almost three decades (Dinsmore, Alexander, & Loughlin, 2008), there is still no universal definition (Roth et al., 2015), and consequently, no universal conceptualization of SRL. Generally, there is a variety of coexisting theoretical approaches to SRL (Puustinen & Pulkkinen, 2001) that, albeit emphasizing on different facets of SRL (Pintrich, 2000), can be subsumed according to basic assumptions about SRL. A rationale to meet the assumptions of all existing conceptualizations of SRL may characterize SRL "as an intermediate construct describing the ways in which individuals regulate their own cognitive processes within an educational setting" (Puustinen & Pulkkinen, 2001, p. 269). Certainly, a general definition as such could be considered the lowest common denominator of all existing conceptualizations of SRL and thus needs to be specified in order to be subject of further discussion.

For an exhaustive discussion of the various existing conceptualizations and theoretical approaches to SRL goes beyond the scope of this master thesis (for a comprehensive review, see e.g. Boekaerts, Pintrich, & Zeidner, 2000; Dinsmore, Alexander, & Loughlin, 2008; Puustinen & Pulkkinen, 2001; Schmitz & Wiese, 2006; Zimmerman & Schunk, 2011), the following chapters will place the focus primarily on the conceptual framework offered by Dresel and colleagues (2015). A brief overview of the basic assumptions underlying the predominant conceptualizations of SRL will be presented in the course of outlining the argumentation inherent in the structural model of competencies for successful SRL in higher education (Dresel et al., 2015).

Definition of SRL as a competence

The central assumption that underlies the structural model offered by Dresel and colleagues (2015) is that SRL can be understood as a competence. Therefore, the authors differentiate between 'SRL on the behavioral level' (e.g. the application of SRL strategies) and the competencies underlying the concrete application of these strategies. However, this distinction is rather theoretical, for Blömeke et al. (2015) argue that performance is the manifestation of competence. Consequently, competencies in SRL go beyond the application of learning strategies and embody concepts like the knowledge about strategies, meta-cognitive knowledge, and the capacity to regulate one's cognition and motivation as a means of execution control (e.g. Wirth & Leutner, 2008). Hence, SRL as a competence can be defined as, "a learner's competence to autonomously plan, execute, and evaluate learning processes, which involves continuous decisions on cognitive, motivational and behavioral aspects of the cyclic process of learning" (Wirth & Leutner, 2008, p. 103).

Dimensional Structure of SRL as a Competence

Based on the definition of SRL as a competence proposed by Wirth and Leutner (2008), the structural model by Dresel and colleagues (2015) combines three approaches to SRL prominent in research, each of which represents an axis on the cube displayed in Figure 1. First, the *strategy dimension* refers to a comprehensive body of component-

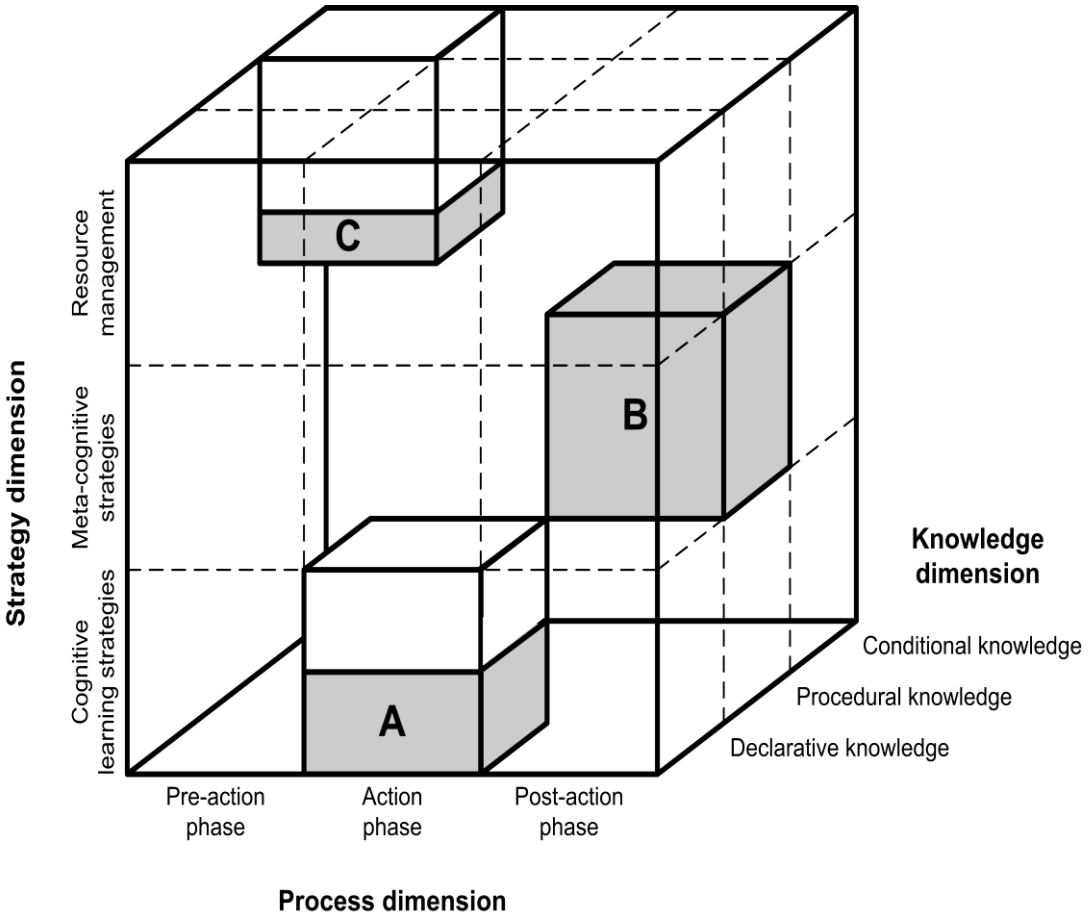

Figure 1. Proposed structural model of SRL competencies (with three marked examples A, B, and C for competence facets pronounced to varying extents) from "Competencies for successful self-regulated learning in higher education: structural model and indications drawn from expert interviews." by Dresel, M., Schmitz, B., Schober, B., Spiel, C., Ziegler, A., Engelschalk, T., … Steuer, G. (2015). Studies in Higher Education, 40(3), 454–470.

oriented models (e.g. Boekaerts, 1997, 1999; Boekaerts, Pintrich, & Zeidner, 2000; Pintrich, 1999, 2000), which aim to define SRL by means of at least three core strategy dimensions, namely cognitive strategies, metacognitive strategies und resource management. Second, the *process dimension* bears on process-oriented approaches (e.g. Schmitz & Wiese, 2006; Wirth, 2004; Zimmerman, 2000; Zimmerman & Schunk, 2001), which model SRL as a cyclic process with consecutive phases (pre-action, action, post-action). Third, the *knowledge dimension* acknowledges that competencies in SRL can be defined by means of three different types of knowledge (declarative, procedural, conditional; e.g. Paris, Lipson, & Wixson, 1983).

The strategy dimension.

According to component-oriented models (e.g. Boekaerts, 1997, 1999; Boekaerts, Pintrich, & Zeidner, 2000; Pintrich, 1999, 2000), SRL can be grouped into different types of strategies. Usually, the categorization of strategy types embodies at least three core strategy dimensions, cognitive strategies, metacognitive strategies, and resource management, respectively.

Cognitive learning strategies designate activities in which a learner uses specific strategies to regulate knowledge acquisition. Cognitive learning strategies embody deep-processing strategies like elaboration and surface-processing strategies like rehearsal (Pintrich, 2000).

Metacognitive learning strategies subsume activities and thoughts in which a learner engages in the process of learning on a higher-level. Metacognition can therefore be regarded as 'cognition about cognition' and refers to two aspects of a learners self-awareness (Flavell, 1979), namely the awareness of a knowledge base containing information about how, when, and where to apply a variety of cognitive strategies, and the awareness of and access to strategies that direct learning (e.g. evaluating the difficulty level of a task, adapting learning activities).

The management of resources include internal (e.g. regulation of motivation and emotion) and external (time management, management of the learning environment, and setting up social support for learning) resource management (see Dresel et al., 2015).

The process dimension.

Approaches focusing on the processes involved in SRL (e.g. Wirth, 2004; Zimmerman, 2000; Zimmerman & Schunk, 2001) typically model SRL as a cyclic process with consecutive phases, the pre-action, the action and the post-action phase, respectively (see Heckhausen & Gollwitzer, 1987; Schmitz & Wiese, 2006; Zimmerman, 2000).

In the pre-action phase, learners analyze situational and task demands, set goals, build up motivation, and make plans with respect to the use of learning strategies and resources (e.g. Schmitz & Wiese, 2006).

The action phase represents the process in which a learner engages in tasks and applies cognitive learning strategies, as well as he/she monitors the effectiveness of the

learning strategies in use. If necessary, learners may modify the learning activity and the learning strategies used. Moreover, students are challenged to maintain motivation and positive emotions during the learning activity (see Dresel et al., 2015).

The post-action phase is characterized by self-reflection and self-evaluation of the learning behavior, where learning outcomes are compared with learning goals and attributions. It is important to note that the cyclic nature of SRL implies that each variable can both be regarded as a predictor and an outcome. Thus, the conclusions drawn from the self-reflection and self-evaluation in the post-action phase are, on the one hand, affected by previous learning activities and, on the other hand, affecting future learning activities in terms of motivation, affect and strategy use (e.g. Schmitz & Wiese, 2006).

The knowledge dimension.

The knowledge dimension distinguishes SRL strategy knowledge by means of three different types of knowledge: declarative, procedural and conditional knowledge, respectively (Paris et al., 1983). First, declarative knowledge refers to factual knowledge (*know what*) about learning strategies (e.g. knowledge about the existence of learning strategies). Second, procedural knowledge comprises action knowledge (*know how*) about the implementation of learning strategies. Third, conditional knowledge embodies knowledge about the (in-) appropriateness of learning strategies for specific learning demands (*know when and why*; Paris, Lipson, & Wixson, 1983), and thus requires the analysis of the specific properties of these demands (see Wirth & Leutner, 2008). Furthermore, the three knowledge dimensions can be considered consecutive stages of competence development (Dresel & Haugwitz, 2006). Thus, the development from one knowledge type to the next implies additional knowledge being available. Thus, the step from declarative knowledge to procedural knowledge includes not only knowing about the existence of a specific SRL strategy but also to know how to apply the strategy. Consequently, the acquisition of conditional knowledge implies to have information available about the existence of various strategies, about how to apply these strategies, and about when and why to apply these strategies depending on the specific properties of the self-regulatory demands in a specific learning context.

Context-specificity of SRL competence.

It can be assumed that conditional knowledge plays an important role in the development of SRL competencies for self-regulatory learning processes may vary across situations and contexts depending on the specific regulatory demands a specific learning task presents (Winne, 2010). Similarly, competences (in general) have been described as "context-specific cognitive dispositions that are acquired and needed to successfully cope with certain situations or tasks in specific domains" (Koeppen, Hartig, Klieme, & Leutner, 2008, p. 62). Consequently, competencies in SRL can be regarded as cognitive dispositions for the autonomous and appropriate execution and control of a student's learning processes as a function of the specific SRL demands of a given learning situation (e.g. Dresel et al., 2015; Wirth & Leutner, 2008). Given that different learning situations can require the use of different self-regulatory strategies, a student competent in SRL might know that the use of the cognitive learning strategy 'memorization' may not be as helpful in a learning situation that requires reading comprehension than when studying for an exam that asks students to list the three dimensions and corresponding components named in the structural model described above.

Assessing Self-Regulated Learning

The number of measurement approaches to SRL is as rich as the number of conceptualizations, not least because different conceptualizations often lead to different measurement approaches (Boekaerts & Corno, 2005). In order to survey the variety of approaches to assess SRL, it can be expedient to categorize them by means of the standards of comparison they imply. These standards can differ in the application of either offline or online standards (e.g. Winne & Perry, 2000; Wirth & Leutner, 2008), as well as quantitative or qualitative standards (Wirth & Leutner, 2008).

Offline vs. Online Standards

Whether an instrument aims to monitor components of SRL (i.e. SRL viewed as an aptitude; see Winne & Perry, 2000) or the processes involved in a learning activity

(i.e. SRL viewed as an event; see Winne & Perry, 2000) determines if the standards which the means of assessment uses are to be labeled *offline* or *online* (e.g. Winne & Perry, 2000; Wirth & Leutner, 2008). *Offline standards* are based on component models of SRL (e.g. Boekaerts, 1997, 1999; Boekaerts, Pintrich, & Zeidner, 2000; Pintrich, 1999, 2000) and can be defined as independent of the learning processes (Wirth & Leutner, 2008). Consequently, they are used to measure if students possess the competencies considered crucial for successful self-regulation in a learning activity (e.g. knowing about learning cognitive and metacognitive learning strategies). *Online standards*, theoretically derived form process models of SRL (e.g. Wirth, 2004; Zimmerman, 2000; Zimmerman & Schunk, 2001), are applied when the process of learning and thus the process of self-regulation in learning are to be captured. Hence, online standards change individually and continuously in answer to alternating conditions and contexts of different learning activities (see Wirth & Leutner, 2008). Therefore, the difference between instruments based on *offline standards* and instruments based on *online standards* is that the former aim to assess the possession of competencies required to learn (i.e. the knowledge of strategies) while the latter aim to capture the process of self-regulated learning in the mastery of a given task, that is, whether or not learners are capable of applying their SRL competencies depending on the requirements of the task (cf. Wirth & Leutner, 2008).

Offline standards are usually applied in instruments like questionnaires, surveys, and interviews (Zimmerman, 2008) while online standards are used in measurement approaches like think aloud protocols, unstructured interviews or traces (Veenman, Bavelaar, De Wolf, & Van Haaren, 2014; Winne, 2010; Winne & Perry, 2000). Traces represent approaches that aim to resolve the challenge that mental processes are not observable by creating opportunities to trace these processes. Therefore, a trace can be viewed as a datum produced by a learner that occurs approximately simultaneous with his or her cognitive operations involved in the learning activity. A learner's trace can be the highlight of words in a text, inter alia (cf. Winne, 2010).

Quantitative vs. Qualitative Standards

Wirth and Leutner's (2008) further propose to distinguish between approaches to SRL that are based on *quantitative* or *qualitative standards*, whereas the former represent

standards that assume that the more SRL strategies a student applies, the better the regulation of the learning process will be. A *quantitative standard* as a maximum view is based on the assumption that the knowledge and use of as much learning strategies as possible is substantial for competent SRL. The standard of these kinds of measurement approaches is defined as the maximum rating across all items (Wirth & Leutner, 2008). Videlicet, the more strategies respondents know and the more they use them, the better is their score (i.e. their SRL competency).

Qualitative standards, on the other hand, follow the rationale that a student's competence in self-regulated learning depends on how well the strategies applied fit the situation- and task-specific demands. Wirth and Leutner (2008) propose that *qualitative standards*, that is, in how well learning activities match the specific self-regulatory demands in a given situation, are to be derived from theoretical models and/or expert judgments. Similarly, Dresel and colleagues (2015) accentuate that applying all possible SRL strategies in a given learning situation may not always be appropriate. Consequently, they argue that standards used to evaluate a learner's SRL competencies should take into account the fit between the specific self-regulatory demands in a learning situation and the extent to which a learner's SRL activities meet these demands. However, the most commonly used instruments to assess SRL are self-report questionnaires using quantitative offline standards (Roth et al., 2015).

Self-Report Questionnaires in SRL research: Using Quantitative Offline Standards

Self-report questionnaires are most commonly used to assess SRL in higher education (Spörer & Brunstein, 2006), largely for economical reasons (e.g. large-scale assessments or easy administration for both participants and practitioners; Cromley & Azevedo, 2011). Further reasons for the popularity of self-report questionnaires in SRL research is that they do not interfere in the learning process (Veenman, 2005) and that certain aspects of SRL processes are not accessible via direct observation (e.g. metacognitive activities; see Floden, 1981). Two of the most commonly used self-report questionnaires in Anglo-American SRL research (Winne & Perry, 2000) are the Motivated Strategies for Learning Questionnaire (MSLQ; Pintrich, Smith, Garcia, &

Mckeachie, 1993) and the Learning and Study Strategies Inventory (LASSI; Weinstein, Palmer, & Schulte, 1987). For German-speaking populations, the Inventory to Assess Learning Strategies of Students (LIST, Inventar zur Erfassung von Lernstrategien im Studium [Learning and Study Strategies Inventory]; Wild & Schiefele, 1994) constitutes a popular instrument to assess SRL in higher education (Spörer & Brunstein, 2006).

Generally, self-report questionnaires in SRL research consist of a number of questionnaire items per type of (sub-) strategy (e.g. "help seeking" as a subscale of the learning strategy "resource management" in MSLQ; see Pintrich et al., 1993), each of which asks respondents to rate their tendency to use the learning strategy represented by the item, typically on a Likert-type scale (see Pintrich, Smith, Garcia, & Mckeachie, 1993; Weinstein et al., 1987; Wild & Schiefele, 1994). Most self-report questionnaires aiming to assess SRL apply quantitative standards (Roth et al., 2015).

Assessing Self-Regulated Learning as a competence

The distinction between offline and online standards (e.g. Winne & Perry, 2000; Wirth & Leutner, 2008) can also be traced back historically in their temporal order of occurrence. According to Panadero, Klug and Järvelä (2015), three different waves of SRL measurements can be identified. The first wave is determined by the assessment of SRL via self-report measurements applying offline standards (e.g. questionnaires, surveys, and interviews). The second wave embodies measurements that aim to capture SRL as a process or an event and thus apply online standards (e.g. thinking aloud protocols, traces). Third, the newest wave implies an innovation in the conceptualization of SRL measurement and consists of measurements that concurrently act as an intervention (e.g. learning diaries).

However, Panadero and his colleagues (2015) do not suggest that distinct times call for distinct measures, or, in other words, that a prior wave of measurements should be superseded by the next one. On the contrary, they argue that these different modes of measurement should be considered complementary in their coexistence and can – in interweaving – offer new insights into the complex nature of SRL.

This integrated view has become relevant particularly since SRL is conceptualized as a competence (Wirth & Leutner, 2008; see also Dresel et al., 2015). Hence, there is a

call for new ways to measure SRL (Schober et al., 2015). The acknowledgement of the multi-dimensional nature of competences (Blömeke, Gustafsson, & Shavelson, 2015) requires an integrational perspective (Schober et al., 2015) that pursues systematic multi-method-multi-informant approaches (Azevedo, 2009). Hence, two major advantages could come along with converging multiple sources of data in the assessment of SRL: (1) Given that different measurement approaches (i.e. qualitative and qualitative) aim to assess different dimensions of SRL, they could therefore provide a comprehensive understanding of the underlying (meta) cognitive and self-regulatory processes involved in SRL (Azevedo, 2009). (2) The examination of student´s SRL competencies from different vantage points using different methods of assessment allows for assessment triangulation (see Pellegrino, Chudowsky, & Glaser, 2001) and, if the results appear consistent, researchers can assume with reasonable confidence that major aspects of reliability and validity of their measures have been attained (Boekaerts & Corno, 2005). Evidence for the adequacy of multi-method approaches on SRL was provided by Souvignier and Roes (2005) who compared a self-report questionnaire on learning strategies and learning diaries. Although the results indicated basically no correlation between the responses to each instrument, both instruments could explain a substantial amount of the variance in learning outcomes on their own.

Consequently, Schober and her colleagues (2015) argue that a comprehensive understanding of SRL as a competence requires an integrated theoretical model of SRL (e.g. Dresel et al., 2015), which can provide a conceptual framework for the derivation of different measurement approaches that focus on different aspects and dimensions of SRL competencies. A comprehensive approach to SRL measurement also requires the application of different combinations of standards (see Wirth & Leutner, 2008), and qualitative standards are particularly interconnected with context-specificity for they take into account the fit between the specific self-regulatory demands in a learning situation and the extent to which a learner's SRL activities meet these demands (Dresel et al., 2015).

In what follows, the measurement approach that is derived from the integrated theoretical model of SRL competencies provided by Dresel and colleagues (2015) will be introduced.

The measurement approach

The measurement approach under consideration (Jöstl et al., 2014, 2015, see Appendix A and B) constitutes a German online (here: internet-based) self-report questionnaire using Unipark (www.unipark.de) that is still in the development phase. In order to impede confusions about self-report questionnaire approaches in general and the self-report questionnaire approach under consideration, the latter will henceforth be referred to as IWAN[1] (Instrument without a name). Together with two online measures, namely a situational judgment test (Engelschalk et al., 2014) and a learning diary (Roth & Schmitz, 2013), it complements a multi-method-multi informant approach (Azevedo, 2009) embedded in the joint research project PRO SRL to assess SRL as a competence as derived from the structural model of competencies for successful SRL (Dresel et al., 2015) described above. IWAN is designed to approach the SRL competence components (strategy dimension in the structural model) cognition, metacognition and emotional regulation (regulation of frustration, regulation of boredom), as well as declarative and procedural knowledge (knowledge dimension in the structural model) concerning corresponding learning situations. The two other complementary instruments of the joint research project both aim to assess SRL competencies in the three phases of the learning cycle (process dimension on the structural model), whereas the situational judgment test (Engelschalk et al., 2014) aims to assess strategies to regulate motivation (strategy dimension in the structural model) while the learning diary (Roth & Schmitz, 2013) is designed to measure metacognitive competencies (strategy dimension in the structural model).

Because IWAN also provides exploratory information (e.g. ranking of reasons for not using a specific strategy; see also Figure 3) and for the scope of this master thesis is limited, the whole instrument in its entirety can only be outlined whereas the aspects relevant for this validation study will be described in detail (for a more detailed description of the whole instrument, see Jöstl et al., 2014, 2015).

[1] For it is still in developmental stage, no final name has been assigned to the instrument yet.

Core Assumptions and Development

The core assumption underlying IWAN is that for assessing SRL competencies, instruments should be tailored to particular tasks or contexts (Bråten & Samuelstuen, 2007; Richardson, 2004). Based on this assumption, Dresel and his colleagues (2015) examined beneficial and adverse strategies for representative learning situations (e.g. exam preparation) in a sequential mixed-method study among experts in the field of SRL (i.e. university teachers and excellent students). Results of this study showed that experts found different sets of learning strategies to be relevant for different learning situations. Thus, the results confirmed the assumption of context- and situation-specificity of SRL and provided meaningful combinations of learning situations and correspondingly beneficial and adverse strategies for the development of a set of measurement approaches (Dresel et al., 2015).

Self-regulatory strategies as a function of situation-specific requirements.

IWAN aims to assess SRL competence components found to be pertinent in specific learning situations (see Dresel et al., 2015). It is important to note that each situation requires a variety of different self-regulatory strategies (Dresel et al., 2015). Thus, it can be assumed that in all of the situations also motivation regulation, among others, can be considered relevant (cf. Dresel et al., 2015). Anyhow, the fit between learning situations and relevant SRL strategies assumed in IWAN is based on the preliminary study described above (Dresel et al., 2015), and the following combinations were found to be of high relevance (cf. Jöstl et al., 2014, 2015; also illustrated in Figure 2): (1) Exam preparation requires the application of cognitive learning strategies (2) Writing a minor scientific work requires metacognitive learning strategies (3) In writing a major scientific work, strategies to regulate frustration has found to be of high relevance (4) Lecture attendance challenges students to apply strategies to regulate boredom (Jöstl et al., 2014, 2015).

The relevant learning situations are characterized as follows: (1) In exam preparation, students are challenged to process and learn large quantities of subject matter autonomously whereat the profound understanding of the subject matter is essential. (2) Writing a minor scientific work (e.g. seminar papers, essays) contains searching and

Learning situation		Relevant SRL strategies
Exam Preparation	×	Cognitive learning strategies
Writing a minor scientific work	×	Metacognitive strategies
Writing a major scientific work	×	Strategies to regulate frustration
Lecture attendance	×	Strategies to regulate boredom

Figure 2. Combination of learning situations and relevant SRL strategies.

processing further literature for a scientific text, which has to be handed in at a fixed date. (3) Writing a major scientific work (e.g. diploma thesis, master thesis, and bachelor thesis) requires working autonomously for several months on an extensive subject matter supported by a docent. (4) Lecture attendance refers to lectures in which a docent discusses a large body of study matter in front of a rather large number of students.

Tailored Vignettes: Introducing the relevant learning situation.

Corresponding with the four situations named above, IWAN is composed of four distinct elements, each of which aims to assess a specific competency in SRL (metacognition, cognition, regulation of frustration, regulation of boredom) that were found relevant in a specific learning situation (see Figure 2). Each learning situation is introduced in a tailored vignette presenting the situation (e.g. exam preparation) and the general learning requirements in the situation (e.g. in exam preparation, students are challenged to process and learn large quantities of subject matter autonomously whereat the profound understanding of the subject matter is essential). Moreover, each of the vignettes contains the remark that in a study, students found the application of a specific set of SRL strategies (e.g. learning strategies) in the specific learning situation (e.g. exam preparation) particularly relevant (Jöstl et al., 2014, 2015).

Making a difference: Strategy knowledge and strategy use.

In each of the vignettes the basic rationale of IWAN is introduced by asking two general questions (that are not to be answered directly): (a) *What is, in your opinion, beneficial in the given learning situation* (e.g. exam preparation)*?* (b) *What do you*

actually do when engaging in the given learning situation (e.g. exam preparation)*?* This distinction made between knowledge and performance is based on the assumption that both constitute two components of competences (cf. Blömeke et al., 2015). Thus, it is assumed insufficient to only have vast strategy knowledge. In order to be considered competent in SRL, it is deemed necessary also to transform this knowledge into action (Jöstl et al., 2014, 2015). IWAN therefore aims to assess both components (knowledge about strategies and use of strategies). Consequently, each item consists of a statement that requires two answers on dichotomous response formats ("yes" or "no") that refer to abbreviations of the two questions mentioned above: (a) *Is this beneficial?* (b) *Do you do this?* (Jöstl et al., 2014, 2015).

The hierarchical order of knowledge: declarative and procedural knowledge.

IWAN also aims to assess declarative and procedural strategy knowledge. Thus, the instrument is composed of two levels of competence that are based on the presumed hierarchical order of declarative and procedural knowledge (Dresel & Haugwitz, 2006). The former is general knowledge about the benefit of a specific strategy in a given situation (e.g. structuring the subject matter in an exam preparation) and the latter is the knowledge of which sub-strategies a strategy consists of (know-how; e.g. summarizing the key points of each paragraph in order to structure the subject matter, cf. Paris et al., 1983). Thus, each procedural knowledge scale (each assessing a relevant sub-strategy) is represented by one single item in the declarative knowledge scale (see Figure 3). Due to this specific property IWAN´s design, only respondents that have reported declarative knowledge of a specific strategy (i.e. they recognize strategy as beneficial and state that they apply the strategy in the given situation) can advance to the procedural level of this specific strategy dimension. As such, the advancement from the declarative knowledge level to the procedural knowledge level in the assessment of a specific sub-strategy can vary inter-individually as a function of the answer to the corresponding item on the declarative knowledge scale. On this account, IWAN provides an individual profile of a respondent's declarative and procedural knowledge rather than a general means of comparison between respondents.

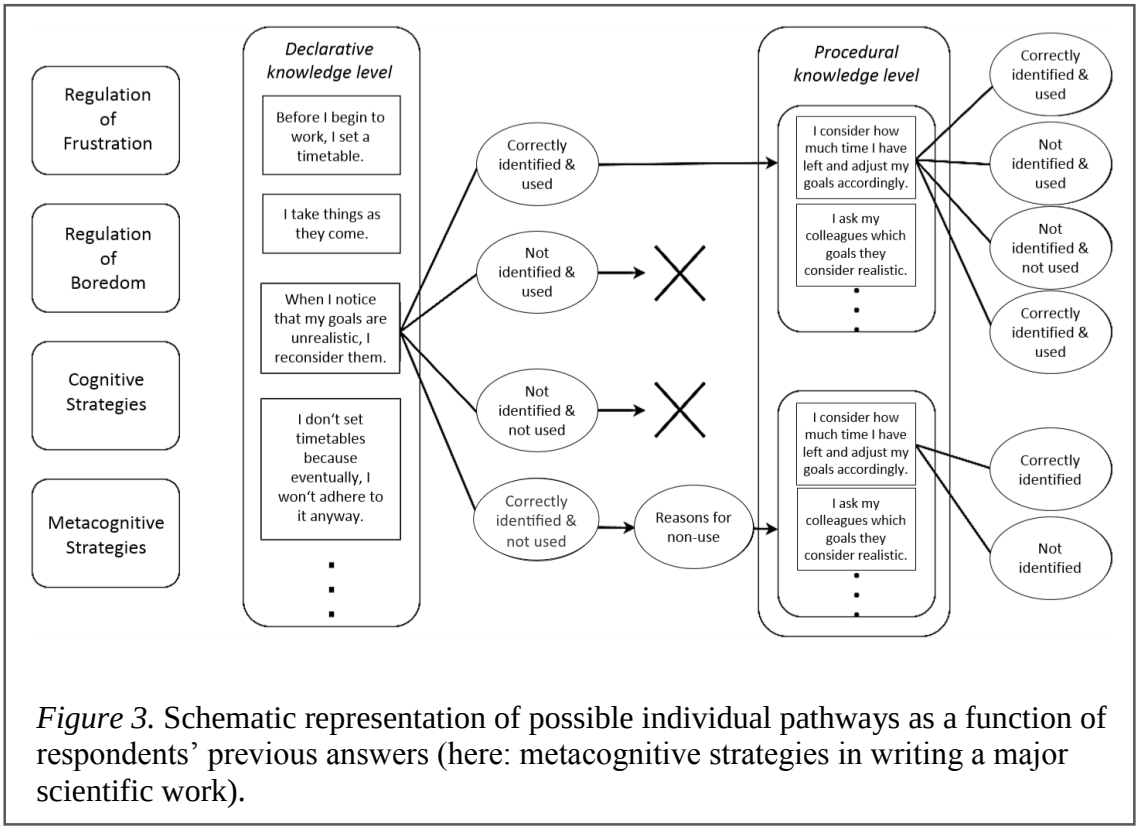

Figure 3. Schematic representation of possible individual pathways as a function of respondents' previous answers (here: metacognitive strategies in writing a major scientific work).

Item and scale development.

The development of the items for IWAN (Jöstl et al., 2014, 2015) was based on the expert's reports obtained from the exploratory study reported by Dresel and colleagues (2015). Hence, the presumed fit between specific SRL strategies and the demands for SRL in a given situation (e.g. cognitive strategies in exam preparation) was considered valid. Consequently, items were created and collected that were assumed to represent the reported beneficial and adverse strategies corresponding with the requirements in a given situation. The item pool consists of self-constructed items both empirically (i.e. based on the sequential mixed-method study among experts in the field of SRL by Dresel et al., 2015) and/or theoretically founded (i.e. based on the conceptualizations underlying the structural model of successful SRL; Dresel et al., 2015). Moreover, items from established self-report questionnaire scales such as the LIST (Wild & Schiefele, 1994), among others, for cognitive and metacognitive strategy scales, or the SVF-KJ (Stressverarbeitungsfragebogen von Janke und Erdmann angepasst für Kinder und Jugendliche [Coping Questionnaire for children and adolescents]; Hampel,

Petermann, & Dickow, 2001), inter alia, for regulation of frustration scales were adapted (for a list of items and detailed information about the source of each item, see Appendix B).

The scoring system

The intended scoring system for IWAN is based on the assumption that strategy knowledge and strategy use both contribute to competences in SRL (cf. Blömeke et al., 2015). Hence, provision is made for creating a new *composite scale* (cf. Raykov, 1997) that is a function of both correct knowledge of beneficial and adverse strategies, and the reported use of beneficial as well as non-use of adverse strategies. It is intended that the new composite scale shall be created for each strategy and knowledge dimension separately and provide the basis for the calculation of the *composite sum score* as an overall indicator of the latent competency. Therefore, four consecutive calculation steps are required:

(1) It is ensured that the responses on each index (strategy knowledge and strategy use) are coded in '0' for '*no*' and '1' for '*yes*'.

(2) In order to maintain unipolarity, the polarity of items that represent adverse strategies is reversed in both indices (strategy knowledge and strategy use).

(3) For creating each item of the new composite scale, a combined competence value (X) is calculated as a multiplicative function of the value in strategy knowledge ($X_{knowledge}$; '0' or '1') and strategy use (X_{use};'0' or '1') that again is dummy coded (i.e. '0' for '*no*', 1'' for '*yes*'). It is calculated as follows:

$$X = X_{knowledge} \times X_{use}$$

(4) The composite sum score (X_{sum}) of a scale with *k* items as an overall indicator of the latent SRL competency component assessed in the scale is calculated as follows:

$$X_{sum} = \sum_{i=1}^{k} X_i$$

The intended scoring system thus applies quantitative standards (cf. Wirth & Leutner, 2008). Before the adequacy of the scoring system and various other assumptions can be validated, it is necessary to outline the concept of validity underlying this master thesis.

The concept of validity

The concept of validity has been subject to much controversy and misconceptions in the last 100 years (Borsboom et al., 2004; Kane, 2013c). The conventional definition of validity is that a test is valid if it measures what it claims to measure (e.g. Kelley, 1927). Contemporary views of instrument validation (Kane, 1992, 2013a, 2013b, 2013c; Messick, 1995), though, note that validity is not a property of tests but that of test score interpretations. Thus, the core idea of test validations is to state the proposed interpretations and use of an instrument explicitly (i.e. IUA, interpretation/use argument; Kane, 2013a) and then to evaluate the plausibility of those proposals (i.e. validity argument; Kane, 2013a). The validation process as illustrated in Figure 4 consists of several steps. (1) The proposed interpretations and use of an instrument are stated. (2) Claims, assumptions, inferences derived from the stated interpretations and used are formulated. (3) The weakest assumptions and claims are tested. (4) The evidence collected in (3) is evaluated. (5) Based on the conclusions drawn from the evaluation (i.e. whether or not the evidence supports the proposed interpretation and use) a decision is made, either to accept or reject the validity argument, or, in case of unfavorable evidence, to revise the instrument or the proposed interpretation and use.

The proposed interpretations and uses can only be valid to the extent that the IUA is complete and coherent. Failing to state the proposed uses and interpretations in the IUA consistently makes a fully adequate and exhaustive validation basically impossible (see Kane, 2013b). Consequently, test validation requires a consistent argumentation that "specifies the sequence or network of inferences and assumptions involved in getting

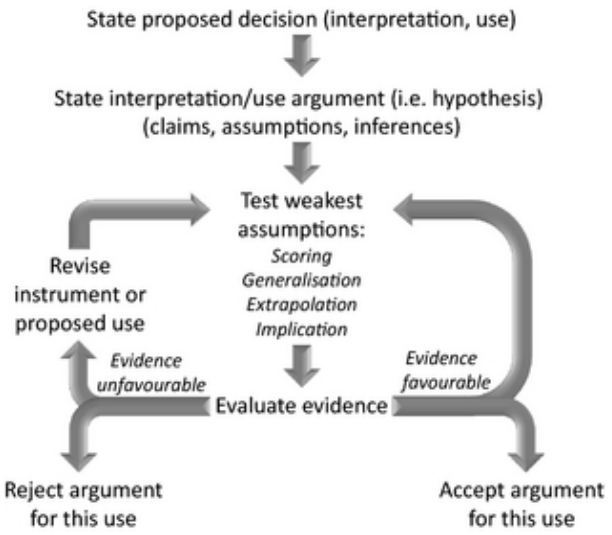

Figure 4. Illustration of the process of validating the validity argument from "A contemporary approach to validity arguments: a practical guide to Kane's framework." by Cook, D. A., Brydges, R., Ginsburg, S., & Hatala, R. (2015). *Medical Education*, *49*(6), 560–575.

from a test taker's observed test performances to the conclusions and decisions to be based on these performances" (Kane, 2013a, p. 448). In recapitulation of Kane's framework, Cook, Brydges, Ginsburg and Hatala (2015) outline that "Kane traces an assessment from the *Scoring* of a single observation […], to using the observation score(s) to generate an overall test score representing performance in the test setting (*Generalization*), to drawing an inference regarding what the test score might imply for the real-life performance (*Extrapolation*), and then to interpreting this information and making a decision (*Implications*)" (p.563; italics in original).

Claims, assumptions, inferences derived from the stated interpretation of the instrument

As noted, the self-report questionnaire (Jöstl et al., 2014, 2015) seeks to assess context-specific competencies for successful SRL in higher education (metacognitive strategies, cognitive strategies, strategies to regulate frustration, strategies to regulate boredom) on two knowledge dimensions, namely declarative and procedural knowledge.

For this master thesis aims to provide first evidence for the validity of the self-report questionnaire (Jöstl et al., 2014, 2015), it is primarily concerned with the weakest assumptions and claims about the validity argument of the self-report instrument. Nonetheless, it is intended to embed this – albeit perforce limited - validation in the broader framework provided by Kane (1992, 2013a, 2013b, 2013c), not least because the self-report questionnaire approach (Jöstl et al., 2014, 2015) is still in the development phase and Kane´s framework regards validation as a crucial part of instrument development (as illustrated in *Figure 3*; see also Cook et al., 2015). Thus, this master thesis cannot claim to represent an exhaustive validation study in the sense of Kane (1992, 2013a, 2013b, 2013c) for three main reasons:

(1) The validity argument is limited to the *proposed interpretation* of the instrument's scores for it is embedded in the joint research project PRO SRL and the *proposed use* of the instrument would be to contribute to a comprehensive understanding of successful SRL in higher education.

(2) Failing to evaluate the proposed use of the instrument perforce makes it impossible to discuss the consequences of the assessment on the learner, the stakeholders or society at large (cf. Cook et al., 2015). Thus, testing the adequacy of the *implication inference* as a crucial part in the evaluation process cannot be pursued in this master thesis.

(3) An exhaustive validation study would include, amongst others, assessment triangulation (cf. Pellegrino et al., 2001) with the other instruments derived from the structural model (Dresel et al., 2015).

Considering these limitations, the IUA argument is now reduced to an *interpretative argument* (note the difference to *interpretive argument*, that is, in Kane´s terminology, exactly the same as the IUA; see Kane, 2013c, p. 2) that can only consider the proposed interpretation of the instrument. Thus, this validation study is limited to the claim that the scales of the instrument measure what they purport to measure (cf. Borsboom et al., 2004). It is important to note that this does not imply the assumption that the instrument is valid when it measures what it claims to measure (cf. Kelley, 1927) but that the proposed interpretation can be considered valid (cf. Kane, 1992, 2013a, 2013b, 2013c). Hence, the *interpretative argument* concerning the claim that the scales measure

the variables they are intended to measure (i.e. context-specific competence components for successful SRL in higher education) is as follows (cf. Hopfenbeck & Maul, 2011; see also Cook et al., 2015; Kane, 1992, 2013a):

(1) The *scoring inference* from observed performances to the observed score is adequate. In support of the scoring inference, evidence is required of the appropriate, coherent, and unbiased application of the scoring rules.

The evidence for the validation argument can come from item analysis (e.g. item difficulties) and/or analysis of the scoring rubric/criteria and the implementation procedures (e.g. Cook et al., 2015).

(2) The generalization inference from the observed score to the "universe score" is adequate (i.e. the selection of items on the self-report questionnaire can be assumed to be representative of the universe of possible similar items). The generalization inference can be supported by evidence coming from expert judgment and logical argumentation. Another source of evidence can come from reliability analysis in terms of internal consistency or retest-reliability.

(3) The *extrapolation inference* from the "universe score" to the target domain is adequate (that is, the universe of possible questionnaire items is reasonably and/or empirically related to the constructs that the scales aim to measure).

The extrapolation inference is typically supported by qualitative evidence and/or argumentation that the processes that are involved in participants' item responses overlap sufficiently with the processes identified in theory about the construct to be assessed.

The adequacy of the extrapolation inference can also be examined by means of the relationship between the observed test scores and other variables and outcomes that are consistent with the assumptions of the underlying theory of the construct (i.e. criterion validity).

In sum, the adequacy of the extrapolation inference of learning strategy scales evaluates whether students respond to the test items in a manner consistent with the predictions of theory, and, whether their responses can be

taken as indicative of their real-world learning behaviors (cf. Hopfenbeck & Maul, 2011).

Validation of Self-Report Questionnaires in SRL research

According to Winne and Perry (2000), the validation of self-report questionnaires in SRL research are validated according to a 3-step script: First, an exploratory factor analysis of a sample's responses is applied in order to investigate the compatibility of those responses and the theoretical model of SRL underlying the development of the single items. Second, almost always coefficients of internal consistency are determined as means of the reliability of the full scale and the subscales (if there are any). Third, correlations between the scales and external variables are calculated. The most commonly used external criteria that SRL self-report scales are correlated with are measurements of achievement.

In sum, the validation of self-report questionnaires in SRL research is mainly concerned with the isolated examination of the adequacy of the *generalization inference* (steps 1 and 2) and of the *extrapolation inference* (step 3). Certainly, these validation steps can be considered crucial in the examination of the validity arguments of a self-report questionnaire (Hartig, Frey, & Jude, 2008). However, there is a number of scholars (e.g. Samuelstuen & Bråten, 2007; Schellings, 2011; Schellings & Van Hout-Wolters, 2011; Spörer & Brunstein, 2006; Veenman, 2011; Winne, 2010; Winne & Jamieson-Noel, 2002; Wirth & Leutner, 2008) that have expressed concerns about the validity of self-report questionnaires aiming to assess SRL. Regarding the generalization inference, published SRL self-report questionnaires show moderate to good coefficients of internal consistency and are not so much subject of validity concerns. The MSLQ subscales have been reported to have Cronbach's α that range between .52 and .92 (Pintrich et al., 1993), and the LASSI subscales show Alphas from .68 to .86 (Weinstein et al., 1987). However, the concerns regarding the validity of self-report questionnaires assessing SRL primarily concern the adequacy of two inferences presented above, that is, the scoring inference, and the extrapolation inference, respectively.

Concerns about the adequacy of the scoring inferences.

As noted above, self-report questionnaires aiming to assess SRL usually apply quantitative standards (Wirth & Leutner, 2008). These standards imply that the more strategies respondents know and the more they use them, the better is their score (i.e. their SRL competency). Although the application of quantitative standards (calculation of a total score across items) can – given that no human error takes place – be considered direct and flawless (Hopfenbeck & Maul, 2011), Wirth and Leutner (2008), however, argue that the application of these kinds of standard can lead to a reduced validity of the instruments for there is no theoretical model underpinning the maximum view (Leopold, den Elzen-Rump, & Leutner, 2007). There are attempts to introduce qualitative standards to offline measures, which often imply expert ratings (see Wirth & Leutner, 2008).

Concerns about the adequacy of the extrapolation inference.

The adequacy of the extrapolation inference is given when the universe of possible questionnaire items is reasonably and/or empirically related to the constructs that the scales aim to measure (Kane, 1992, 2013a, 2013b, 2013c). The validation of self-report questionnaires measuring SRL usually embodies the examination of the relationship between the scales of the instrument under consideration and some external criteria, typically learning outcomes (i.e. measures of achievement; Winne & Perry, 2000). Self-regulated learning has been reported to have great relevance for learning success (e.g. Robbins et al., 2004). Despite these assumptions, correlations between SRL self-report questionnaires and learning outcomes have been reported to be very low (Veenman, 2005). In a meta-analysis of 21 studies using self-report questionnaires Veenman (2002; as cited in Veenman, 2005) found out that on average not more than 3% of the total variance in learning outcomes could be explained by the self-report instruments. Pintrich and his colleagues (1993) reported moderate correlations between learning strategy scales and final course grades for the MSLQ that range from .02 to .41. However, the validity of grades as a predictor for learning achievement is controversial (see Tight, Mok, Huisman, & Morphew, 2009). Thus, Boerner, Seeber, Keller and Beinborn (2005), for instance, used subjective measures of learning achievement in addition to grades and found out that the reported use of learning strategies collected with the LIST (Wild &

Schiefele, 1994) could predict the grades of students (R^2 = .067) to a substantial lesser extent than the subjective learning achievement (R^2 = .255).

Anyway, the reasons for low validity in general and little to modest predictive power of SRL self-report questionnaires on learning achievement in particular have been discussed to be manifold (e.g. Spörer & Brunstein, 2006). Generally, they concern the question about how self-report questionnaires lead to responses and how to interpret these responses (van Hout-Wolters, 2009; as cited in Schellings, 2011). Winne and Perry (2000), for instance, outline the fact that most of the instruments ask respondents to generalize their actions across situations and therefore disregard the situation-specificity of SRL (Artelt, 2000). In a study, Leopold and her colleagues (2007) showed that enhanced situation-specificity results in higher correlations with learning outcomes.

Another source of concern about the adequacy of the extrapolation inference can be illustrated by means of Hasselhorn's (1995) concept of strategy acquisition. As Spörer and Brunstein (2006) point out, the fact that strategies are usually pre-formulated could conceal possible knowledge deficits and/or lack of actual action although the application of learning strategies was reported. Consequently, although they might respond to an item that the use of the strategy is adequate in the situation and/or that they use the strategy in this situation because they know or believe this strategy to be effective, that needs not necessarily signify that they actually use those strategies (Samuelstuen & Bråten, 2007). The fact that self-report questionnaires may not be appropriate means of distinguishing between different stages of strategy acquisition could furthermore be accompanied by challenges inherent in the method of self-report questionnaires (Veenman et al., 2003), namely the tendency to choose ‘agree’ or ‘true’ as a response category (“acquiescence”; see Samuelstuen, 2003) and the tendency to affirm response categories that reflect socially approved behaviors (Samuelstuen, 2003). Veenman and his colleagues (2003) further indicate that respondents may differ in the choice of reference points. Students may base their answers on their own individual standards, the (alleged) viewpoint of a teacher or the (alleged) ideal student or some other reference point. Consequently, this could lead to a reduced informative value of the responses and to a reduced validity, respectively. Moreover, Leopold and Leutner (2002) argued that self-report questionnaires rather draw on learners’ recognition of strategies than on their declarative strategy knowledge.

Aims of this study

Although this master thesis acknowledges the validity concept proposed by modern scholars like Kane (1992, 2013a, 2013b, 2013c) or Messick (1995), it does not and can not claim to represent an exhaustive validation. Thus, this master thesis is limited to the aim of providing first empirical evidence and is therefore primarily concerned with the weakest assumptions of the instruments validity argument (cf. Kane, 1992, 2013a; see also Cook et al., 2015).

Considering the concerns towards the validity of self-report questionnaire approaches to SRL presented above, this master thesis aims to evaluate the validity argument of IWAN (i.e. the adequacy of the scoring inference, of the generalization inference and of the extrapolation inference) by putting the following questions up for discussion:

1. *Can the scoring inference from the observed performances to the observed score be assumed to be adequate?*

   The adequacy of the scoring criteria will be discussed on the basis of evidence drawn from a quantitative expert rating (inter-rater-reliability), analyses of item difficulties, and by logical argumentation.

2. *To which extent can it be assumed that the items of a scale account for one underlying factor?*

   This question refers to the adequacy of the generalization inference, and will be discussed by means of coefficients of internal consistency, and by logical argumentation.

3. *To which extent can the responses provided by students be assumed to be related to academic achievement?*

   The relationship between competencies in SRL assessed by the instrument (composite sum scores) and academic achievement (mean of the grade in the last three courses, grade in the corresponding learning situation, subjective learning achievement, subjective content with learning achievement) will be determined via correlation analyses ($\alpha = .05$).

Due to specific qualities in the methodological design of the self-report questionnaire (i.e. limited comparability of procedural scales due to different number of respondents (n) in each procedural scale as a function of respondents' previous answers on the declarative scales, see above), the analysis is limited to the examination of the four declarative knowledge scales corresponding with the context-specific SRL competence components assessed by the instrument. Moreover, the validity arguments concerning each of the four declarative knowledge scales (metacognition, cognition, regulation of frustration, regulation of boredom) can only be analyzed separately due to specific qualities of the validation study on which this study is based.

Method

Design

This master thesis employs a multi-informant approach. On the one hand, data collected by the developers of IWAN in a pilot study among 408 students was used. On the other hand, a quantitative expert rating with an adapted version of IWAN was conducted among 11 renowned scholars in German-speaking SRL research. In each of the two cross-sectional studies, respondents were informed that the ethical standards for anonymity and confidentiality are ensured in all steps of the research process.

Pilot Study

Sample and procedure.

The sample of the pilot study consisted of 408 German-speaking students (224 female, 93 male, 91 not stated; age: *M*= 25.36 years, *SD* = 4.88), 175 (43%) of which were inscribed at the department of Psychology and 233 (57%) at the department of economic sciences of the University of Vienna. Of the 408 students, 226 (55%) were studying in the Bachelor program, 89 (22%) in the Master program of one of the two fields of study, 93 (23%) studied Psychology in the Diploma program. It was intended to examine whether students of different disciplines differ in their SRL competencies. For

this study aims to focus on the basic assumptions of the instrument, the examination of these assumptions was realized with the overall sample.

The sample was acquired via formal email requests to participate in the study in the winter semester of 2015 addressing students of both faculties. The e-mail contained a link to the online (here: internet-based) questionnaire created in Unipark (www.unipark.de). In courses and lectures, the invitation to participate in the study was emphasized.

Due to specific qualities of the pilot study´s design (for reasons of reasonableness, the entire instrument was segmented and the elements were allocated to sub-samples), none of the 408 students participated in all of the four elements of the self-report questionnaire corresponding to the four relevant situations (exam preparation, writing a minor scientific work, writing a major scientific work, lecture attendance). However, in some cases a combination of two elements was presented to the students, which is why the total number of participants responding to each the element surpasses the overall sample size (N=408). In sum, the sample that reported knowledge and use of strategies in exam preparation (cognitive strategies) consisted of 144 students, in writing a minor scientific work (metacognitive strategies) of 158 students, in writing a major scientific work (regulation of frustration) of 76; and lecture attendance (regulation of boredom) of 109 respondents.

Measures.

Context-specific competencies in SRL.

To assess students' competencies in SRL, IWAN was used. Students were instructed to report their learning strategy knowledge and strategy use in four relevant situations (exam preparation, writing a minor scientific work, writing a major scientific work, lecture attendance). As this master thesis is a validation study, IWAN was described above in the connection with its underlying theoretical and conceptional framework (see also Appendix A and B).

Academic Achievement.

In order to assess learning outcomes, students were asked to report objective and subjective indicators for academic achievement. *Objective indicators* were the mean of the grades in the last three courses (regardless of course type) and the grade in the relevant learning situation. The grades in the last three courses (regardless of course type), as well as the grade in the relevant situation, were assessed on a scale that ranges from 1 to 5 whereas '1' reflects the best and '5' the worst grade. *Subjective indicators* consisted of a subjective estimation of the own learning achievement in the relevant situation and of a subjective estimation of the content with the own learning achievement in the relevant situation. In exam preparation, for instance, the latter refers to the degree of satisfaction students experience when evaluating their knowledge gain. Subjective indicators were both assessed via a visual analogue scales ranging from 1 to 101 ('1' represent the worst possible, and '101' the best possible estimate).

Quantitative Expert Rating

Sample and procedure.

The sample size of the quantitative expert rating consisted of 11 renowned scholars in SRL research in German-speaking countries. The contact was realized via e-mail acquisition that consisted of a formal request to participate in the study, of an introduction in how to participate in the expert rating, as well as information about the basic assumptions of the instrument. The e-mail further contained a link to the online (here: internet-based) questionnaire created in Unipark (www.unipark.de). The standardized e-mail request can be found in Appendix D.

Measures.

Adequacy of the scoring system.

In order to obtain a rating of the presumed quality (beneficial/adverse) of the named self-regulatory strategies in the given situation, the design of the self-report questionnaire described above was slightly adjusted. First, the instruction and tailored

vignettes were modified in terms of a specification of what the scales were supposed to measure (e.g. metacognitive strategies in the writing a minor scientific work). Second, the basic rationale of the instrument was reduced to the question: *What is, in your opinion, beneficial in the given learning situation* (e.g. exam preparation)*?* The use of respective self-regulatory strategies by the experts was considered irrelevant for the rating of the adequacy of the scoring system. Consequently, each item consisted of a statement that required an answer in dichotomous response format ("yes" or "no") that refers to abbreviation of the question (*Is this beneficial?)* mentioned above.

Estimation of the impact of SRL strategies on the success in mastering a relevant learning situation.

In order to obtain an estimate of the presumed impact of the use of specific learning strategies on the success in mastering a relevant learning situation, the experts were asked to estimate the extent to which the self-regulatory strategies rated to be beneficial/adverse may foster/thwart a successful mastery of the relevant learning situation. Therefore, for every strategy an expert rated as beneficial in the given situation, the expert was asked to estimate the potential to foster a successful mastery of the learning situation and for every strategy rated as adverse in the given situation, the experts were instructed to estimate its potential to thwart the successful mastery of the given situation. Thus, the estimation of the impact of a strategy can vary across experts depending on the previous rating of the strategy as beneficial/adverse. For the estimation of the impact of each strategy, a visual analogue scales ranging from 1 to 101 ('1' represents 'not at all beneficial/adverse' and '101' 'very beneficial/adverse) was used. The data collected in these scales were not used in this validation study.

Statistical analysis

Generally, the statistical analysis performed in this study was – if procurable - computed with SPSS-22. Thus, descriptive statistics, item analysis, correlative analysis were executed in SPSS. Anyway, the inter-rater-reliability and composite reliability could not be computed using SPSS. Thus, some information about the theoretical foundation

and actual computations of inter-rater-reliability and composite reliability performed in the course of this master thesis will be provided hereafter.

Inter-rater-reliability.

For determining the reliability of agreement between a fixed number of raters on binary or nominal scale ratings, Fleiss´ kappa (henceforth: κ; Fleiss, 1971) can be used. It can be regarded as the proportion of agreement beyond that expected by chance, or, in other words, "the *achieved* beyond-chance agreement as a proportion of the *possible* beyond-chance agreement" (Sim & Wright, 2005, p. 258; italics in original). It can be calculated as follows:

$$\kappa = \frac{\sum_{i=1}^{N}\sum_{j=1}^{k} n_{ij}^2 - Nn[1+(n-1)\sum_{j=1}^{k} p_j^2]}{Nn(n-1)(1-\sum_{j=1}^{k} p_j^2)} \quad (1)$$

with *N* representing the total number of subjects (here: items), *n* the number of ratings per subject, p_j the overall rates of assignment to every subject, and *k* the number of categories into which assignments are made (Fleiss, 1971). In order to infer that the overall agreement in the assignment to the *k* categories is significantly greater than chance, one can calculate κ/*SE*(κ), which is, by the central limit theorem, approximately distributed as a standard normal variate (Fleiss, 1971). The value of κ/S*E*(κ) can then be compared to a threshold (1,96, which is the critical *z*-value of the normal distribution at α= .05). The standard error *SE*(κ) can be calculated as follows:

$$SE(\kappa) = \sqrt{\frac{2}{Nn(n-1)} \times \frac{\sum_{j}^{k} p_j^2 - (2n-3)(\sum_{j}^{k} p_j^2)^2 + 2(n-2)\sum_{j=1}^{k} p_j^3}{(1-\sum_{j=1}^{k} p_j^2)^2}} \quad (2)$$

Anyway, when it can be assumed a priori that raters will agree significantly more than just by chance, Sim and Wright (2005) recommend to test the significance of κ against a value representing a minimum acceptable level of agreement. Therefore, a confidence interval (CI) around the determined κ is constructed in order to test whether the plausible values of κ lie above the acceptable threshold. The two-sided 95% confidence interval can be calculated as: CI(κ*)* = κ ± (1.96 × *SE*(κ)), with 1.96 being the *z*-score corresponding to a 95% 2-sided confidence level (Sim & Wright, 2005).

In order to interpret κ, Landis and Koch (1977) provided a classification of κ which specifies values below cero as indicators of 'poor agreement', vales from .01 to .20 as 'slight agreement', from .21 to .40 as 'fair agreement', from .41 to .60 as 'moderate agreement', from .61 to .80 as 'substantial agreement' and from .81 to 1.00 as 'almost perfect agreement'. Thus, given that κ tends to be higher when there are fewer categories (Sim & Wright, 2005) and that the expert rating used binary scales, the "critical value" against the obtained κ was tested in this study was set to κ = .61. Setting this threshold was deemed reasonable for the appropriateness of the scoring inference (i.e. the beneficial/adverse impact of each strategy) was tested. The statistical analysis to determine Fleiss´ κ as a means of rater-agreement in this study was performed in Windows Excel by applying equations 1 and 2 described above.

Item difficulties.

In order to calculate the difficulties of items on a binary scale, the number of people answering the item correctly is divided by the total number of people answering the item (Mummendey & Grau, 2014). Hence, the term *item difficulty* can be misleading for it actually indicates the percentage of participants who answered the item correctly. The optimal item difficulty (i.e. the maximum information) a dichotomous item can have is $p = 0.5$. Conventionally, items with *p*s lower than .20 and greater than .80 are excluded from the scale or revised (Mummendey & Grau, 2014). However, as Leopold and Leutner (2002) remarked, self-report questionnaires rather draw on learners' recognition of strategies than on their declarative strategy knowledge, it is assumed to be likely that the respondents recognize the quality (beneficial/adverse) of the named self-regulatory strategies in the relevant situation on some of the dichotomous items rather easily. Thus, in this study, items with item difficulties between $p = .10$ and $p = .90$ are regarded as sufficiently informative.

Coefficients of internal consistency.

In order to analyze the reliability as a means of internal consistence, coefficient alpha (also known as Cronbach's α) is the most commonly used measure (Graham, 2006). However, coefficient alpha is based on a measurement model that requires some

assumptions for the estimate to accurately reflect the true reliability of a scale (Raykov, 1997). In sum, the application of coefficient alpha requires that "all items measure the same latent trait on the same scale, with the only variance unique to an item being comprised wholly of error" (Graham, 2006, p. 935) or, in other words, that "all items load on a single underlying construct and (…) all items represent that construct equally well" (Geldhof, Preacher, & Zyphur, 2014, p. 73). A violation of these assumptions (e.g. When estimating the reliability of scales which aim to measure a construct, i.e. a competence, that is assumed to be multidimensional; see Blömeke et al., 2015) can cause coefficients alpha to underestimate the true reliability of the scale (Miller, 1995, as cited in Graham, 2006). A restrictive reliability model such as Cronbach's α applied for congeneric scales (i.e. scales in which heterogeneous item-construct relations are plausible or expected; see Raykov, 1997) may be inappropriate to estimate the scale´s reliability. However, it must be noted that reliability does not provide a measure of test scale unidimensionality; instead, it assumes that unidimensionality exists (Miller, 1995). That is, coefficients of internal consistency are based on the assumption of one underlying factor. Hence, an alternative to Cronbach's α is ω (Raykov, 1997). It can be calculated from factor loadings of a confirmatory factor analysis (CFA) that allows for heterogeneous correlations between indicators (i.e. items) and their underlying latent construct (i.e. heterogeneous factor loadings; cf. Geldhof et al., 2014).

A congeneric scale with a standardized latent construct (with the variance fixed to 1), *composite reliability* (ω) can be calculated as

$$\omega = \frac{\left(\sum_{i=1}^{k} \lambda_i\right)^2}{\left(\sum_{i=1}^{k} \lambda_i\right)^2 + \sum_{i=1}^{k} \theta_{ii}} \quad (3)$$

with λ_i representing the factor loading of item *i* onto a single common factor and θ_{ii} representing the unique variance of item *i* (Geldhof et al., 2014). In order to calculate the overall amount of variance in the items accounted for by the latent construct (VE, variance extracted), Hair (2006) proposes the following equation:

$$VE = \frac{\sum_{i=1}^{k} \lambda_i^2}{n_k} \quad (4)$$

Hair (2006) further recommended setting the threshold of composite reliability above .70 and an extracted variance (VE) above .50 in order to assume with reasonable confidence that the validity of the construct has been attained. However, each SRL component (metacognition, cognition, regulation of frustration, regulation of boredom) was assessed by a variety of items in the declarative scales, each of which represent one single item of the sub-strategies of the components (which were measured by a variety of items on the corresponding procedural scales). Consequently, the scales on the declarative level aim to assess the components on a more general level, whereas the scales of the procedural level aim to measure one sub-strategy at a time. Hence, given that reliability assumes that unidimensionality exists (Miller, 1995), it was considered rather likely that the coefficients of internal consistency for the scales on the declarative level were modest as a function of the heterogeneity of items (i.e. sub-strategies), which each SRL components is consisting of. The assumption of a possible multidimensionality of the declarative scales is also derived from the results coming conventional SRL questionnaires like the LIST, for which it was shown that the sub-strategies of metacognition (e.g. planning, monitoring) or of cognition (e.g. elaboration, rehearsal) constitute distinct factors (Boerner et al., 2005). Accordingly, in this study, the threshold for composite reliability (ω) was set to .60 and for VE to .40.

Thus, in order to calculate the internal consistency of the scales, composite reliability (ω) was identified for the four scales (cognitive strategies in exam preparation; metacognitive strategies in writing a minor scientific work; regulation of frustration in writing a major scientific work; regulation of boredom in lecture attendance). To estimate the standardized factor loadings required for determining composite reliability (ω), the CFA was computed via latent variable modeling (cf. Graham, 2006) in MPlus version 7.3 (Muthén & Muthen, 2015). Therefore, the manifest variables (items) were assumed to be ordered categorical (here: binary) and the parameters were estimated via Maximum-Likelihood-Estimation that is robust to non-normality (MLR; cf. Muthén & Muthen, 2015). The calculation of composite reliability (ω) was performed in Windows Excel by applying equations 3 and 4 described above.

Correlations.

In order to examine the relationship between the situation-specific competence scores and academic achievement (objective and subjective), correlation analyses were performed. Correlation coefficients (*r*) are standardized measures of an observed effect and range from -1 to + 1. Values of r ± .1 indicate a small effect, ± .3 a medium effect and ± .5 a large effect (cf. Field, 2013). For some of the variables included in the correlation analyses cannot be treated as interval-scaled (objective indicators of learning achievement) and some are significantly (α= .05) skewed (subjective indicators of learning achievement), Spearman's r_s was used as a correlation coefficient. In detail, the mean grade in the last three courses, as well as the grade across the last three grades in the specific learning situation can be regarded ordinal scaled (ranging from 1 to 5). The variables 'subjective estimation of the own learning achievement' and 'subjective estimation of the content with the own learning achievement' were, in all four scales, significantly skewed (α= .05). Whether or not the skewness/kurtosis of a scale is significant is particularly relevant for the adequate interpretation of the significance of the Pearson's correlation coefficient *r*. When the skewness and/or the kurtosis of a variable in the correlation analysis is significant, it is reasonable to use Spearman's correlation coefficient r_s (Field, 2013). In order to examine the significance of the skewness, the skewness'/kurtosis' *z*-score is compared against what would be expected to be obtained by chance alone. Thus, an absolute value of the skewness' *z*-score greater than 1.96 or 2.33 (which are the critical values of the normal distribution at α= .05 and α= .01) is significant (cf. Field, 2013). The skewness' *z*-scores, as well as the means, standard deviations and skewness of each variable in each sample can be found in Tables 6-9.

Results

Inter-rater-reliability

The analysis of the inter-rater agreement from the expert rating resulted in a Fleiss κ = .73 (*SE*(κ) = 0.023), with an overall number of items across learning strategies and corresponding situations of *N* = 35, with an overall number of ratings per item of *n* = 11 and the number of categories of *k* = 2. Thus, the examination of whether the obtained

rater agreement is significantly greater than chance resulted in $\kappa/SE(\kappa) = .73/.023 = 23.10$. Given that 23.10 > 1,96, the obtained rater-agreement is significantly greater than chance ($\alpha = .05$).

The 2-sided 95% confidence interval CI(κ) = [.69 - .78] indicate that a 'substantial agreement' as proposed by Landis and Koch (1977) could be attained. The predefined "critical value" against the obtained κ was tested in this study was set to κ = .61, which lies below the obtained CI(κ). Thus, the obtained value of κ = .73 significantly lies above the critical value ($\alpha = 0.05$). The table of frequencies of ratings per item and category drawn from the expert rating can be found in Appendix E.

Item analysis

In order to calculate the item difficulties of each scale, the number of people answering the item correctly was divided by the total number of people answering the item (cf. Mummendey & Grau, 2014). The item difficulties of each scale (strategy knowledge scale, strategy use scale, multiplicative competence scale; see intended scoring system of the instrument) will be presented in the following chapters. For this master thesis is written in English language, the original items were analogously translated from German into English for the purpose of maintaining common understanding. Thus, the items presented in Tables 1 to 4 do not represent the original items that were used in the studies. The original items in German language can be found in Appendix A and B.

Scales for Metacognitive strategies in writing a minor scientific work.

In Table 1, the item difficulties of the three scales (knowledge scale, use scale, multiplicative competence scale) for 'Metacognitive strategies in writing a minor scientific work' are shown. The item difficulties of the knowledge scale range from $p = .51$ to $p = .99$ ($\bar{p} = .87$) and the use scale shows item difficulties from $p = .54$ to $p = .92$ ($\bar{p} = .73$). The results indicate that all items in the initial scales (knowledge scale and use scale) show item difficulties above the optimum of $p = .50$. Thus, each item could consistently be answered correctly by at least more than 50% of the participants.

Table 1

Item difficulties: Metacognitive strategies in writing a minor scientific work

Items	Initial scales		Multiplicative competence scale
	Knowledge	Use	
	p	*p*	*p*
Before I begin to work, I set a timetable.	.96	.70	.68
I take things as they come.	.82	.54	.51
I don't think a lot about the work before I begin.	.89	.69	.66
I think in advance about which recourses (e.g. learning material, help, etc.) I need for my work.	.99	.85	.85
I plan in advance about how to proceed best in writing the scientific work.	.96	.77	.75
I don‘t set timetables because eventually, I won‘t adhere to it anyway.	.94	.73	.70
My only goal is to hand in the work.	.80	.58	.54
During writing I regularly ask myself if my proceeding is useful.	.86	.76	.69
I regularly check if I achieve what I had intended.	.96	.70	.68
During working I count on my initial timetable to	.51	.54	.35
When I notice that I don't proceed, I change my	.96	.85	.84
When I notice that my goals are unrealistic, I reconsider them.	.96	.92	.89
I stick to my initial goals by all means.	.63	.78	.58
Before writing I consider which goals I have for the work.	.97	.75	.73

Note. N= 158. Item difficulty (*p*) = number of people answering the item correctly divided by the total number of people answering item (cf. Mummendey & Grau, 2014).

Furthermore, eight items of the knowledge scale (p = .94 - .99) and one item of the use scale (p = .92) show item difficulties above the defined critical value of p = .90. However, the instrument aims to assess SRL competences as a combination of strategy knowledge and strategy use. In the multiplicative competence scale, the range of item difficulties (p = .35 - .89; $\bar{p}$ = .67) lies within the predefined threshold of .10 > p < .90. Results further indicate that only one item on the multiplicative competence scale shows an item difficulty (p = .35) below p = .50.

Scales for cognitive strategies in exam preparation.

In Table 2, the item difficulties of the three scales (knowledge scale, use scale, multiplicative competence scale) for ‘Cognitive strategies in exam preparation’ are shown. The item difficulties of the knowledge scale range from p = .55 to p = .98 ($\bar{p}$ =

Table 2

Item difficulties: Cognitive strategies in exam preparation

	Initial scales		Multiplicative competence scale
	Knowledge	use	
Items	*p*	*p*	*p*
I structure the study matter.	.95	.78	.76
I follow up the subject matter thoroughly.	.92	.55	.53
I start with learning directly. Everything else will follow.	.69	.72	.52
I repeat the subject matter.	.98	.87	.85
I study a summary.	.55	.40	.31
I read through the lecture notes only once.	.90	.74	.69
I learn the subject matter by heart.	.79	.73	.65

Note. N= 144. Item difficulty (*p*) = number of people answering the item correctly divided by the total number of people answering item (Mummendey & Grau, 2014).

.83) and the use scale shows item difficulties from $p = .40$ to $p = .87$ ($\overline{p} = .68$). The results indicate that only one item of the use scale ($p = .40$) shows an item difficulty below the optimum of $p = .50$. All other items in the initial scales (knowledge scale and use scale) show item difficulties ($p = .55$ - .98) above the optimum of $p = .50$. Thus, all items but one could consistently be answered correctly by at least more than 50% of the participants. Furthermore, three items of the knowledge scale ($p = .92$ - .98) show item difficulties above the defined critical value of $p = .90$. However, the instrument aims to assess SRL competences as a combination of strategy knowledge and strategy use. In the multiplicative competence scale, the range of item difficulties ($p = .31$ - .85; $\overline{p} = .62$) lies within the defined threshold of $.10 > p < .90$.

Scales for regulation of boredom in lecture attendance.

In Table 3, the item difficulties of the three scales (knowledge scale, use scale, multiplicative competence scale) for ‘Regulation of boredom in lecture attendance’ are shown. The item difficulties of the knowledge scale range from $p = .37$ to $p = .95$ ($\overline{p} = .79$) and the use scale shows item difficulties from $p = .17$ to $p = .73$ ($\overline{p} = .41$). The results display that one item of the knowledge scale ($p = .37$) shows an item difficulty below $p = .50$. Three items of the knowledge scale ($p = .91$ - .95) show item difficulties above the

Table 3

Item difficulties: Regulation of boredom lecture attendance

Items	Initial scales		Multiplicative competence scale
	Knowledge	Use	
	p	*p*	*p*
I make myself aware of the benefits I can extract from the lecture.	.95	.73	.72
I prepare for the lecture.	.86	.17	.17
I make myself aware that as a student, I have to attend to lectures.	.37	.42	.28
I engage actively in the lecture.	.91	.35	.34
I distract myself during the lecture.	.95	.41	.39
I go to lectures without high expectations.	.72	.39	.35

Note. N= 109. Item difficulty (*p*) = number of people answering the item correctly devided by the total number of people answering item (cf. Mummendey & Grau, 2014).

predefined critical value of $p = .90$. All items on the use scale but one, on the other hand, show item difficulties ($p = .17$ - .42) below $p = 0.50$. Thus, it can be assumed that a large majority of the items could consistently be answered correctly by less than 50% of the participants. Furthermore, three items of the knowledge scale ($p = .92$ - .98) show item difficulties above the defined critical value of $p = .90$. In the multiplicative competence scale, the range of item difficulties ($p = .17$ - .72; $\overline{p} = .37$) lies within the predefined threshold of $.10 > p < .9$.

Scales for regulation of frustration in writing a major scientific work.

In Table 4, item difficulties of the three scales (knowledge scale, use scale, multiplicative competence scale) for 'Regulation of frustration in writing a major scientific work' are shown. The item difficulties of the knowledge scale range from $p = .39$ to $p = .93$ ($\overline{p} = .68$) and the use scale shows item difficulties from $p = .26$ to $p = .86$ ($\overline{p} = .51$). The results display that two items of the knowledge scale ($p = .92$ - .93) show item difficulties above the predefined critical value of $p = .90$. Furthermore, three items of the knowledge scale ($p = .92$ - .98) show item difficulties above the defined critical value of $p = .90$. In the use scale ($p = .26$ - $p = .86$), and in the multiplicative competence scale ($p = .17$ - .72; $\overline{p} = .45$), the range of item difficulties lies within the predefined threshold of $.10 > p < .9$.

Table 4

Item difficulties: Regulation of frustration in writing a major scientific work

	Initial scales		Multiplicative competence scale
	Knowledge	Use	
Items	*p*	*p*	*p*
I set a timetable.	.87	.51	.49
I consider what I need to work well.	.93	.86	.84
I make myself aware of how much is depending on the work.	.41	.30	.21
I set realistic aspirations.	.92	.68	.67
I postpone the work again and again.	.88	.51	.49
I take things as they come.	.64	.47	.42
I have confidence in mastering the task without difficulties.	.39	.50	.29
I make myself aware of the fact that I don't have a choice other than being tenacious.	.39	.26	.18

Note. N= 76. Item difficulty (*p*) = number of people answering the item correctly divided by the total number of people answering item (cf. Mummendey & Grau, 2014).

Coefficients of internal consistency

As a means of the internal consistency, composite reliability (ω) was calculated for the four scales (cognitive strategies in exam preparation; metacognitive strategies in writing a minor scientific work; regulation of frustration in writing a major scientific work; regulation of boredom in lecture attendance). Because of a negative λ (-.245), the item "I learn the subject matter by heart." from the scale assessing cognitive strategies in

Table 5

Coefficients of internal consistency

Scales	*N*	*k*	*CR* (ω)	*VE*
Cognitive Strategies in exam preparation	144	6[a]	.67[a]	.30[a]
Metacognitive strategies in writing a minor scientific work	158	14	.66	.16
Regulation of frustration in writing a major scientific work	76	8	.68	.24
Regulation of Boredom in lecture attendance.	109	6	.62	.25

Note. CR (ω) = composite reliability (see equation 3); *VE* = variance extracted (see equation 4); *N*= sample size; *k* = number of items per scale.
[a]Because of a negative λ (-.245), the item "I learn the subject matter by heart." was excluded from the analysis (see Appendix F for standardized factor loadings).

exam preparation was excluded from any further analysis (see Appendix F for standardized factor loadings). As Table 5 shows, the coefficients of composite reliability ω for the four scales range from .62 to .68, and thus lie above the predefined critical value of .60. The extracted variances (*VE*s) range from .16 to .30, all of which are below the acceptable value of .40 set as a threshold.

Relationship between SRL competencies and academic achievement

Statistical characteristics of variables used in correlation analyses.

Metacognitive competencies and academic achievement in writing a minor scientific work.

As Table 6 illustrates, the mean sum score of the multiplicative competence scale for metacognitive strategies in writing a minor scientific work is $M = 9.44$ ($SD = 2.47$) with a possible range from 0 and 14 (as a function of the number of items). The average mean grade in the last three courses of this sample ($N = 140$) is $M = 2.36$ ($SD = 0.81$) and the mean grade in writing a minor scientific work is $M = 1.71$ ($SD = 0.82$), both with a possible range from 1 to 5 (1 being the best grade, 5 the worst). Concerning the grade in writing a minor scientific work, the *z*-value of the skewness indicates a significantly skewed distribution ($z_s = 3.97$, $p < .01$). The positive value indicates a build-up of low grades (cf. Field, 2013). Also, the distribution of the estimates of subjective learning achievement in writing a scientific work is significantly and negatively skewed ($z_s = -3.15$, $p < .01$) with a mean of $M = 75.24$ ($SD = 18.16$). The skewness' *z*-score ($z_s = -3.15$, $p < .01$) of the distribution of the estimates of the content with the learning achievement in writing a scientific work ($M = 75.24$; $SD = 18.16$) is significant and negative, indicating a build-up of high estimates (cf. Field, 2013).

Table 6

Statistical characteristics of variables used in correlation analyses between metacognitive competencies in writing a minor scientific work and academic achievement

Variables	*M*	*SD*	Range	*S*	SE_S	*K*	SE_K	z_s	z_k
Multiplicative competence score Metacognition	9.44	2.47	0-14	-.20	.19	-.58	.38	-1.01	-1.52
Mean grade in the last three courses	2.36	0.81	1-5	.40	.21	-.46	.41	1.93	-1.12
Grade in writing of a minor scientific work	1.71	0.82	1-5	-.81	.21	-.31	.41	3.97**	-0.75
Subjective learning achievement in writing a minor scientific work	75.24	18.16	1-101	-.65	.21	.21	.41	-3.15**	0.52
Content with learning achievement in writing a minor scientific work	73.30	23.61	1-101	-.77	.21	-.12	.41	-3.76**	-0.28

Note. $N = 140$; S = skewness; SE_S = Standard error of skewness; K = kurtosis; SE_K = standard error of kurtosis; z_s= standardized skewness = Skewness/SE_S; z_k= standardized kurtosis = Kurtosis/SE_K. **$p < .01$.

Cognitive competencies and academic achievement in exam preparation.

As Table 7 illustrates, the mean sum score of the multiplicative competence scale for cognitive strategies in exam preparation is $M = 3.67$ ($SD = 1.42$) with a possible range from 0 and 6 (as a function of the number of items). The significant *z*-value of the skewness ($z_s = -2.00$, $p < .05$) indicates a build up in high scores (cf. Field, 2013). In this sample ($N = 129$), the average mean grade in the last three courses is $M = 2.48$ ($SD = 0.82$) and the mean exam grade is $M = 2.54$ ($SD = 1.21$), both with a possible range from 1 to 5 (1 being the best grade, 5 the worst). Concerning the average mean grade in in the last three courses, the negative *z*-value of the kurtosis ($z_s = -2.32$, $p < .05$) is significant, indicating a flat and light-tailed distribution (cf. Field, 2013). Considering the exam grades, the *z*-value of the skewness indicates a significantly skewed distribution ($z_s = 2.04$, $p < .05$) with a build-up of low grades (cf. Field, 2013). Also, the distribution of the estimates of subjective learning achievement in exam preparation is significantly and negatively skewed ($z_s = -3.69$, $p < .01$) with a mean of $M = 65.33$ ($SD = 19.29$). Also, the skewness' *z*-score ($z_s = -3.29$, $p < .01$) of the distribution of the estimates of the content

Table 7

Statistical characteristics of variables used in correlation analyses between competencies in boredom regulation in lecture attendance and academic achievement

Variables	*M*	*SD*	Range	*S*	SE_S	*K*	SE_K	z_s	z_k
Multiplicative competence score Regulation of Boredom	2.25	1.41	0-6	.58	.23	-.32	.46	2.52**	-0.70
Mean grade in the last three courses	2.46	0.82	1-5	.41	.25	-.62	.49	1.63	-1.26
Grade in exam	2.48	0.96	1-5	-.42	.24	-.31	.47	1.78	-0.67
Subjective knowledge gain in lecture attendance	64.27	19.45	1-101	-.57	.25	-.05	.49	-2.29*	-0.10
Content with knowledge gain in lecture attendance	62.19	25.31	1-101	-.53	.25	-.55	.49	-2.14*	-1.12

Note. $N = 94$; S = skewness; SE_S = Standard error of skewness; K = kurtosis; SE_K = standard error of kurtosis; z_s= standardized skewness = Skewness/SE_S; z_k= standardized kurtosis = Kurtosis/SE_K. *$p < .05$. **$p < .01$.

with the learning achievement in exam preparation ($M = 66.57$; $SD = 27.19$) is significant and negative, indicating a build-up of high estimates (cf. Field, 2013).

Boredom regulation competencies and academic achievement in lecture attendance.

As shown in Table 8, the mean sum score of the multiplicative competence scale for strategies of boredom regulation in lecture attendance is $M = 2.25$ ($SD = 1.41$) with a possible range from 0 and 6 (as a function of the number of items). The significant *z*-value of the skewness ($z_s = 2.52$, $p < .01$) indicates a build up in low scores (cf. Field, 2013). The average mean grade in the last three courses of this sample ($N = 94$) is $M = 2.46$ ($SD = 0.82$) and the mean exam grade is $M = 2.48$ ($SD = 0.96$) with a possible range from 1 to 5 (‘1’ being the best grade, ‘5’ the worst). Also, the distribution of the estimates of subjective knowledge gain in lecture attendance is significantly and negatively skewed ($z_s = -2.29$, $p < .05$) with a mean of $M = 64.27$ ($SD = 19.45$). Also, the skewness’ *z*-score ($z_s = -2.14$, $p < .05$) of the distribution of the estimates in regards to the content with the subjective knowledge gain in lecture attendance ($M = 62.19$; $SD = 25.31$) is significant and negative, indicating a build-up of high estimates (cf. Field, 2013).

Table 8

Statistical characteristics of variables used in correlation analyses between competencies in frustration regulation in writing a major scientific work and academic achievement

Variables	*M*	*SD*	Range	*S*	SE_S	*K*	SE_K	z_s	z_k
Multiplicative competence score Regulation of Frustration	3.59	1.72	0-8	-.10	.28	-.50	.55	-0.35	-0.92
Mean grade in the last three courses	2.26	0.87	1-5	.71	.31	-.17	.62	-1.91	-0.28
Grade in writing of a major scientific work	1.77	0.71	1-5	-.62	.29	.17	.57	2.15*	0.30
Subjective learning achievement in writing a major scientific work	73.99	17.41	1-101	-.68	.29	.55	.57	-2.37**	0.96
Content with learning achievement in writing a major scientific work	76.09	21.68	1-101	-.76	.29	.12	.57	-2.62**	0.20

Note. $N = 69$; S = skewness; SE_S = Standard error of skewness; K = kurtosis; SE_K = standard error of kurtosis; z_s= standardized skewness = Skewness/SE_S; z_k= standardized kurtosis = Kurtosis/SE_K. *$p < .05$. **$p < .01$.

Frustration regulation competencies and academic achievement in writing a major scientific work.

As illustrated in Table 9, the mean sum score of the multiplicative competence scale for strategies of frustration regulation in writing a major scientific work is $M = 3.59$ ($SD = 1.72$) with a possible range from 0 and 8. The average mean grade in the last three courses of this sample ($N = 129$) is $M = 2.26$ ($SD = 0.87$) and the mean grade on the major scientific work is $M = 1.77$ ($SD = 0.71$), both with a possible range from 1 to 5 (‘1’ being the best grade, ‘5’ the worst). Concerning the grade on the major scientific work, the significant *z*-value of the skewness ($z_s = 2.15$, $p < .05$) indicates a build-up of low grades (cf. Field, 2013). Also, the distribution of the estimates of subjective learning achievement in writing a scientific work is significantly and negatively skewed ($z_s = -2.37$, $p < .01$) with a mean of $M = 73.99$ ($SD = 17.41$). Also, the skewness’ *z*-score ($z_s = -2.62$, $p < .01$) of the distribution of the estimates regarding the content with the learning achievement in writing a major scientific work ($M = 76.09$; $SD = 21.68$) is significant and negative, indicating a build-up of high estimates (cf. Field, 2013).

Table 9

Statistical characteristics of variables used in correlation analyses between cognitive competencies in exam preparation and academic achievement

Variables	*M*	*SD*	Range	*S*	SE_S	*K*	SE_K	z_s	z_k
Multiplicative competence score	3.67	1.42	0-6	-0.40	0.20	-.32	.40	-2.00*	-0.80
Mean grade in the last three courses	2.48	0.82	1-5	-0.01	0.22	-.99	.43	-0.04	-2.32*
Grade in exam	2.54	1.21	1-5	-0.43	0.21	-.72	.42	2.04*	-1.70
Subjective learning achievement in exam preparation	65.33	19.29	1-101	-0.79	0.21	.34	.42	-3.69**	0.80
Content with learning achievement in exam preparation	66.57	27.19	1-101	-0.70	0.21	-.38	.42	-3.29**	-0.90

Note. $N = 129$; S = skewness; SE_S = Standard error of skewness; K = kurtosis; SE_K = standard error of kurtosis; z_s= standardized skewness = Skewness/SE_S; z_k= standardized kurtosis = Kurtosis/SE_K.
*$p < .05$. **$p < .01$.

Correlation analyses.

In order to establish the relationship of the composite scores on each of the situation-specific SRL competence component scales with subjective and objective indicators of academic achievement, Spearman's correlation coefficients were calculated (assumptions for calculating Pearson's correlation coefficient, such as non-significant skewness, were not met; see statistical characteristics of variables used in correlation analyses). For it was assumed that the relationship is directional, the significance tests of the correlations were one-tailed. As Table 10 illustrates, all correlations showed in the expected direction (for low values in the mean grade in the last three courses and the grade in the learning situations indicate good grades, the polarity of the correlation coefficient was reversed). That is, it can be assumed that a modest to moderate positive relationship exists between the scales of the instrument and subjective and objective indicators of academic achievement.

The scale concerning cognitive strategies in exam preparation appears to be related significantly only to subjective indicators of learning achievement. A positive correlative relationship was found between cognitive strategies and subjective learning achievement

Table 10

Correlations between SRL competencies and academic achievement

		objective		subjective	
Scales	N	MG	GRS	CLA	SLA
Cognitive Strategies in exam preparation	129	.08	.16	.23**	.35**
Metacognitive strategies in writing a minor scientific work	140	.31**	.32**	.38**	.38**
Regulation of frustration in writing a major scientific work	76	.18	.06	.24	.33**
Regulation of boredom in lecture attendance.	109	.39**	.28**	.39**	.22*

Note. N= sample size; MG = Mean grade in the last three courses; GRS = Grade in relevant learning situation; CLA = Content with learning achievement in relevant learning situation; SLA = subjective learning achievement in relevant learning situation.
*p < .05. ** p < .01.

(SLA) in exam preparation (r_s = .23, p < .01) and content with the learning achievement (CLA) in exam preparation (r_s = .35, p < .01).

The competence component scale concerning metacognitive learning strategies in writing a minor scientific work appears to be positively and significantly related to all indicators of learning achievement (r_s = .32 - r_s = .38, p < .01), indicating a moderate effect. In detail, metacognitive learning strategies in writing a minor scientific work appear to be positively related to the mean grade in the last three courses (MG) with a correlation coefficient of r_s = .31 (p < .01), to the grade in the relevant situation (GRS; here: writing a minor scientific work) with r_s = .32 (p < .01) and CLA (r_s = .38, p < .01). Moreover, a moderate correlation with SLA in the writing of a minor scientific work could be found (r_s = .38, p < .01). Scores on the scale assessing frustration of regulation in writing a major scientific work seem to be related significantly only to SLA with a correlation coefficient r_s = .33 (p < .01) that indicates a moderate effect.

Scores on the scale concerning the regulation of boredom in lecture attendance is significantly correlated to all indicators of learning achievement. The regulation of boredom in lecture attendance seems to be in a moderate and positive relationship with MG (r_s = .39, p < .01), GRS (r_s = .28 p < .01) and CLA (r_s = .39, p < .01). Moreover, a modest effect with SLA could be found (r_s = .22, p < .05).

Discussion

The aim of this master thesis was to examine first empirical evidence for the validity argument of a new self-report questionnaire approach that aims to assess context-specific competence components for SRL in higher education. Therefore, some of the weakest assumptions and inferences about the scoring, the generalization and the extrapolation were tested (cf. Kane, 1992, 2013a; see also Cook et al., 2015). Favorable evidence for the scoring inference and the extrapolation inference was found. The results of this study further indicate that some assumptions concerning the scoring inference and the generalization inference may not be fully adequate. It must be acknowledged that this master thesis was limited to the examination of the scales on the declarative level. Thus, major components of the instrument (e.g. the scales on the procedural level) cannot be subject of discussion in this study.

Adequacy of the scoring inference. Given that no human error took place, it was assumed that the (technical) application of the scoring system (e.g. the calculation of a total score across items) was direct and flawless (cf. Hopfenbeck & Maul, 2011).

Support for the adequacy of the scoring system was provided by item analyses. Results showed that the item difficulties of all multiplicative competence scales of IWAN lie in-between the predefined thresholds of $.10 > p < .90$. Thus, the multiplicative competence scales were considered sufficiently informative.

Furthermore, the adequacy of the scoring criteria was supported by the data coming from the quantitative expert rating. For the rationale of IWAN is that self-regulatory strategies can either foster or thwart the successful mastery of a given learning situation (i.e. answers can either be right or wrong), a "substantial agreement" (cf. Landis & Koch, 1977) among renowned SRL researchers was considered necessary to assume with reasonable confidence that SRL strategies represented by the items can – without doubt - be classified into beneficial or adverse strategies. The overall agreement among renowned experts in SRL research about the presumed impact of SRL strategies in a given learning situation ($\kappa = .73$; CI = [.69 - .78]) fell significantly above the predefined threshold of $\kappa = .61$. Thus, it can be assumed that experts in SRL research substantially agree in the presumed classification of beneficial and adverse strategies.

The scoring system implies that the more beneficial strategies a respondent knows to be beneficial and the more beneficial strategies he or she uses, as well as the more adverse strategies a respondent assumes to be adverse and the fewer adverse strategies he or she uses, the better is the score. Thus, another concern about the adequacy of the scoring inference may come from the fact that the scoring system applies quantitative standards (Wirth & Leutner, 2008), whereas no existing theoretical model underpins this maximum view (cf. Leopold, den Elzen-Rump, & Leutner, 2007). However, the scale for metacognitive strategies, for instance, aims to assess a variety of different strategies (e.g. setting goals, setting timetables, monitor progress, etc.), each of which were assumed to be of high relevance for the mastery of writing a minor scientific work in a preliminary study (Dresel et al., 2015). Consequently, it could be argued that the scales rather assess the declarative knowledge of situation-specific requirements rather than global knowledge about SRL strategies. Thus, the quantitative standards applied in the scoring system may be more justified than in global SRL questionnaires. Moreover, although Wirth and Leutner (2008) reason that quantitative standards could lead to reduced validity, the chances for new insights on SRL opened by large-scale assessments (Schellings & Van Hout-Wolters, 2011) may compensate for the reduced validity resulting from the application of quantitative standards.

In sum, the findings from item analysis and from the expert rating support the assumptions about the adequacy of the scoring inference. However, considering the concerns about the quantitative standards applied, further evidence is necessary to conclude that the adequacy of the scoring inference is given.

However, there are some limitations originating from the method of the expert rating to be acknowledged. For the expert rating consisted of a reduced form of IWAN (only strategy knowledge scales and no strategy use scales were used), it is conceivable that this might not be an appropriate way of examining the adequacy of the scoring inference. The method only provides a rating of the items of each scale as being beneficial or adverse for the successful mastery of a given situation. However, the expert rating does not provide information about how well the items represent beneficial/adverse SRL strategies, nor does it offer differentiated feedback about why experts assume a strategy to be beneficial or adverse. Moreover, it may be that some experts have responded in

accordance with the more plausible alternative (beneficial/adverse). Thus, the limited informative value of the obtained results in regards to the adequacy of the scoring criteria must be acknowledged.

Adequacy of the generalization inference. In order to examine whether responses to the items of each scale indicate an underlying factor (e.g. metacognitive competence), coefficients of internal consistency have been calculated. Results showed that the obtained values of composite reliability in all four declarative scales (ω = .62 - .68) of IWAN lie above the predefined threshold (ω = .60) necessary to assume with reasonable confidence that the validity of a construct has been attained. As discussed above, the critical value was set to ω = .60 instead of ω = .70 as proposed by Hair (2006) because it was considered plausible that the scales on the declarative level may not be one-dimensional. Anyhow, although composite reliability (ω) allows for items to contribute to a single factor with different precision (Graham, 2006), it still assumes that unidimensionality exists (Miller, 1995). Thus, the attained coefficients of internal consistency indicate that the scales on the declarative level – albeit conceivably multidimensional – are composed of items that may contribute substantially to one underlying factor (e.g. metacognitive competence). Moreover, given that binary scales have per se limited variability (cf. Cohen, 1983), and considering that composite reliability (ω) is a variance-based coefficient (cf. Graham, 2006), the obtained values may as well even underestimate the "true reliability" (i.e. that the items represent one underlying construct).

Thus, results from the reliability analyses showed decent coefficients of internal consistency, which supports the assumption of adequate generalization inference. Further empirical evidence and/or additional measures to evaluate the construct validity of the scales are needed, though, to examine whether the items of IWAN can be considered representative for the universe of possible similar items.

Adequacy of the extrapolation inference. Results of the correlation analyses in this master thesis indicate that the sum scores on each SRL component assessed by IWAN are in many regards modestly and/or moderately related to objective (r_s = .38 - .39) and subjective estimates of academic achievement (r_s = .22 - .39). In detail, more significant relationships could be found with subjective than with objective estimates of learning

achievement. The results of the correlation analyses indicate a moderate relationship between the scales of IWAN and academic achievement, which is in line with the study of Pintrich and his colleagues' (1993), which found modest to moderate correlations (r = .13 - .42) between MSLQ learning strategy scales and final course grades. Similarly, correlations between learning achievement and metacognitive learning strategies (r = .20), as well as resource management strategies (r = .30) for the LIST could be found, albeit only for sub-strategies "monitoring" and "effort", respectively (Schiefele, Streblow, Ermgassen, & Moschner, 2003).

Given that the adequacy of the extrapolation inference depends on whether the universe of possible questionnaire items can be assumed to be reasonably and/or empirically related to the constructs that the scales aim to measure, the results support the assumption about the adequacy of the extrapolation inference.

Anyhow, there are some limitations originated from the measures and methods used in the correlation analyses. Arguably, the use of grades as estimates of learning achievement may not be appropriate since the validity of grades as a predictor for learning achievement has been subject of much controversy (see Tight, Mok, Huisman, & Morphew, 2009). Indications for eventual limited informational value of grades can be found in the present study. Results suggest that it might be more difficult to get a good grade on an exam (M = 2.54/2.48, SD = 1.21/0.96) than on a major/minor scientific work (M = 1.77/1.71, SD = 0.71/0.82). Anyhow, all estimates of learning achievement used in the correlation analyses of the present study have been self-reported. Considering the fact that students might be slightly overconfident about their achievement (cf. Winne & Jamieson-Noel, 2002), self-reports on their achievement may not be a reliable source for academic achievement. This is particularly relevant for the subjective estimates of learning achievement used in this study for the significant negative skewnesses of each estimate in all samples indicate a build up in high estimates (cf. Field, 2013).

Another source of concern about the validity of the obtained correlation coefficients is the fact that possible effects of moderating or confounding variables (e.g. motivation; cf. Efklides, 2011) on the relationship between the scale's sum scores and academic achievement have not been controlled or partialed out. Thus, the obtained

relationship between SRL components assessed by IWAN and academic achievement might be inaccurate.

Furthermore, the instruments design embodies several innovative features (e.g. declarative and procedural knowledge scales; reasons for the non-use of strategies, etc.), which are not subject of discussion in this master thesis. However, there is reason to believe that one innovation in particular may be strongly related to the results reported above, namely the response format that asks students to answer to questions about one statement on two parallel binary scales (i.e. strategy knowledge scale and strategy use scale, respectively). Although the assumptions underlying this new design have not been tested in this study, it will be discussed subsequently if specific qualities in the instruments design may affect the processes involved in answering the items. In detail, concerns about these features particularly reflect the question about whether students' responses to the scales in IWAN can be taken as indicative of their real-world learning behaviors (cf. Hopfenbeck & Maul, 2011).

A major validity concern about self-report questionnaires in SRL research was expressed by Leopold and Leutner (2002), who criticized that these approaches drew on learners' recognition of strategies rather than on their declarative strategy knowledge. It has been argued that this may be due to the fact that items on these scales are usually pre-formulated, and thus possible knowledge deficits may be concealed (cf. Spörer & Brunstein, 2006). Hence, a student may not actually have a specific strategy in mind when confronted with a learning situation that may requires using this strategy, but when faced with an item, he or she mighty recognize that a strategy *is very unlikely* not to be adequate in the given learning situation (production deficit; cf. Hasselhorn, 1995). Accordingly, a closer look at the characteristics of the items in the initial scales (strategy knowledge, strategy use) of IWAN revealed considerable asymmetries in item difficulties, whereas particularly items in the strategy knowledge scales exceeded the predefined threshold. These results indicate that a substantial number of items on the strategy knowledge scales may be too easy, or, in other words, may not provide sufficient information about students' knowledge of strategies. However, it may be that high item difficulties in the strategy knowledge scales were also originated from the combination of relevant learning situations and corresponding learning strategies, whereas students may recognize

beneficial/adverse strategies easier than when asked if a specific strategy is beneficial/adverse in general. Thus, the binary scales used in IWAN may have encouraged students to rather respond in favor of the more plausible alternative (beneficial/adverse) in doubt about the impact of the proposed strategies in a learning situation. Consequently, it seems reasonable that IWAN cannot overcome the validity concerns in regards to the informative value self-report questionnaires in SRL research can provide about student's "actual" strategy knowledge. In fact, it must be questioned if self-report questionnaire scales can be the means of choice to assess declarative knowledge at all (Leopold & Leutner, 2002). Most of the concerns (e.g. recognition of strategies rather than knowledge) could not be overcome yet (cf. Wirth & Leutner, 2008). Still, as Samuelstuen and Bråten (2007) pointed out, students may report to use a certain strategy in conventional questionnaires because they believe or know the strategy to be effective although they do not actually use them. Thus, it must be acknowledged that IWAN represents an approach that aims to incorporate this critique by asking students to report both their strategy knowledge and their strategy use.

However, the design of two parallel scales might have an impact on the process of answering the items. On the one hand, the distinction between strategy knowledge and strategy use could lead to students feeling encouraged to report their actual (non-) use of the strategies since they can show that they do know better. On the other hand, one could argue that respondents may become aware of possible discrepancies between knowing what would be beneficial to do and what they actually do in a learning situation. Consequently, an argument based loosely on Cognitive Dissonance Theory (cf. Festinger, 1962) could be that the parallel scales might lead to responses that are based on the reduction of cognitive dissonance rather than on the report of student's actual real-life learning behavior. Thus, "actual" discrepancies between strategy knowledge and strategy use could be more prominent than what results found in this study indicate. Additionally, there is reason to believe that respondents may differ in their choice of reference points when asked about the (non-) use of a strategy. As Veenman and his colleagues (2003) have indicated, students may not base their answers only on their own individual standards when asked about their strategy use, which they assumed to be a validity concern inherent in the method of self-report questionnaires. Strategy use is assessed by IWAN in form of the question *Do you do this?*. In line with Veenman and his colleagues

(2003), a respondent thus may answer this question by referring to what he or she actually does or what he or she assumes he or she should be doing. However, the question *Do you do this?* may also lead to distorted answers for it does not specify what is meant by *doing this*. It is possible to understand the question in many different ways. One respondent may refer to it as "*Do you use this strategy frequently?".* Another respondent might base his or her answer rather on the reading of the question: "*Have you ever used this strategy?"* However, it seems plausible that some respondents choose to negate the question because they assume it is not appropriate to report that they use the strategy when they only have applied it a couple of times. Other respondents may affirm the same question in accordance with their memory of having used the strategy once. Thus, there is reason to believe that respondents may use different reference points and different readings to answer the question *Do you do this?*.

Moreover, asking students about which strategies they use without giving a concrete context may encourage respondents to generalize their actions across situations (cf. Winne & Perry, 2000). IWAN aims to acknowledge the situation-specificity of SRL. For the items serve to maintain or reinstate the psychological contexts in which self-regulatory processes occurred (cf. Richardson, 2004), responses to scales that imply contextualized situations may be more valid. Consequently, it seems plausible that the items presented in IWAN may trigger more vivid memories about the actual use of learning strategies in a given situation than a global learning strategy questionnaire would do (cf. Samuelstuen & Bråten, 2007).

In sum, this master thesis provided both empirical evidence, as well as theoretical and practical arguments that support, as well as challenge the claim that the instrument measures what it purports to measure (i.e. context-specific competence components in SRL as a function of strategy knowledge and strategy use). It can be assumed, though, that IWAN constitutes a self-report questionnaire approach assessing SRL that aims to acknowledge the validity concerns regarding conventional approaches by posing an alternative assessment method to the large body of self-report questionnaire approaches. However, it may very well be that introducing innovative designs in response to justified concerns about the validity of conventional measures could entail further challenges only to be overcome with unconventional measures. Thus, this master thesis can only constitute the gambit in the investigation of the assumptions and claims underlying the

instrument. Therefore, it is considered essential to further investigate the adequacy of the scoring inference, of the generalization inference and of the extrapolation inference, as well as the implication inference.

Implications

Considering the limitations of the empirical evidence provided in this study, and the validity concerns presented above, it is deemed expedient to further investigate basic assumptions of this instrument. Certainly, the implications that can be drawn from this study are pre-eminently limited to the issues discussed in this thesis. One also want to keep in mind that the instrument is still in developmental phase, thus also modifications in the instrument's design will be proposed.

Arguably, the most critical issue concerning the design of the instrument may be the vagueness that is implied in the question *Do you do this?* for assessing strategy use. Although no empirical evidence was provided to augment the arguments about eventual inconsistencies in students' response processes that may result from this question, it is considered indispensable to specify what *doing this* is meant to mean. Thus, it seems reasonable either to add more information to the tailored vignette in order to specify what is to be assessed (e.g. frequency of use, consistency of use) and continue using the dichotomous response format, or to specify strategy use by offering a more differentiated scale (e.g. Likert-type scales or visual analogue scales) that may provoke more distinct/interpretable information about students' strategy use. It may be instrumental to evaluate and compare different modes of measurement and decide upon the design that best suits the purpose (i.e. the measurement design that provides the most accurate/plausible measure of the constructs under consideration). Therefore, it seems reasonable to investigate the processes involved in giving answers to the parallel scales. This could be realized by means of think-aloud protocols (cf. Hopfenbeck & Maul, 2011) or semi-standardized interviews among a reasonable number of students. Thereby, it may be investigated how the items are interpreted, if knowledge scales conceal knowledge deficits (production deficit; cf. Hasselhorn, 1995) and whether or not the binary scales may facilitate this concealment, or, which reference points respondents use to report their strategy knowledge and strategy use. Additionally, and most importantly, students could

be asked to report how they respond to the parallel scales when faced with possible discrepancies between strategy knowledge and strategy use.

Moreover, it may be necessary to apply measures that allow for multidimensionality in the scales other than reliability coefficients in order to provide convincing evidence that the responses to the items indicate an underlying, yet multidimensional construct. Additionally, the execution of standard procedures in the validation of a self-report questionnaire, such as the analysis of the factor structure or the determination of the retest-reliability, requires further empirical studies.

Moreover, regarding the limitations of the expert rating used in this study, it is considered adjuvant for the instrument development to conduct differentiated and differentiating expert interviews that focus on how well the items represent the construct to be measured and how well the items represent beneficial or adverse strategies.

Furthermore, in order to explore the relationship between the scales of IWAN and learning achievement, longitudinal studies that both monitor SRL by means of IWAN and imply a differentiated assessment of learning achievement (e.g. an elaborate scoring rubric for the quality of a major/minor scientific work) could augment the correlative relationships found in this study.

Last but not least, it must be acknowledged that the instrument is part of a larger project to assess competences for successful self-regulated learning in higher education. Therefore, once sufficient empirical evidence was provided in support of the basic assumptions of the instrument, assessment triangulation with the other instruments of the project that apply qualitative offline and online stands could bring further information about the validity argument of IWAN (cf. Pellegrino et al., 2001). Moreover, given that in this master thesis, the proposed use of the instrument was specified as the contribution to a comprehensive understanding of successful SRL in higher education, the adequacy of the implication inference can only be tested within the joint research project PRO-SRL. Additionally, it is conceivable that once the validity argument of the revised IWAN has been tested and evidence consistently supports its assumptions and claims, IWAN could be also used in academic research in order to examine competencies in SRL, especially for large-scale assessments. Furthermore, it may be used in teaching as a diagnostic instrument to detect eventual deficits in the SRL competencies of students and/or to derive tailored interventional trainings.

The project PRO-SRL

The questionnaire was developed within the project PRO-SRL (Product- and process-oriented modeling and assessment of competencies in self-regulated learning (SRL) in higher education [Produkt- und prozessorientierte Modellierung und Erfassung von Kompetenzen zum Selbstregulierten Lernen (SRL) im tertiären Bereich], which was funded by the initiative KOKOHs (Kompetenzmodellierung und Kompetenzerfassung im Hochschulsektor [Modelling and Assessment of Competencies in Higher Education]) by the German Federal Ministry of Education and Research (BMBF, Bundesministerium für Bildung und Forschung).

The collaborative project PRO-SRL consists of four subprojects of universities in Augsburg, Darmstadt, Erlangen-Nurenberg und Vienna, each of which aim to assess different aspects of the model proposed by (Dresel et al., 2015). The project leadership of the subproject Vienna was held by Univ.-Prof. Dipl.-Psych. Dr. Barbara Schober and Univ.-Prof. Dr. Dr. Christiane Spiel.

References

Artelt, C. (2000). Wie prädiktiv sind retrospektive Selbstberichte über den Gebrauch von Lernstrategien für strategisches Lernen [How predictive are self-reported strategies for their actual use?]. *Zeitschrift Für Pädagogische Psychologie*, *14*(2/3), 72–84. doi.org/10.1024//1010-0652.14.23.72.

Azevedo, R. (2009). Theoretical, conceptual, methodological, and instructional issues in research on metacognition and self-regulated learning: A discussion. *Metacognition and Learning*, *4*(1), 87–95. doi.org/10.1007/s11409-009-9035-7

Baumert, J. & Köller, O. (1996). Lernstrategien und schulische Leistung [Learning strategies and school achievement]. In J. Möller & O. Köller (Hrsg.), *Emotionen, Kognitionen und Schulleistung* (S. 137 – 154). Weinheim: Psychologie Verlags Union.

Blömeke, S., Gustafsson, J.-E., & Shavelson, R. J. (2015). Beyond Dichotomies. Competence Viewed as a Continuum. *Zeitschrift Für Psychologie*, *223*(1), 3–13. doi.org/10.1027/2151-2604/a000194

Boekaerts, M. (1997). Self-regulated learning: A new concept embraced by researchers, policy makers, educators, teachers, and students. *Learning and Instruction*, *7*(2), 161–186. doi.org/10.1016/S0959-4752(96)00015-1

Boekaerts, M. (1999). Self-regulated learning: where we are today. *International Journal of Educational Research*, *31*(6), 445–457. doi.org/10.1016/S0883-0355(99)00014-2

Boekaerts, M. (2007). What have we learned about the link between motivation and learning/performance? *Zeitschrift für Pädagogische Psychologie*, *21*(3), 263–269.

Boekaerts, M., & Corno, L. (2005). Self-regulation in the classroom: A perspective on assessment and intervention. *Applied Psychology: An International Review*, *54*(2), 199–231.

Boekaerts, M., Pintrich, P. R., & Zeidner, M. (Eds.). (2000). *Handbook of self-regulation*. San Diego: Academic Press.

Boerner, S., Seeber, G., Keller, H., & Beinborn, P. (2005). Lernstrategien und Lernerfolg im Studium: Zur Validierung des LIST bei berufstätigen Studierenden [Learning strategies and success: Validation of the LIST with part-time students]. *Zeitschrift Für Entwicklungspsychologie Und Pädagogische Psychologie*, *37*(1), 17–26. doi.org/10.1026/0049-8637.37.1.17

Borsboom, D., Mellenbergh, G. J., & van Heerden, J. (2004). The Concept of Validity. *Psychological Review*, *111*(4), 1061–1071. doi.org/10.1037/0033-295X.111.4.1061

Bråten, I., & Samuelstuen, M. S. (2007). Measuring strategic processing: comparing task-specific self-reports to traces. *Metacognition and Learning*, *2*(1), 1–20. doi.org/10.1007/s11409-007-9004-y

Cohen, J. (1983). The cost of dichotomization. *Applied Psychological Measurement*, *7*(3), 249–253. doi.org/10.1177/014662168300700301

Cook, D. A., Brydges, R., Ginsburg, S., & Hatala, R. (2015). A contemporary approach to validity arguments: a practical guide to Kane's framework. *Medical Education*, *49*(6), 560–575. doi.org/10.1111/medu.12678

Cromley, J., & Azevedo, R. (2011). Measuring strategy use in context with multiple-choice items. *Metacognition and Learning*, *6*(2), 155–177. doi.org/10.1007/s11409-011-9070-z

Dinsmore, D. L., Alexander, P. A., & Loughlin, S. M. (2008). Focusing the Conceptual Lens on Metacognition, Self-regulation, and Self-regulated Learning. *Educational Psychology Review*, *20*(4), 391–409. doi.org/10.1007/s10648-008-9083-6

Dresel, M., & Haugwitz, M. (2006). The relationship between cognitive abilities and self- regulated learning: evidence for interactions with academic self- concept and gender. *High Ability Studies*, *16*(2), 201–218. doi.org/10.1080/13598130600618066

Dresel, M., Schmitz, B., Schober, B., Spiel, C., Ziegler, A., Engelschalk, T., … Steuer, G. (2015). Competencies for successful self-regulated learning in higher education: structural model and indications drawn from expert interviews. *Studies in Higher Education*, *40*(3), 454–470. doi.org/10.1080/03075079.2015.1004236

Efklides, A. (2011). Interactions of metacognition with motivation and affect in self-regulated learning: The MASRL model. *Educational Psychologist*, *46*(1), 6–25. doi.org/dx.doi.org/10.1080/00461520.2011.538645

Engelschalk, T., Steuer, G., & Dresel, M. (2014, September). *Kompetenzen zur Motivationsregulation bei Studierenden: Erfassung mittels Situational Judgement Test als gangbarer Weg?* [Competencies in motivation regulation: Assessment via situationional judgement test as feasable way?]. Paper presented at the 49th DGPs Congress, Bochum, Germany.

Festinger, L. (1962). *A Theory of Cognitive Dissonance*. Stanford University Press.

Field, A. P. (2013). *Discovering statistics using IBM SPSS statistics*. London: SAGE.

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive–developmental inquiry. *American Psychologist*, *34*(10), 906–911. doi.org/10.1037/0003-066X.34.10.906

Fleiss, J. L. (1971). Measuring nominal scale agreement among many raters. *Psychological Bulletin*, *76*(5), 378–382. doi.org/0.1037/h0031619

Floden, R. E. (1981). The Logic of Information-Processing Psychology in Education. *Review of Research in Education*, *9*, 75–109. doi.org/10.2307/1167183

Geldhof, J. G., Preacher, K. J., & Zyphur, M. J. (2014). Reliability estimation in a multilevel confirmatory factor analysis framework. *Psychological Methods*, *19*(1), 72–91. doi.org/10.1037/a0032138

Graham, J. M. (2006). Congeneric and (Essentially) Tau-Equivalent Estimates of Score Reliability What They Are and How to Use Them. *Educational and Psychological Measurement*, *66*(6), 930–944. doi.org/10.1177/0013164406288165

Hair, J. F. (2006). *Multivariate data analysis* (6. ed.). Upper Saddle River, NJ: Pearson Prentice-Hall.

Hampel, P., Petermann, F., & Dickow, B. (2001). *SVF-KJ. Stressverarbeitungsfragebogen von Janke und Erdmann angepasst für Kinder und Jugendliche* [SVF-KJ. Coping Questionnaire for children and adolescents]. Göttingen: Hogrefe.

Hartig, J., Frey, A., & Jude, N. (2008). Validität [Validity]. In H. Moosbrugger & A. Kelava (Eds.), *Testtheorie und Fragebogenkonstruktion* (pp. 135–163). Springer Berlin Heidelberg. Retrieved from link.springer.com/chapter/10.1007/978-3-540-71635-8_7

Hasselhorn, M. (1995). Beyond production deficiency and utilization inefficiency: Mechanisms of the emergence of strategic categorization in episodic memory tasks. In F. E. Weinert & W. Schneider (Eds.), *Memoryperformance and competencies issues in growth and development*. Mahwah, NJ: Erlbaum.

Heckhausen, H., & Gollwitzer, P. M. (1987). Thought contents and cognitive functioning in motivational versus volitional states of mind. *Motivation and Emotion*, *11*(2), 101–120. doi.org/10.1007/BF00992338

Hopfenbeck, T. N., & Maul, A. (2011). Examining Evidence for the Validity of PISA Learning Strategy Scales Based on Student Response Processes. *International Journal of Testing*, *11*(2), 95–121. doi.org/10.1080/15305058.2010.529977

Jamieson-Noel, D., & Winne, P. H. (2003). Comparing Self-Reports to Traces of Studying Behavior as Representations of Students' Studying and Achievement. *Zeitschrift für Pädagogische Psychologie*, *17*(3/4), 159–171. doi:10.1024/1010-0652.17.3.159

Jöstl, G., Klug, J., Pichler, M., Schober, B., & Spiel, C. (2014, September). *Messung der Kompetenzen zum Selbstregulierten Lernen im Studium – Die Auswahl passender Strategien in relevanten Anforderungssituationen und deren Zusammenhang mit Studienerfolg* [Assessment of competencies in self-regulated learning in higher education - the selection of adequate strategies in relevant learning situations and their relationship with academic achievement]. Paper presented at the 49th DGPs Congress, Bochum, Germany.

Jöstl, G., Klug, J., Schober, B., & Spiel, C. (2015, March). *Wie viel Wissen haben Studierende über SRL und warum wenden sie dieses nicht an? Ein Ansatz zur Aufklärung widersprüchlicher Be- funde* [How much knowledge do students have about SRL and why don't they apply it? An approach to contradictory findings]. Paper presented at the 3rd GEBF Congress, Bochum, Germany.

Kane, M. T. (1992). An argument-based approach to validity. *Psychological Bulletin*, *112*(3), 527–535. doi.org/10.1037/0033-2909.112.3.527

Kane, M. T. (2013a). The argument-based approach to validation. *School Psychology Review*, *42*(4), 448–457.

Kane, M. T. (2013b). Validating the interpretations and uses of test scores. *Journal of Educational Measurement*, *50*(1), 1–73. doi.org/0.1111/jedm.12000

Kane, M. T. (2013c). Validation as a pragmatic, scientific activity. *Journal of Educational Measurement*, *50*(1), 115–122. doi.org/10.1111/jedm.12007

Kelley, T. L. (1927). *Interpretation of educational measurements*. Oxford, England: World Book Company.

Koeppen, K., Hartig, J., Klieme, E., & Leutner, D. (2008). Current Issues in Competence Modeling and Assessment. *Zeitschrift Für Psychologie / Journal of Psychology*, *216*(2), 61–73. doi.org/10.1027/0044-3409.216.2.61

Landis, J. R., & Koch, G. G. (1977). The Measurement of Observer Agreement for Categorical Data. *Biometrics*, *33*(1), 159–174. doi.org/10.2307/2529310

Leopold, C., den Elzen-Rump, V., & Leutner, D. (2007). Self-regulated learning from science texts. In M. Prenzel (Ed.), *Studies on the educational quality of schools. The final report on the DFG Priority Program* (pp. 221–238). Münster: Waxmann.

Leopold, C., & Leutner, D. (2002). Der Einsatz von Lernstrategien in einer konkreten Lernsituation bei Schülern unterschiedlicher Jahrgangsstufen [The use of learning strategies in a concrete learning situation among students from different grades]. *Zeitschrift Für Pädagogik*, *45*(Beiheft), 240–258.

Messick, S. (1995). Validity of psychological assessment: validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist*, *50*(9), 741.

Miller, M. B. (1995). Coefficient alpha: A basic introduction from the perspectives of classical test theory and structural equation modeling. *Structural Equation Modeling: A Multidisciplinary Journal*, *2*(3), 255–273. doi.org/10.1080/10705519509540013

Mummendey, H. D., & Grau, I. (2014). *Die Fragebogen-Methode: Grundlagen und Anwendung in Persönlichkeits-, Einstellungs- und Selbstkonzeptforschung* [The questionnaire approach: Basics and application in personality -, attitude - and self-concept research]. Hogrefe Verlag.

Muthén, L. K., & Muthen, B. O. (2015). *Mplus User's Guide. Seventh Edition.* Los Angeles, CA: Muthén & Muthén.

Panadero, E., Klug, J., & Järvelä, S. (2015). Third wave of measurement in the self-regulated learning field: when measurement and intervention come hand in hand. *Scandinavian Journal of Educational Research*, 1–13. doi.org/10.1080/00313831.2015.1066436

Paris, S. G., Lipson, M. Y., & Wixson, K. K. (1983). Becoming a strategic reader. *Contemporary Educational Psychology*, *8*(3), 293–316. doi.org/10.1016/0361-476X(83)90018-8

Pellegrino, J. W., Chudowsky, N., & Glaser, R. (Eds.). (2001). *Knowing what students know the science and design of educational assessment*. Washington, DC: The National Academy Press. Retrieved from site.ebrary.com/id/10032455

Peverly, S. T., Brobst, K. E., Graham, M., & Shaw, R. (2003). College adults are not good at self-regulation: A study on the relationship of self-regulation, note taking, and test taking. *Journal of Educational Psychology*, *95*(2), 335–346. doi.org/10.1037/0022-0663.95.2.335

Pintrich, P. R. (1999). The role of motivation in promoting and sustaining self-regulated learning. *International Journal of Educational Research*, *31*(6), 459–470. doi.org/10.1016/S0883-0355(99)00015-4

Pintrich, P. R. (2000). The role of goal orientation in self-regulated learning. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 451–502). San Diego, CA, US: Academic Press. Retrieved from dx.doi.org/10.1016/B978-012109890-2/50043-3

Pintrich, P. R., Smith, D. A. F., Garcia, T., & Mckeachie, W. J. (1993). Reliability and predictive validity of the motivated strategies for learning questionnaire (MSLQ). *Educational and Psychological Measurement*, *53*(3), 801–813. doi.org/10.1177/0013164493053003024

Puustinen, M., & Pulkkinen, L. (2001). Models of Self-regulated Learning: A review. *Scandinavian Journal of Educational Research*, *45*(3), 269–286. doi.org/10.1080/00313830120074206

Raykov, T. (1997). Estimation of Composite Reliability for Congeneric Measures. *Applied Psychological Measurement*, *21*(2), 173–184. doi.org/10.1177/01466216970212006

Richardson, J. T. E. (2004). Methodological Issues in Questionnaire-Based Research on Student Learning in Higher Education. *Educational Psychology Review*, *16*(4), 347–358. doi.org/10.1007/s10648-004-0004-z

Robbins, S. B., Lauver, K., Le, H., Davis, D., Langley, R., & Carlstrom, A. (2004). Do Psychosocial and Study Skill Factors Predict College Outcomes? A Meta-Analysis. *Psychological Bulletin*, *130*(2), 261–288. doi.org/10.1037/0033-2909.130.2.261

Roth, A., Ogrin, S., & Schmitz, B. (2015). Assessing self-regulated learning in higher education: a systematic literature review of self-report instruments. *Educational Assessment, Evaluation and Accountability*, 1–26. doi.org/10.1007/s11092-015-9229-2

Roth, A., & Schmitz, B. (2013, March). *Eine Herausforderung: Die valide Erfassung von selbstreguliertem Lernen mit Selbstberichtsverfahren* [A challenge: The valid assessment of self-regulated learning with self-report measures]. Paper presented at the 1st GEBF Congress, Kiel, Germany.

Samuelstuen, M. S. (2003). Psychometric properties and item-keying direction effects for the Learning and Study Strategies Inventory-High School Version with norwegian students. *Educational and Psychological Measurement*, *63*(3), 430–445.

Samuelstuen, M. S., & Bråten, I. (2007). Examining the validity of self-reports on scales measuring students' strategic processing. *British Journal of Educational Psychology*, *77*(2), 351–378. doi.org/10.1348/000709906X106147

Schellings, G. (2011). Applying learning strategy questionnaires: problems and possibilities. *Metacognition and Learning*, *6*(2), 91–109. doi.org/10.1007/s11409-011-9069-5

Schellings, G., & Van Hout-Wolters, B. (2011). Measuring strategy use with self-report instruments: theoretical and empirical considerations. *Metacognition and Learning*, *6*(2), 83–90. doi.org/10.1007/s11409-011-9081-9

Schiefele, U., Streblow, L., Ermgassen, U., & Moschner, B. (2003). Lernmotivation und Lernstrategien als Bedingungen der Studienleistung [The Influence of Learning Motivation and Learning Strategies on College Achievement]. *Zeitschrift Für Pädagogische Psychologie*, *17*(3/4), 185–198. doi.org/10.1024//1010-0652.17.3.185

Schmitz, B. (2003). Selbstregulation-Sackgasse oder Weg mit Forschungsperspektive? [Self Regulation: Dead End or a Path with a Research Perspective?]. *Zeitschrift für pädagogische Psychologie, 17*(3), 221–232.

Schmitz, B., & Wiese, B. S. (2006). New perspectives for the evaluation of training sessions in self-regulated learning: Time-series analyses of diary data. *Contemporary Educational Psychology, 31*(1), 64–96. doi.org/10.1016/j.cedpsych.2005.02.002

Schober, B., Klug, J., Jöstl, G., Spiel, C., Dresel, M., Steuer, G., … Ziegler, A. (2015). Gaining substantial new insights into university students' self-regulated learning competencies: How can we succeed? *Zeitschrift Für Psychologie*, *223*(1), 64–65. doi.org/10.1027/2151-2604/a000201

Sim, J., & Wright, C. C. (2005). The Kappa Statistic in Reliability Studies: Use, Interpretation, and Sample Size Requirements. *Physical Therapy*, *85*(3), 257–268.

Sitzmann, T., & Ely, K. (2011). A meta-analysis of self-regulated learning in work-related training and educational attainment: What we know and where we need to go. *Psychological Bulletin*, *137*(3), 421–442. doi.org/10.1037/a0022777

Souvignier, E., & Roes, K. (2005). Lernstrategien und Lernerfolg bei komplexen Leistungsanforderungen: Analysen mit Fragebogen und Lerntagebuch [Learning strategies and performance in complex tasks: Analyses with questionnaires and learning diaries]. In C. Artelt & B. Moschner (Eds.), *Lernstrategien und Metakognition: Implikationen für Forschung und Praxis.* (pp. 65–76). Münster: Waxmann.

Spörer, N., & Brunstein, J. C. (2006). Erfassung selbstregulierten Lernens mit Selbstberichtsverfahren: Ein Überblick zum Stand der Forschung [Assessing self-regulated learning with self-report measures: A state-of-the-art review]. *Zeitschrift für Pädagogische Psychologie*, *20*(3), 147–160. doi.org/10.1024/1010-0652.20.3.147

Tight, M., Mok, K. H., Huisman, J., & Morphew, C. (2009). *The Routledge International Handbook of Higher Education*. Routledge.

Veenman, M. V. J. (2005). The assessment of metacognitive skills: What can be learned from multi-method designs? In C. Artelt & B. Moschner (Eds.), *Lernstrategien und Metakognition: Implikationen für Forschung und Praxis* (pp. 77–99). Münster: Waxmann.

Veenman, M. V. J. (2011). Alternative assessment of strategy use with self-report instruments: a discussion. *Metacognition and Learning*, *6*(2), 205–211. doi.org/10.1007/s11409-011-9080-x

Veenman, M. V. J., Bavelaar, L., De Wolf, L., & Van Haaren, M. G. P. (2014). The on-line assessment of metacognitive skills in a computerized learning environment. *Learning and Individual Differences*, *29*, 123–130. doi.org/10.1016/j.lindif.2013.01.003

Veenman, M. V. J., Prins, F. J., & Verheij, J. (2003). Learning styles: Self-reports versus thinking-aloud measures. *British Journal of Educational Psychology*, *73*(3), 357–372.

Weinstein, C. E., & Hume, L. M. (1998). *Study strategies for lifelong learning*. Washington, DC: American Psychological Association.

Weinstein, C. E., Palmer, D. R., & Schulte, A. C. (1987). *Learning and study strategies inventory (LASSI)*. Clearwater, FL: H & H Publishing.

Wild, K.-P. (2000). *Lernstrategien im Studium Strukturen und Bedingungen* [Learning strategies in undergraduate and graduate education]. Münster: Waxmann.

Wild, K.-P., & Schiefele, U. (1994). Lernstrategien im Studium: Ergebnisse zur Faktorenstruktur und Reliabilität eines neuen Fragebogens [Learning strategies in college education: Results on the factor structure and test reliability of a new questionnaire]. *Zeitschrift Für Differentielle Und Diagnostische Psychologie*, *15*, 185–200.

Winne, P. H. (2010). Improving Measurements of Self-Regulated Learning. *Educational Psychologist*, *45*(4), 267–276. doi.org/10.1080/00461520.2010.517150

Winne, P. H., & Jamieson-Noel, D. (2002). Exploring students' calibration of self reports about study tactics and achievement. *Contemporary Educational Psychology*, *27*(4), 551–572. doi.org/10.1016/S0361-476X(02)00006-1

Winne, P. H., & Perry, N. E. (2000). Measuring self-regulated learning. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of Self-Regulation* (pp. 531–566). San Diego: Academic Press.

Wirth, J. (2004). *Selbstregulation von Lernprozessen* [Self-regulation of learning processes]. Münster: Waxmann Verlag.

Wirth, J., & Leutner, D. (2008). Self-Regulated Learning as a Competence: Implications of Theoretical Models for Assessment Methods. *Zeitschrift Für Psychologie / Journal of Psychology*, *216*(2), 102–110. doi.org/10.1027/0044-3409.216.2.102

Zimmerman, B. J. (2000). Attaining Self-Regulation: A Social Cognitive Perspective. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of Self-Regulation* (pp. 13–39). San Diego: Academic Press. Retrieved from www.sciencedirect.com/science/article/pii/B9780121098902500317

Zimmerman, B. J. (2008). Investigating Self-Regulation and Motivation: Historical Background, Methodological Developments, and Future Prospects. *American Educational Research Journal*, *45*(1), 166–183. doi.org/10.3102/0002831207312909

Zimmerman, B. J., & Schunk, D. H. (2001). *Self-Regulated Learning and Academic Achievement: Theoretical Perspectives*. Routledge.

Zimmerman, B., & Schunk, D. H. (2011). *Handbook of Self-Regulation of Learning and Performance*. Taylor & Francis.

List of Figures

List of Tables

Abstract

This master thesis aims to provide first empirical evidence for the validity argument of a new context-specific self-report questionnaire approach to assess competencies for successful self-regulated learning (SRL) in higher education. The instrument measures SRL by means of two parallel scales per strategy - metacognition, cognition and emotion regulation (frustration, boredom) - in the situations exam preparation, lecture attendance and writing of a minor/major scientific work. The two scales are: (a) knowledge about the adequacy of strategies in given learning situations and (b) reported use of these strategies in the given situation. This master thesis is limited to testing the weakest assumptions underlying the claim that the instrument measures what it purports to measure (i.e. context-specific competence components in SRL as a function of strategy knowledge and strategy use). Results from an expert rating (N = 12) support assumptions about the adequacy of the scoring criteria. Based on data coming from a pilot study (N = 408), item analyses indicate sufficient informative value of the items on the multiplicative competence scales (strategy knowledge times strategy use). Composite reliability coefficients indicate acceptable internal consistency of the scales. Correlation analyses suggest moderate relationships between the scales and objective/subjective estimates of academic achievement and therefore provide evidence for the adequacy of the extrapolation inference. Thus, this study found first empirical evidence supporting the assumptions about the adequacy of the scoring, the generalization and the extrapolation inference. On the basis of these results, limitations and further implications are discussed.

Zusammenfassung

Diese Masterarbeit berichtet erste empirische Befunde hinsichtlich des Validitätsarguments eines neuen kontextspezifischen Selbstberichtsfragebogens zur Erfassung von Kompetenzen zum Selbstregulierten Lernen (SRL) im tertiären Bildungsbereich. Das Instrument misst SRL anhand zweier paralleler Skalen je Strategie – Metakognition, Kognition und Emotionsregulation (Frustration, Langeweile) – in den Situationen Vorbereitung auf eine Prüfung, Vorlesungsbesuch und Schreiben einer kleineren/größeren wissenschaftlichen Arbeit. Die zwei Skalen sind: (a) Wissen um die Angemessenheit dieser Strategien in der gegebenen Situation (b) berichtete Verwendung dieser Strategien in der gegeben Situationen. Die Masterarbeit ist limitiert auf die Überprüfung der schwächsten Annahmen, die der Behauptung zugrunde liegen, das Instrument erfasse, was es vorgibt zu erfassen (kontext-spezifische Kompetenzkomponenten im SRL in Abhängigkeit von Strategiewissen und –anwendung). Die Ergebnisse eines ExpertInnenratings (N = 12) unterstützen die Annahme der Angemessenheit der Verrechnungskriterien. Die auf Basis der Daten einer Pilotstudie (N = 408) durchgeführten Itemanalysen sprechen für eine ausreichende Aussagekraft der Items der multiplikativen Kompetenzskalen (Strategiewissen mal Strategieanwendung). Koeffizienten der Faktorreliabilität sprechen ferner für eine akzeptable innere Konsistenz der Skalen. Korrelationsanalysen weisen auf moderate Zusammenhänge zwischen den Skalen und objektiven/subjektiven Leistungsmaßen hin. Diese Studie berichtet daher erste empirische Befunde für die Angemessenheit der Inferenzen bezüglich der Verrechnung, der Generalisierung und der Extrapolation. Auf Basis der Ergebnisse werden Limitationen und mögliche Implikationen diskutiert.

Appendix A

The self-report questionnaire (IWAN)

Fragebogen

1 Startseite

Herzlich Willkommen zur Online-Befragung von **Pro-SRL**!

Es freut uns, dass Sie an unserer Befragung zum Thema "**Lernen im Studium**" teilnehmen. Sie werden dabei viel über Ihr eigenes Lernverhalten reflektieren und uns in unserer Forschungsarbeit unterstützen.

Pro-SRL ist ein großes internationales Forschungsprojekt, das sich mit den **Lernkompetenzen von Studierenden** befasst. Es geht darum, herauszufinden, mit welchen **Anforderungen** Studierende konfrontiert sind und wie ihr **Lernverhalten** aussieht. Die Ergebnisse sollen dazu dienen, gezielte **Fördermaßnahmen** zu entwickeln. Mit Ihrer Teilnahme (Dauer der Bearbeitung ca. **25 Minuten**) helfen Sie uns dabei, dieses Anliegen zu verwirklichen.

Bei etwaigen Fragen wenden Sie sich bitte an michaela.pichler@univie.ac.at

Und jetzt geht es los!

2 Filterfragen

Ich studiere...

- Wirtschaftswissenschaften
- Psychologie

Ich befinde mich zur Zeit im...

- Bachelorstudium
- Masterstudium
- Diplomstudium 1. Abschnitt
- Diplomstudium 2. Abschnitt

3 Instruktion allgemein

Im Folgenden geht es darum, welche **Strategien für ein erfolgreiches Lernen im Studium** günstig sind. Es werden Fragen zu verschiedenen Situationen gestellt. Bei manchen Bereichen werden wir etwas genauer nachfragen. Es kann daher vorkommen, dass einige Fragen ähnlich klingen.

Die Fragen beziehen sich einerseits darauf, was man eigentlich **machen sollte** und andererseits darauf, was Sie in der konkreten Situation **wirklich machen**. In der Praxis ist es oft nicht möglich, alles so zu machen wie es am besten wäre. Es ist daher ganz normal, wenn Ihre Antworten „Das ist günstig" und „Das mache ich" nicht immer übereinstimmen.

Die Fragen sind jeweils mit **Ja oder Nein** zu beantworten. Falls Sie sich bei einer Frage **nicht sicher** sind, kreuzen Sie bitte an, was **eher** zutrifft.

4 Metakognition Vignette

SCHREIBEN EINER KLEINEREN WISSENSCHAFTLICHEN ARBEIT

Denken Sie an das Erstellen einer **kleineren wissenschaftlichen Arbeit** im Studium. Damit ist die Erstellung eines wissenschaftlichen Textes gemeint, für den selbstständig Literatur recherchiert und aufgearbeitet werden muss und der bis zu einem vorgegebenen Zeitpunkt fertiggestellt werden muss (z.B. Seminararbeit, schriftliche Ausarbeitung etc.).

Eine Befragung von Studierenden hat ergeben, dass es hierbei vor allem wichtig ist, **über das eigene Vorgehen nachzudenken**.

Was ist aus Ihrer Sicht günstig, wenn man eine **kleinere wissenschaftliche Arbeit** schreibt?

Was machen Sie wirklich, wenn Sie eine kleinere wissenschaftliche Arbeit schreiben?

4.1 Metakognition deklarativ

In Bezug auf eine **kleinere wissenschaftliche Arbeit:**

	Das ist günstig.		Ich mache das.	
	Ja	Nein	Ja	Nein
Bevor ich zu arbeiten beginne, überlege ich mir einen Zeitplan.	○	○	○	○
Ich lasse die Dinge auf mich zukommen.	○	○	○	○
Ich denke nicht viel über die Arbeit nach, bevor ich beginne.	○	○	○	○
Ich überlege mir vorab, welche Mittel (z.B. Lernmaterial, Hilfe, etc.) ich für meine Arbeit brauche.	○	○	○	○
Ich plane vorab, wie ich am besten beim Schreiben der wissenschaftlichen Arbeit vorgehen will.	○	○	○	○
Ich mache keine Zeitpläne, da man sich im Endeffekt ja doch nie daran hält.	○	○	○	○
Mein einziges Ziel ist es, die Arbeit abzugeben.	○	○	○	○
Während des Schreibens frage ich mich immer wieder, ob mein Vorgehen gerade sinnvoll ist.	○	○	○	○
Ich überprüfe immer wieder, ob ich das, was ich mir vorgenommen habe, auch erreiche.	○	○	○	○
Während des Arbeitens verlasse ich mich darauf, dass mein ursprünglicher Zeitplan aufgeht.	○	○	○	○
Wenn ich merke, dass ich nicht weiterkomme, ändere ich meine Vorgehensweise.	○	○	○	○
Wenn ich merke, dass meine Ziele unrealistisch sind, überdenke ich sie nochmal.	○	○	○	○
Ich halte unter allen Umständen an meinen ursprünglichen Zielen fest.	○	○	○	○
Ich überlege mir vor dem Schreiben, welche Ziele ich für die Arbeit habe.	○	○	○	○

4.2 Erfolg spezifisch

	nicht genügend	genügend	befriedigend	gut	sehr gut
Wie schätzen Sie Ihre eigene Leistung bei kleinen wissenschaftlichen Arbeiten ein?	○	○	○	○	○

Wie zufrieden waren Sie mit Ihrer letzten kleinen wissenschaftlichen Arbeit?	gar nicht zufrieden		völlig zufrieden

Welche Note haben Sie auf Ihre letzte kleine wissenschaftliche Arbeit bekommen (bzw. auf das Seminar, wo diese Leistung in die Note eingeflossen ist)?

- Nicht Genügend
- Genügend
- Befriedigend
- Gut
- Sehr Gut

5 Kognition Vignette

VORBEREITUNG AUF EINE PRÜFUNG

Denken Sie an eine **Prüfungsvorbereitung**. Dabei müssen Sie große Stoffmengen selbständig aufbereiten und lernen, wobei das Verständnis des Lernstoffs wesentlich ist.

Eine Befragung von Studierenden hat ergeben, dass hierbei vor allem **Lernstrategien** wichtig sind.

Was ist aus Ihrer Sicht günstig bei der **Vorbereitung auf eine Prüfung?**

Was machen Sie wirklich, wenn Sie sich auf eine Prüfung vorbereiten?

5.1 Kognition deklarativ

In Bezug auf die **Vorbereitung auf eine Prüfung:**

	Das ist günstig.		**Ich mache das.**	
	Ja	**Nein**	**Ja**	**Nein**
Ich strukturiere den zu lernenden Stoff.				
Ich beschäftige mich tiefgehend und ausführlich mit dem Stoff.				
Ich fange direkt an zu lernen, der Rest ergibt sich.				
Ich wiederhole den Stoff.				
Ich lerne eine Zusammenfassung.				
Ich lese mir nur einmal das gesamte Skript durch.				
Ich lerne den Stoff auswendig.				

5.2 Erfolg spezifisch

Wie schätzen Sie Ihre eigene Leistung bei Prüfungsvorbereitungen ein?	nicht genügend	genügend	befriedigend	gut	sehr gut

Wie zufrieden waren Sie mit Ihrer letzten Prüfungsvorbereitung?	gar nicht zufrieden	|	völlig zufrieden

Welche Note haben Sie auf Ihre letzte Prüfung bekommen?

- Nicht Genügend
- Genügend
- Befriedigend
- Gut
- Sehr Gut

6 Langeweile Vignette

LANGEWEILE IN EINER VORLESUNG

Denken Sie an den Besuch einer **Vorlesung**. Damit ist eine Lehrveranstaltung mit einer größeren Anzahl von Studierenden gemeint, in der große Stoffmengen durch eine/n Dozierende/n vorgetragen werden.

Eine Befragung von Studierenden hat ergeben, dass hierbei vor allem **Langeweile** ein Problem ist.

Was ist aus Ihrer Sicht günstig, um **Langeweile in einer Vorlesung entgegenzuwirken**?

Was machen Sie wirklich, damit Ihnen in einer Vorlesung nicht langweilig wird?

6.1 Langeweile deklarativ

In Bezug auf **Langeweile in einer Vorlesung**:

	Das ist günstig.		**Ich mache das.**	
	Ja	**Nein**	**Ja**	**Nein**
Ich mache mir bewusst, welchen Nutzen ich aus der Vorlesung ziehen kann.				
Ich bereite mich auf die Vorlesung vor.				
Ich mache mir bewusst, dass ich als Studierende/r Vorlesungen besuchen muss.				
Ich beteilige mich aktiv an der Vorlesung.				
Ich lenke mich während der Vorlesung ab.				
Ich gehe ohne große Erwartungen in die Vorlesung.				

6.2 Erfolg spezifisch

Wie schätzen Sie Ihren Wissensgewinn bei Vorlesungsbesuchen ein?	nicht genügend	genügend	befriedigend	gut	sehr gut
Wie zufrieden waren Sie mit dem erlangten Wissensgewinn in Ihrer letzten besuchten Vorlesung?	gar nicht zufrieden				völlig zufrieden

Welche Note haben Sie auf Ihre letzte Vorlesungsprüfung bekommen, deren Vorlesung Sie besucht haben?

- Nicht Genügend
- Genügend
- Befriedigend
- Gut
- Sehr Gut

7 Frustration Vignette

FRUSTRATION BEI EINER GROSSEN WISSENSCHAFTLICHEN ABSCHLUSSARBEIT

Denken Sie an das Erstellen einer **großen wissenschaftlichen Abschlussarbeit** im Studium (Diplomarbeit, Masterarbeit, eventuell auch Bachelorarbeit). Dazu muss eine umfangreiche Thematik selbstständig über mehrere Monate hinweg unter Betreuung eines/einer Dozierenden bearbeitet werden.

Eine Befragung von Studierenden hat ergeben, dass hierbei vor allem **Frustration** ein Problem ist.

Was ist aus Ihrer Sicht günstig, um beim Erstellen einer **großen wissenschaftlichen Abschlussarbeit** das Entstehen von **Frustration** zu **verhindern**?

Was machen Sie wirklich, um beim Schreiben einer Abschlussarbeit nicht frustriert zu werden?

7.1 Frustration deklarativ

In Bezug auf **Frustration bei einer großen wissenschaftlichen Abschlussarbeit:**

	Das ist günstig.		**Ich mache das.**	
	Ja	**Nein**	**Ja**	**Nein**
Ich mache mir einen Zeitplan.				
Ich überlege mir, was ich brauche, um gut arbeiten zu können.				
Ich mache mir bewusst, wie viel von der Arbeit abhängt.				
Ich stelle an mich selbst realistische Ansprüche.				
Ich verschiebe die Arbeit immer weiter nach hinten.				

Ich lasse die Dinge auf mich zukommen.	○	○	○	○
Ich vertraue darauf, dass mir die Aufgabe keine Schwierigkeiten bereiten wird.	○	○	○	○
Ich mache mir bewusst, dass ich gar keine andere Wahl habe, als mich durchzubeißen.	○	○	○	○

7.2 Erfolg spezifisch

	Nicht Genugend	Genugend	Befriedigend	Gut	Sehr Gut
Wie schätzen Sie Ihre eigene Leistung bei großen wissenschaftlichen Arbeiten ein?	○				○

	gar nicht zufrieden				völlig zufrieden
Wie zufrieden waren Sie mit Ihrer letzten großen wissenschaftlichen Arbeit?	○				○

Welche Note haben Sie auf Ihre letzte große wissenschaftliche Arbeit bekommen (bzw. auf das Seminar, wo diese Leistung in die Note eingeflossen ist)?

- ○ Nicht Genügend
- ○ Genügend
- ○ Befriedigend
- ○ Gut
- ○ Sehr Gut

8 Leistung

Welche Noten haben Sie auf Ihre letzten 3 Prüfungen erhalten?

Nicht Genügend = 5
Genügend = 4
Befriedigend = 3
Gut = 2
Sehr Gut = 1

letzte Lehrveranstaltungsnote ☐ vorletzte Lehrveranstaltungsnote ☐ drittletzte Lehrveranstaltungsnote ☐

9 Infos TeilnehmerInnen

Wir möchten Sie zum Abschluss bitten, einige Angaben zu Ihrer Person zu machen. Aus diesen Angaben ist kein Rückschluss auf Ihre Identität möglich.

Bitte geben Sie Ihr Alter in Jahren an.

☐ Jahre

Sie sind...

○ männlich ○ weiblich

10 Endseite

Appendix B

Codebook: Self-Report Questionnaire (IWAN)

Herzlich Willkommen zur Online-Befragung von **Pro-SRL**!

Es freut uns, dass Sie an unserer Befragung zum Thema "**Lernen im Studium**" teilnehmen. Sie werden dabei viel über Ihr eigenes Lernverhalten reflektieren und uns in unserer Forschungsarbeit unterstützen.

Pro-SRL ist ein großes internationales Forschungsprojekt, das sich mit den **Lernkompetenzen von Studierenden** befasst. Es geht darum, herauszufinden, mit welchen **Anforderungen** Studierende konfrontiert sind und wie ihr **Lernverhalten** aussieht. Die Ergebnisse sollen dazu dienen, gezielte **Fördermaßnahmen** zu entwickeln. Mit Ihrer Teilnahme (Dauer der Bearbeitung ca. **25 Minuten**) helfen Sie uns dabei, dieses Anliegen zu verwirklichen.

Bei etwaigen Fragen wenden Sie sich bitte an michaela.pichler@univie.ac.at

Und jetzt geht es los!

Instruktion allgemein

Im Folgenden geht es darum, welche **Strategien für ein erfolgreiches Lernen im Studium** günstig sind. Es werden Fragen zu verschiedenen Situationen gestellt. Bei manchen Bereichen werden wir etwas genauer nachfragen. Es kann daher vorkommen, dass einige Fragen ähnlich klingen.

Die Fragen beziehen sich einerseits darauf, was man eigentlich **machen sollte** und andererseits darauf, was Sie in der konkreten Situation **wirklich machen**. In der Praxis ist es oft nicht möglich, alles so zu machen wie es am besten wäre. Es ist daher ganz normal, wenn Ihre Antworten „Das ist günstig“ und „Das mache ich“ nicht immer übereinstimmen.

Die Fragen sind jeweils mit **Ja oder Nein** zu beantworten. Falls Sie sich bei einer Frage **nicht sicher** sind, kreuzen Sie bitte an, was **eher** zutrifft.

Item	Antwortformat	Quelle	Variablenname	Variablenlabel (außer = Item)	Wertelabel
Ich studiere…	Auswahl	Eigenkonstruktion	fach	Studium	1 = Wirtschafts-wissenschaften 2 = Psychologie
Ich befinde mich zurzeit im…	Auswahl	Eigenkonstruktion	abschn	Bachelor-Master-Diplom	1 = Bachelorstudium 2 = Masterstudium 3 = Diplomstudium 1. Abschnitt 4 = Diplomstudium 2. Abschnitt

FRUSTRATION BEI EINER GROSSEN WISSENSCHAFTLICHEN ABSCHLUSSARBEIT

Denken Sie an das Erstellen einer **großen wissenschaftlichen Abschlussarbeit** im Studium (Diplomarbeit, Masterarbeit, eventuell auch Bachelorarbeit). Dazu muss eine umfangreiche Thematik selbstständig über mehrere Monate hinweg unter Betreuung eines/einer Dozierenden bearbeitet werden.

Eine Befragung von Studierenden hat ergeben, dass hierbei vor allem **Frustration** ein Problem ist.

Was ist aus Ihrer Sicht günstig, um beim Erstellen einer **großen wissenschaftlichen Abschlussarbeit** das Entstehen von **Frustration** zu **verhindern**?

Was machen Sie wirklich, um beim Schreiben einer Abschlussarbeit nicht frustriert zu werden?

Frustration deklarativ

In Bezug auf **Frustration bei einer großen wissenschaftlichen Abschlussarbeit:**

Das ist günstig.

Ich mache das.

Item	Distraktor	Antwort-format	Quelle	Variablen-name	Wertelabel
Ich mache mir einen Zeitplan.		dichotom (*ja – nein*)	angelehnt an Survivalguide Schreiben: Eine dezidierte Planung (Zeit, Ort usw.) entscheidet mit über den Erfolg von Schreibprojekten.	frdek1g frdek1a	1 = Ja 2 = Nein
Ich überlege mir, was ich brauche, um gut arbeiten zu können.		dichotom (*ja – nein*)	Eigenkonstruktion (Planung in Bezug auf externe Ressourcen)	frdek2g frdek2a	1 = Ja 2 = Nein
Ich mache mir bewusst, wie viel von der Arbeit abhängt.	x	dichotom (*ja – nein*)	Distraktor angelehnt an SVF-KJ: Emotionsregulierende Bewältigung: Bagatellisierung (BAG): Die Stärke, Dauer oder Gewichtung einer Belastung wird abgewertet.	frdek3g frdek3a	1 = Ja 2 = Nein

Ich stelle an mich Eigenkonstruktion realistische Ansprüche.		dichotom (*ja – nein*)	adaptiert nach Handbuch Lernstrategien, S. 234	frdek4g frdek4a	1 = Ja 2 = Nein
Ich verschiebe die Arbeit immer weiter nach hinten.	x	dichotom (*ja – nein*)	Handbuch Lernstrategien, S. 234	frdek5g frdek5a	1 = Ja 2 = Nein
Ich lasse die Dinge auf mich zukommen.	x	dichotom (*ja – nein*)	Disktraktor angelehnt an Pekrun (2006) Control-Value Theory of Achievement Emotions: geringe Kontrolle	frdek6g frdek6a	1 = Ja 2 = Nein
Ich vertraue darauf, dass mir die Aufgabe keine Schwierigkeiten bereiten wird.	x	dichotom (*ja – nein*)	Disktraktor angelehnt an SVF-KJ: Problemlösende Bewältigung: Situationskontrolle (STK): Die Situation analysieren, Handlungen zur Kontrolle/Problemlösung planen und ausführen	frdek7g frdek7a	1 = Ja 2 = Nein
Ich mache mir bewusst, dass ich gar keine andere Wahl habe, als mich durchzubeißen.	x	dichotom (*ja – nein*)	Distraktor mit Bezug auf SVF-KJ: Emotionsregulierende Bewältigung: Bagatellisierung (BAG): Die Stärke, Dauer oder Gewichtung einer Belastung wird abgewertet. Pekrun (2006) Control-Value Theory of Achievement Emotions: geringe Kontrolle	frdek8g frdek8a	1 = Ja 2 = Nein

Frustration subjektiver Studienerfolg

Zuletzt stellen wir Ihnen noch einige Fragen zu Ihrer Leistung im Studium.

Bitte schieben Sie den Balken auf die für Sie zutreffende Position.

Item	Antwortformat	Quelle	Variablenname	Wertelabel
Wie schätzen Sie Ihre eigene Leistung bei großen wissenschaftlichen Arbeiten ein?	Analogskala von 1 (*Nicht Genügend*) - 101 (*Sehr Gut*)	Eigenkonstruktion	subjgwA	Keine
Wie zufrieden waren Sie mit Ihrer letzten großen wissenschaftlichen Arbeit?	Analogskala von 1 (*gar nicht zufrieden*) - 101 (*völlig zufrieden*)		zufrgwA	Keine

Frustration objektiver Studienerfolg

Bitte bewerten Sie nach dem Notensystem.

Item	Antwortformat	Quelle	Variablenname	Wertelabel
Welche Note haben Sie auf Ihre letzte große wissenschaftliche Arbeit bekommen (bzw. auf das Seminar, wo diese Leistung in die Note eingeflossen ist)?	Auswahl	Eigenkonstruktion	notegwA	1 = Nicht Genügend 2= Genügend 3 = Befriedigend 4 = Gut 5 = Sehr Gut

Metakognition

SCHREIBEN EINER KLEINEREN WISSENSCHAFTLICHEN ARBEIT

Denken Sie an das Erstellen einer **kleineren wissenschaftlichen Arbeit** im Studium. Damit ist die Erstellung eines wissenschaftlichen Textes gemeint, für den Eigenkonstruktionständig Literatur recherchiert und aufgearbeitet werden muss und der bis zu einem vorgegebenen Zeitpunkt fertiggestellt werden muss (z.B. Seminararbeit, schriftliche Ausarbeitung etc.).

Eine Befragung von Studierenden hat ergeben, dass es hierbei vor allem wichtig ist, **über das eigene Vorgehen nachzudenken**.

Was ist aus Ihrer Sicht günstig, wenn man eine **kleinere wissenschaftliche Arbeit** schreibt?

Was machen Sie wirklich, wenn Sie eine kleinere wissenschaftliche Arbeit schreiben?

Metakognition deklarativ

In Bezug auf eine **kleinere wissenschaftliche Arbeit:**

Das ist günstig.

Ich mache das.

Item	Distraktor	Antwortformat	Quelle	Variablenname	Wertelabel
Bevor ich zu arbeiten beginne, überlege ich mir einen Zeitplan.		dichotom (*ja – nein*)	angelehnt an Wild & Schiefele (1994), Planung von Lernschritten	metdek1g metdek1a	1 = Ja 2 = Nein
Ich lasse die Dinge auf mich zukommen.	x	dichotom (*ja – nein*)	Eigenkonstruktion	metdek2g metdek2a	1 = Ja 2 = Nein

Ich denke nicht viel über die Arbeit nach, bevor ich beginne.	x	dichotom (*ja – nein*)	Distraktor angelehnt an Streblow & Schiefele (2006), Auswahl geeigneter Strategien	metdek3g metdek3a	1 = Ja 2 = Nein
Ich überlege mir vorab, welche Mittel (z.B. Lernmaterial, Hilfe, etc.) ich für meine Arbeit brauche.		dichotom (*ja – nein*)	Eigenkonstruktion? (Planung externer Ressourcen)	metdek4g metdek4a	1 = Ja 2 = Nein
Ich plane vorab, wie ich am besten beim Schreiben der wissenschaftlichen Arbeit vorgehen will.		dichotom (*ja – nein*)	angelehnt an Streblow & Schiefele (2006), Auswahl geeigneter Strategien	metdek5g metdek5a	1 = Ja 2 = Nein
Ich mache keine Zeitpläne, da man sich im Endeffekt ja doch nie daran hält.	x	dichotom (*ja – nein*)	Distraktor angelehnt an Wild & Schiefele (1994), Planung von Lernschritten	metdek6g metdek6a	1 = Ja 2 = Nein
Mein einziges Ziel ist es, die Arbeit abzugeben.	x	dichotom (*ja – nein*)	Eigenkonstruktion?	metdek7g metdek7a	1 = Ja 2 = Nein
Während des Schreibens frage ich mich immer wieder, ob mein Vorgehen gerade sinnvoll ist.		dichotom (*ja – nein*)	Eigenkonstruktion	metdek8g metdek8a	1 = Ja 2 = Nein
Ich überprüfe immer wieder, ob ich das, was ich mir vorgenommen habe, auch erreiche		dichotom (*ja – nein*)	angelehnt an Wild & Schiefele (1994) Überprüfung eigener Lernfortschritte; Einschätzung der eigenen Lernziele	metdek9g metdek9a	1 = Ja 2 = Nein
Während des Arbeitens verlasse ich mich darauf, dass mein ursprünglicher Zeitplan aufgeht.	x	dichotom (*ja – nein*)	Disktraktor angelehnt an Wild & Schiefele (1994) Überprüfung eigener Lernfortschritte; Einschätzung der eigenen Lernziele	metdekAg metdekAa	1 = Ja 2 = Nein
Wenn ich merke, dass ich nicht weiterkomme, ändere ich meine Vorgehensweise.		dichotom (*ja – nein*)	angelehnt an Wild & Schiefele (1994) adaptive Regulation durch	metdekBg metdekBa	1 = Ja 2 = Nein

			Änderung des Lernverhaltens		
Wenn ich merke, dass meine Ziele unrealistisch sind, überdenke ich sie nochmal.		dichotom (*ja – nein*)	Eigenkonstruktion	metdekCg metdekCa	1 = Ja 2 = Nein
Ich halte unter allen Umständen an meinen ursprünglichen Zielen fest.	x	dichotom (*ja – nein*)	Distraktor Eigenkonstruktion	metdekDg metdekDa	1 = Ja 2 = Nein
Ich überlege mir vor dem Schreiben, welche Ziele ich für die Arbeit habe.		dichotom (*ja – nein*)	angelehnt an Streblow & Schiefele (2006), Formulierung von Lernzielen	metdekEg metdekEa	1 = Ja 2 = Nein

Metakognition subjektiver Studienerfolg

Zuletzt stellen wir Ihnen noch einige Fragen zu Ihrer Leistung im Studium.

Bitte schieben Sie den Balken auf die für Sie zutreffende Position.

Item	Antwortformat	Quelle	Variablenname	Wertelabel
Wie schätzen Sie Ihre eigene Leistung bei kleinen wissenschaftlichen Arbeiten ein?	Analogskala von 1 (nicht genügend) - 101 (sehr gut)	Eigenkonstruktion	subjkwA	Keine
Wie zufrieden waren Sie mit Ihrer letzten kleinen wissenschaftlichen Arbeit?	Analogskala von 1 (gar nicht zufrieden) - 101 (völlig zufrieden)		zufrkwA	Keine

Metakognition objektiver Studienerfolg

Bitte bewerten Sie nach dem Notensystem.

Item	Antwortformat	Quelle	Variablenname	Wertelabel
Welche Note haben Sie auf Ihre letzte kleine wissenschaftliche Arbeit bekommen (bzw. auf das Seminar, in dem diese Leistung in die Note eingeflossen ist)?	Auswahl	Eigenkonstruktion	notekwA	1 = Nicht Genügend 2= Genügend 3 = Befriedigend 4 = Gut 5 = Sehr Gut

Kognition

VORBEREITUNG AUF EINE PRÜFUNG

Denken Sie an eine **Prüfungsvorbereitung**. Dabei müssen Sie große Stoffmengen selbstständig aufbereiten und lernen, wobei das Verständnis des Lernstoffs wesentlich ist.

Eine Befragung von Studierenden hat ergeben, dass hierbei vor allem **Lernstrategien** wichtig sind.
Was ist aus Ihrer Sicht günstig bei der **Vorbereitung auf eine Prüfung?**

Was machen Sie wirklich, wenn Sie sich auf eine Prüfung vorbereiten?

Kognition deklarativ

In Bezug auf die **Vorbereitung auf eine Prüfung:**

Das ist günstig.

Ich mache das.

Item	Ungünstige Strategie	Antwortformat	Quelle	Variablenname	Wertelabel
Ich strukturiere den zu lernenden Stoff.		dichotom (*ja – nein*)	angelehnt an Streblow & Schiefele (2006): Kernaussagen aus Texten herausarbeiten Inhalte neu strukturieren LIST: „ Für größere Stoffmengen fertige ich eine Gliederung an, die die Struktur des Stoffs am besten wiedergibt.“	kodek1g kodek1a	1 = Ja 2 = Nein

Ich beschäftige mich tiefgehend und ausführlich mit dem Stoff.		dichotom (*ja – nein*)	Tiefenstrategien – Zitat allgemein	kodek2g kodek2a	1 = Ja 2 = Nein
Ich fange direkt an zu lernen, der Rest ergibt sich.	x	dichotom (*ja – nein*)	Ungünstige Strategie Eigenkonstruktion	kodek3g kodek3a	1 = Ja 2 = Nein
Ich wiederhole den Stoff.		dichotom (*ja – nein*)	angelehnt an Wild & Schiefele (1994) aktives Wiederholen einzelner Fakten, wiederholte Durcharbeitung und Streblow & Schiefele (2006): Lernstoff immer wieder laut oder leise aufsagen	kodek4g kodek4a	1 = Ja 2 = Nein
Ich lerne eine Zusammenfassung.	x	dichotom (*ja – nein*)	Ungünstige Strategie Eigenkonstruktion	kodek5g kodek5a	1 = Ja 2 = Nein
Ich lese mir nur einmal das gesamte Skript durch.	x	dichotom (*ja – nein*)	Ungünstige Strategie Eigenkonstruktion	kodek6g kodek6a	1 = Ja 2 = Nein
Ich lerne den Stoff auswendig.	x	dichotom (*ja – nein*)	Ungünstige Strategie Eigenkonstruktion	kodek7g kodek7a	1 = Ja 2 = Nein

Kognition subjektiver Studienerfolg

Zuletzt stellen wir Ihnen noch einige Fragen zu Ihrer Leistung im Studium.

Wir bitten Sie, die zutreffende Einschätzung abzugeben:

Bitte schieben Sie den Balken auf die für Sie zutreffende Position.

Item	Antwortformat	Quelle	Variablenname	Wertelabel
Wie schätzen Sie Ihre eigene Leistung bei Prüfungsvorbereitungen ein?	Schieberegler von 1 (nicht genügend) - 101 (sehr gut)	Eigenkonstruktion	subjPr	
Wie zufrieden waren Sie mit Ihrer letzten Prüfungsvorbereitung?			zufrPr	

Kognition objektiver Studienerfolg

Instruktion: Bitte bewerten Sie nach dem Notensystem.

Item	Antwortformat	Quelle	Variablenname	Wertelabel
Welche Note haben Sie auf Ihre letzte Prüfung bekommen?	Auswahl	Eigenkonstruktion	notePr	1 = Nicht Genügend 2= Genügend 3 = Befriedigend 4 = Gut 5 = Sehr Gut

Langeweile

LANGEWEILE IN EINER VORLESUNG

Denken Sie an den Besuch einer **Vorlesung**. Damit ist eine Lehrveranstaltung mit einer größeren Anzahl von Studierenden gemeint, in der große Stoffmengen durch eine/n Dozierende/n vorgetragen werden.

Eine Befragung von Studierenden hat ergeben, dass hierbei vor allem **Langeweile** ein Problem ist.

Was ist aus Ihrer Sicht günstig, um **Langeweile in einer Vorlesung entgegenzuwirken**?

Was machen Sie wirklich, damit Ihnen in einer Vorlesung nicht langweilig wird?

Langeweile deklarativ

In Bezug auf **Langeweile in einer Vorlesung:**

Das ist günstig.

Ich mache das.

Item	Ungünstige Strategie	Antwortformat	Quelle	Variablenname	Wertelabel
Ich mache mir bewusst, welchen Nutzen ich aus der Vorlesung ziehen kann.		dichotom (*ja – nein*)	Subjektiv höher extrinsisch valent uminterpretieren; Götz (2007)	ladek1g ladek1a	1 = Ja 2 = Nein
Ich bereite mich auf die Vorlesung vor.		dichotom (*ja – nein*)	Eigenkonstruktion angelehnt an Götz (2007): Aktivierung metakognitiver Strategie; Coping	ladek2g ladek2a	1 = Ja 2 = Nein

Ich mache mir bewusst, dass ich als Studierende/r Vorlesungen besuchen muss.	x	dichotom (*ja* – *nein*)	Eigenkonstruktion nach Pekrun (2006) Control-Value Theory of Achievement Emotions: geringer Wert	ladek3g ladek3a	1 = Ja 2 = Nein
Ich beteilige mich aktiv an der Vorlesung.		dichotom (*ja* – *nein*)	adaptiert nach Widulle (2009) Frontalunterricht aktiv nutzen	ladek4g ladek4a	1 = Ja 2 = Nein
Ich lenke mich während der Vorlesung ab.	x	dichotom (*ja* – *nein*)	Avoidance, Nett (2010)	ladek5g ladek5a	1 = Ja 2 = Nein
Ich gehe ohne große Erwartungen in die Vorlesung.	x	dichotom (*ja* – *nein*)	Eigenkonstruktion nach Pekrun (2006) Control-Value Theory of Achievement Emotions: geringer Wert	ladek6g ladek6a	1 = Ja 2 = Nein

Langeweile subjektiver Studienerfolg

Zuletzt stellen wir Ihnen noch einige Fragen zu Ihrer Leistung im Studium.

Wir bitten Sie, die zutreffende Einschätzung abzugeben:

Bitte schieben Sie den Balken auf die für Sie zutreffende Position.

Item	Antwortformat	Quelle	Variablenname	Wertelabel
Wie schätzen Sie Ihren Wissensgewinn bei Vorlesungsbesuchen ein?	Schieberegler von 1 (*nicht genügend*) - 101 (*sehr gut*) mit markierter Mitte (51)	Eigenkonstruktion	subjPr	Keine
Wie zufrieden waren Sie mit dem erlangten Wissensgewinn in Ihrer letzten besuchten Vorlesung?	Schieberegler von 1 (*gar nicht zufrieden*) - 101 (*völlig zufrieden*) mit markierter Mitte (51)		zufrPr	Keine

Langeweile objektiver Studienerfolg

Bitte bewerten Sie nach dem Notensystem.

Item	Antwortformat	Quelle	Variablenname	Wertelabel
Welche Note haben Sie auf Ihre letzte Vorlesungsprüfung bekommen, deren Vorlesung Sie besucht haben?	Auswahl	Eigenkonstruktion	notePr	1 = Nicht Genügend 2= Genügend 3 = Befriedigend 4 = Gut 5 = Sehr Gut

Gesamt: objektiver Studienerfolg

Instruktion: Bitte bewerten Sie nach dem Notensystem.

Welche Noten haben Sie auf Ihre letzten 3 Prüfungen erhalten?

Item	Antwortformat	Quelle	Variablenname	Wertelabel
letzte Lehrveranstaltungsnote	offen	Eigenkonstruktion	noteges1	
vorletzte Lehrveranstaltungsnote			noteges2	
drittletzte Lehrveranstaltungsnote			noteges3	

Infos TeilnehmerInnen

Wir möchten Sie zum Abschluss bitte, einige Angaben zu Ihrer Person zu machen. Aus diesen Angaben ist kein Rückschluss auf Ihre Identität möglich.

Item	Antwortformat	Quelle	Variablenname	Wertelabel
Bitte geben Sie Ihr Alter in Jahren an.	freie Eingabe	Eigenkonstruktion	alter	
Sie sind...	Auswahl	Eigenkonstruktion	geschl	1 = männlich 2 = weiblich

Was ist Ihr höchster Abschluss?	Auswahl	Eigenkonstruktion	abschl abschla abschlb	1 = Matura/Abitur 2 = Fachmatura/ Fachabitur (BHS etc.) 3 = anderer Schulabschluss (freie Eingabe) 4 = Bachelor in (freie Eingabe) 5 = abgeschlossenes Studium in (freie Eingabe)
Maturanote/ Abiturnote (Abschlussnote bei anderen Abschlüssen) Bitte geben Sie Ihren Notendurchschnitt bei der Matura/ beim Abitur an. Falls Sie sich nicht mehr genau erinnern, reicht eine Schätzung.	freie Eingabe	Eigenkonstruktion	notemat	
In welchem Semester befinden Sie sich aktuell? (Gesamtzahl der Semester in allen je studierten Fächern)	freie Eingabe	Eigenkonstruktion	semester	

Die Daten werden selbstverständlich nicht mit Ihren Angaben in Verbindung gebracht und streng vertraulich behandelt.

Danke, dass Sie sich Zeit genommen haben!

Literatur Codebook Self-Report Questionnaire (IWAN) and Expert Rating

Bensberg, Gabriele. (2013). Survivalguide Schreiben. Berlin: Springer.

Bliuc, A., Ellis, R. A., Goodyear, P. & Hendres, D. M. (2011). The role of social identification as university student in learning: relationships between students‘ social identity, approaches to learning and academic achievement. *Educational Psychology, 31(5)*, S. 559-574.

Hampel, P., Petermann, F. & Dickow, B. (2001). SVF-KJ. Stressverarbeitungsfragebogen von Janke und Erdmann angepasst für Kinder und Jugendliche. Göttingen: Hogrefe.

Dickhäuser, O., Schöne, C., Spinath, B., Stiensmeier-Pelster, J. (2002). Die Skalen zum akademischen Selbstkonzept. Konstruktion und Überprüfung eines neuen Konzeptes. *Zeitschrift für Differentielle und Diagnostische Psychologie, 23(4),* S. 393–405.

Doosje, B., Ellemers, N., & Spears, R. (1995). Perceived Intragroup Variability as a Function of Group Status and Identification. *Journal of Experimental Social Psychology, 31(5), S.* 410-436.

Götz, T., Frenzel, A. C., & Pekrun, R. (2007). Regulation von Langeweile im Unterricht Was Schülerinnen und Schüler bei der „Windstille der Seele "(nicht) tun. *Unterrichtswissenschaft*, 35(4), 312-333.

Grob, A. & Smolenski, C. (2005). *Fragebogen zur Erhebung der Emotionsregulation bei Kindern und Jugendlichen*. Göttingen: Verlag Hans Huber.

Knigge-Illner, H. (2009). Prüfungsangst bewältigen. Psychotherapeut 9, 334-345.

Krapp, A., Schiefele, U., Wild, K.P., Winteler, A. (1993). Fragebogen zum Studieninteresse (FSI). *Diagnostika, 39,* S. 335-351.

Lompscher, J. (1995). Erfassung von Lernstrategien mittels Fragebogen. *Lern-und Lehrforschung, LLF-Berichte Nr, 10*.

Mandl, H., & Friedrich, H. F. (2006). Handbuch Lernstrategien. Hogrefe Verlag.

Nett (2010) Self-regulation in School. Disseration.

Pekrun, R. (2006). The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. *Educational Psychological Review*, *18*, 315–341.

Pekrun, R., & Götz, T. (2006). Emotionsregulation: Vom Umgang mit Prüfungsangst. *Handbuch Lernstrategien*, 248-258.

Pohlmann, B. & Möller, J. (2010). Fragebogen zur Erfassung der Motivation für die Wahl des Lehramtsstudiums (FEMOLA). *Zeitschrift für Pädagogische Psychologie, 24(1),* S. 73-84.

Prenzel, M., Kristen, A., Dengler, P., Ettle, R. & Beer, T. (1996). Selbstbestimmt motiviertes und interessiertes Lernen in der kaufmännischen Erstausbildung. *Zeitschrift für Berufs- und Wirtschaftspädagogik, Beiheft 13,* 108-127.

Renz, K. C. (2013). *Das 1 x 1 der Präsentation*. Springer Fachmedien Wiesbaden.

Satow, L. (2012). Stress- und Coping-Inventar (SCI): Test- und Skalendokumentation. URL: www.drsatow.de.

Schober, Barbara, Dresel, M., & Ziegler, A. (2007). Warum Elterneinflüsse in der Koedukationsdebatte berücksichtigt werden sollten: Erwartungen, Einstellungen und Überzeugungen von Eltern von monoedukativ und von koedukativ unterrichteten Mädchen. In P. Ludwig & H. Ludwig (Eds.), *Erwartungen in himmelblau und rosarot – Auslöser lebenslanger Geschlechterdifferenzen im Lernen?* Opladen: Verlag für Sozialwissenschaften.

Schwartzer, R. & Jerusalem, M. (Hrsg.) (1999). Skalen zur Erfassung von Lehrer- und Schülermerkmalen. Dokumentation der psychometrischen Verfahren im Rahmen der Wissenschaftlichen Begleitung des Modellversuchs selbstwirksame Schulen. Berlin: Freie Universität Berlin.

Streblow, L.; Schiefele, U. (2006). Lernstrategien im Studium. In: Heinz Mandel; Helmut Felix Friedrich: Handbuch Lernstrategien. Hogrefe Verlag

Widulle, W. (2009). Handlungsorientiert Lernen im Studium. *Arbeitsbuch für soziale und pädagogische Berufe. Wiesbaden: Verlag für Sozialwissenschaften*.

Wild, K. P., & Schiefele, U. (1994). Lernstrategien im Studium: Ergebnisse zur Faktorenstruktur und Reliabilität eines neuen Fragebogens. *Zeitschrift für Differentielle und Diagnostische Psychologie*.

Appendix C

Expert Rating

Fragebogen

1.1 Metakognition Vignette

SCHREIBEN EINER KLEINEREN WISSENSCHAFTLICHEN ARBEIT

Denken Sie an das Erstellen einer **kleineren wissenschaftlichen Arbeit** im Studium **(z.B. Seminararbeit, schriftliche Ausarbeitung etc.)**. Damit ist die Erstellung eines wissenschaftlichen Textes gemeint, für den selbstständig Literatur recherchiert und aufgearbeitet werden muss und für den es einen festgelegten Abgabetermin gibt.

Eine Befragung von Studierenden hat ergeben, dass es hierbei vor allem wichtig ist, **metakognitive Strategien anzuwenden**.

1.1.1 deklarativ

Welche der folgenden Strategien sind aus Ihrer Sicht **für das erfolgreiche Erstellen einer kleineren wissenschaftlichen Arbeit günstig**?

- ☐ Bevor ich zu arbeiten beginne, überlege ich mir einen Zeitplan.
- ☐ Ich lasse die Dinge auf mich zukommen.
- ☐ Ich denke nicht viel über die Arbeit nach, bevor ich beginne.
- ☐ Ich überlege mir vorab, welche Mittel (z.B. Lernmaterial, Hilfe, etc.) ich für meine Arbeit brauche.
- ☐ Ich plane vorab, wie ich am besten beim Schreiben der wissenschaftlichen Arbeit vorgehen will.
- ☐ Ich mache keine Zeitpläne, da man sich im Endeffekt ja doch nie daran hält.
- ☐ Mein einziges Ziel ist es, die Arbeit abzugeben.
- ☐ Während des Schreibens frage ich mich immer wieder, ob mein Vorgehen gerade sinnvoll ist.
- ☐ Ich überprüfe immer wieder, ob ich das, was ich mir vorgenommen habe, auch erreiche.
- ☐ Während des Arbeitens verlasse ich mich darauf, dass mein ursprünglicher Zeitplan aufgeht.
- ☐ Wenn ich merke, dass ich nicht weiterkomme, ändere ich meine Vorgehensweise.
- ☐ Wenn ich merke, dass meine Ziele unrealistisch sind, überdenke ich sie nochmal.
- ☐ Ich halte unter allen Umständen an meinen ursprünglichen Zielen fest.
- ☐ Ich überlege mir vor dem Schreiben, welche Ziele ich für die Arbeit habe.

1.1.2 wie günstig

Sie haben angegeben, dass folgende Strategien für das erfolgreiche Erstellen einer kleineren wissenschaftlichen Arbeit günstig sind.

Wir bitten Sie nun einzuschätzen, **wie relevant** die einzelnen Strategien **für das erfolgreiche Erstellen einer kleineren wissenschaftlichen Arbeit** sind.

Bevor ich zu arbeiten beginne, überlege ich mir	kaum relevant	|	sehr relevant

einen Zeitplan.

Ich überlege mir vorab, welche Mittel (z.B. Lernmaterial, Hilfe, etc.) ich für meine Arbeit brauche.

kaum relevant | sehr relevant

Ich plane vorab, wie ich am besten beim Schreiben der wissenschaftlichen Arbeit vorgehen will.

kaum relevant | sehr relevant

Während des Schreibens frage ich mich immer wieder, ob mein Vorgehen gerade sinnvoll ist.

kaum relevant | sehr relevant

Ich überprüfe immer wieder, ob ich das, was ich mir vorgenommen habe, auch erreiche.

kaum relevant | sehr relevant

Wenn ich merke, dass ich nicht weiterkomme, ändere ich meine Vorgehensweise.

kaum relevant | sehr relevant

Wenn ich merke, dass meine Ziele unrealistisch sind, überdenke ich sie nochmal.

kaum relevant | sehr relevant

Ich überlege mir vor dem Schreiben, welche Ziele ich für die Arbeit habe.

kaum relevant | sehr relevant

Ich lasse die Dinge auf mich zukommen.

kaum relevant | sehr relevant

Ich denke nicht viel über die Arbeit nach, bevor ich beginne.

kaum relevant | sehr relevant

Ich mache keine Zeitpläne, da man sich im Endeffekt ja doch nie daran hält.	kaum relevant	I	sehr relevant
Mein einziges Ziel ist es, die Arbeit abzugeben.	kaum relevant	I	sehr relevant
Während des Arbeitens verlasse ich mich darauf, dass mein ursprünglicher Zeitplan aufgeht.	kaum relevant	I	sehr relevant
Ich halte unter allen Umständen an meinen ursprünglichen Zielen fest.	kaum relevant	I	sehr relevant

1.1.3 wie ungünstig

Sie haben angegeben, dass folgende Strategien nicht günstig sind, um beim Schreiben einer kleineren wissenschaftlichen Arbeit erfolgreich zu sein.

Wir bitten Sie nun einzuschätzen, **wie hinderlich** die einzelnen Strategien **für das erfolgreiche Erstellen einer kleineren wissenschaftlichen Arbeit** sind.

Ich lasse die Dinge auf mich zukommen.	kaum hinderlich	I	sehr hinderlich
Ich denke nicht viel über die Arbeit nach, bevor ich beginne.	kaum hinderlich	I	sehr hinderlich
Ich mache keine Zeitpläne, da man sich im Endeffekt ja doch nie daran hält.	kaum hinderlich	I	sehr hinderlich
Mein einziges Ziel ist es, die Arbeit abzugeben.	kaum hinderlich	I	sehr hinderlich

Während des Arbeitens verlasse ich mich darauf, dass mein ursprünglicher Zeitplan aufgeht.	kaum hinderlich	I	sehr hinderlich
Ich halte unter allen Umständen an meinen ursprünglichen Zielen fest.	kaum hinderlich	I	sehr hinderlich
Bevor ich zu arbeiten beginne, überlege ich mir einen Zeitplan.	kaum hinderlich	I	sehr hinderlich
Ich überlege mir vorab, welche Mittel (z.B. Lernmaterial, Hilfe, etc.) ich für meine Arbeit brauche.	kaum hinderlich	I	sehr hinderlich
Ich plane vorab, wie ich am besten beim Schreiben der wissenschaftlichen Arbeit vorgehen will.	kaum hinderlich	I	sehr hinderlich
Während des Schreibens frage ich mich immer wieder, ob mein Vorgehen gerade sinnvoll ist.	kaum hinderlich	I	sehr hinderlich
Ich überprüfe immer wieder, ob ich das, was ich mir vorgenommen habe, auch erreiche.	kaum hinderlich	I	sehr hinderlich
Wenn ich merke, dass ich nicht weiterkomme, ändere ich meine Vorgehensweise.	kaum hinderlich	I	sehr hinderlich
Wenn ich merke, dass meine Ziele unrealistisch sind, überdenke ich sie nochmal.	kaum hinderlich	I	sehr hinderlich

Ich überlege mir vor dem Schreiben, welche Ziele ich für die Arbeit habe.	kaum hinderlich	I	sehr hinderlich

1.2 Kognition Vignette

VORBEREITUNG AUF EINE PRÜFUNG

Denken Sie an eine **Prüfungsvorbereitung**. Dabei müssen Studierende große Stoffmengen selbständig aufbereiten und lernen, wobei das Verständnis des Lernstoffs wesentlich ist.

Eine Befragung von Studierenden hat ergeben, dass hierbei vor allem **kognititve Lernstrategien** wichtig sind.

1.2.1 deklarativ

Welche der folgenden Strategien sind aus Ihrer Sicht **für die erfolgreiche Vorbereitung auf eine Prüfung günstig?**

- Ich strukturiere den zu lernenden Stoff.
- Ich beschäftige mich tiefgehend und ausführlich mit dem Stoff.
- Ich fange direkt an zu lernen, der Rest ergibt sich.
- Ich wiederhole den Stoff.
- Ich lerne eine Zusammenfassung.
- Ich lese mir nur einmal das gesamte Skript durch.
- Ich lerne den Stoff auswendig.

1.2.2 wie günstig

Sie haben angegeben, dass folgende Strategien für die erfolgreiche Vorbereitung auf eine Prüfung günstig sind.

Wir bitten Sie nun einzuschätzen, **wie relevant** die einzelnen Strategien **für die erfolgreiche Vorbereitung auf eine Prüfung** sind.

Ich strukturiere den zu lernenden Stoff.	kaum relevant	I	sehr relevant
Ich beschäftige mich tiefgehend und ausführlich mit dem Stoff.	kaum relevant	I	sehr relevant

Ich wiederhole den Stoff.	kaum relevant	|	sehr relevant
Ich fange direkt an zu lernen, der Rest ergibt sich.	kaum relevant	|	sehr relevant
Ich lerne eine Zusammenfassung.	kaum relevant	|	sehr relevant
Ich lese mir nur einmal das gesamte Skript durch.	kaum relevant	|	sehr relevant
Ich lerne den Stoff auswendig.	kaum relevant	|	sehr relevant

1.2.3 wie ungünstig

Sie haben angegeben, dass folgende Strategien nicht günstig für die erfolgreiche Vorbereitung auf eine Prüfung sind.

Wir bitten Sie nun einzuschätzen, **wie hinderlich** die einzelnen Strategien **für die erfolgreiche Vorbereitung auf eine Prüfung** sind.

Ich fange direkt an zu lernen, der Rest ergibt sich.	kaum hinderlich	|	sehr hinderlich
Ich lerne eine Zusammenfassung.	kaum hinderlich	|	sehr hinderlich
Ich lese mir nur einmal das gesamte Skript durch.	kaum hinderlich	|	sehr hinderlich
Ich lerne den Stoff auswendig.	kaum hinderlich	|	sehr hinderlich

Ich strukturiere den zu lernenden Stoff.	kaum hinderlich	|	sehr hinderlich
Ich beschäftige mich tiefgehend und ausführlich mit dem Stoff.	kaum hinderlich	|	sehr hinderlich
Ich wiederhole den Stoff.	kaum hinderlich	|	sehr hinderlich

1.3 Langeweile Vignette

LANGEWEILE IN EINER VORLESUNG

Denken Sie an den Besuch einer **Vorlesung**. Damit ist eine Lehrveranstaltung mit einer größeren Anzahl von Studierenden gemeint, in der große Stoffmengen durch eine/n Dozierende/n vorgetragen werden.

Eine Befragung von Studierenden hat ergeben, dass hierbei vor allem die **Regulation von Langeweile** von Bedeutung ist.

1.3.1 deklarativ

Welche der folgenden Strategien sind aus Ihrer Sicht bei der **Regulation von Langeweile für einen ergiebigen Vorlesungsbesuch günstig?**

- Ich mache mir bewusst, welchen Nutzen ich aus der Vorlesung ziehen kann.
- Ich bereite mich auf die Vorlesung vor.
- Ich mache mir bewusst, dass ich als Studierende/r Vorlesungen besuchen muss.
- Ich beteilige mich aktiv an der Vorlesung.
- Ich lenke mich während der Vorlesung ab.
- Ich gehe ohne große Erwartungen in die Vorlesung.

1.3.2 wie günstig

Sie haben angegeben, dass folgende Strategien günstig sind, um Langeweile bei einer Vorlesung zu regulieren.

Wir bitten Sie nun einzuschätzen, **wie relevant** die einzelnen Strategien **bei der Regulation von Langeweile** für einen **ergiebigen Vorlesungsbesuch** sind.

Ich mache mir bewusst, welchen Nutzen ich aus der	kaum relevant	|	sehr relevant

Vorlesung ziehen kann.			
Ich bereite mich auf die Vorlesung vor.	kaum relevant		sehr relevant
Ich beteilige mich aktiv an der Vorlesung.	kaum relevant		sehr relevant
Ich mache mir bewusst, dass ich als Studierende/r Vorlesungen besuchen muss.	kaum relevant		sehr relevant
Ich lenke mich während der Vorlesung ab.	kaum relevant		sehr relevant
Ich gehe ohne große Erwartungen in die Vorlesung.	kaum relevant		sehr relevant

1.3.3 wie ungünstig

Sie haben angegeben, dass folgende Strategien nicht günstig sind, um Langeweile bei einer Vorlesung zu regulieren.

Wir bitten Sie nun einzuschätzen, **wie hinderlich** die einzelnen Strategien **bei der Regulation von Langeweile** für einen **ergiebigen Vorlesungsbesuch** sind.

Ich mache mir bewusst, dass ich als Studierende/r Vorlesungen besuchen muss.	kaum hinderlich		sehr hinderlich
Ich lenke mich während der Vorlesung ab.	kaum hinderlich		sehr hinderlich
Ich gehe ohne große Erwartungen in die Vorlesung.	kaum hinderlich		sehr hinderlich

Ich mache mir bewusst, welchen Nutzen ich aus der Vorlesung ziehen kann.	kaum hinderlich	|	sehr hinderlich
Ich bereite mich auf die Vorlesung vor.	kaum hinderlich	|	sehr hinderlich
Ich beteilige mich aktiv an der Vorlesung.	kaum hinderlich	|	sehr hinderlich

1.4 Frustration Vignette

FRUSTRATION BEI EINER GROSSEN WISSENSCHAFTLICHEN ABSCHLUSSARBEIT

Denken Sie an das Erstellen einer **großen wissenschaftlichen Abschlussarbeit** im Studium (Diplomarbeit, Masterarbeit, eventuell auch Bachelorarbeit). Dazu muss eine umfangreiche Thematik selbstständig über mehrere Monate hinweg unter Betreuung eines/einer Dozierenden bearbeitet werden.

Eine Befragung von Studierenden hat ergeben, dass hierbei vor allem die **Regulation von Frustration** von Bedeutung ist.

1.4.1 deklarativ

Welche der folgenden Strategien sind aus Ihrer Sicht **günstig, um beim Erstellen einer großen wissenschaftlichen Abschlussarbeit das Entstehen von Frustration zu verhindern?**

- ☐ Ich mache mir einen Zeitplan.
- ☐ Ich überlege mir, was ich brauche, um gut arbeiten zu können.
- ☐ Ich mache mir bewusst, wie viel von der Arbeit abhängt.
- ☐ Ich stelle an mich selbst realistische Ansprüche.
- ☐ Ich verschiebe die Arbeit immer weiter nach hinten.
- ☐ Ich lasse die Dinge auf mich zukommen.
- ☐ Ich vertraue darauf, dass mir die Aufgabe keine Schwierigkeiten bereiten wird.
- ☐ Ich mache mir bewusst, dass ich gar keine andere Wahl habe, als mich durchzubeißen.

1.4.2 wie günstig

Sie haben angegeben, dass folgende Strategien günstig sind, um Frustration beim Verfassen einer großen wissenschaftlichen Abschlussarbeit zu regulieren.

Wir bitten Sie nun einzuschätzen, **wie relevant** die einzelnen Strategien **bei der Regulation von Frustration** für das **erfolgreiche Schreiben einer großen wissenschaftlichen Abschlussarbeit** sind.

Ich mache mir einen Zeitplan.	kaum relevant	I	sehr relevant
Ich überlege mir, was ich brauche, um gut arbeiten zu können.	kaum relevant	I	sehr relevant
Ich stelle an mich selbst realistische Ansprüche.	kaum relevant	I	sehr relevant
Ich mache mir bewusst, wie viel von der Arbeit abhängt.	kaum relevant	I	sehr relevant
Ich verschiebe die Arbeit immer weiter nach hinten.	kaum relevant	I	sehr relevant
Ich lasse die Dinge auf mich zukommen.	kaum relevant	I	sehr relevant
Ich vertraue darauf, dass mir die Aufgabe keine Schwierigkeiten bereiten wird.	kaum relevant	I	sehr relevant
Ich mache mir bewusst, dass ich gar keine andere Wahl habe, als mich durchzubeißen.	kaum relevant	I	sehr relevant

1.4.3 wie ungünstig

Sie haben angegeben, dass folgende Strategien nicht günstig sind, um Frustration beim Verfassen einer großen wissenschaftlichen Abschlussarbeit zu regulieren.

Wir bitten Sie nun einzuschätzen, **wie hinderlich** die einzelnen Strategien **bei der Regulation von**

Frustration für das **erfolgreiche Schreiben einer großen wissenschaftlichen Abschlussarbeit** sind.

Ich mache mir bewusst, wie viel von der Arbeit abhängt.	kaum hinderlich		sehr hinderlich
Ich verschiebe die Arbeit immer weiter nach hinten.	kaum hinderlich		sehr hinderlich
Ich lasse die Dinge auf mich zukommen.	kaum hinderlich		sehr hinderlich
Ich vertraue darauf, dass mir die Aufgabe keine Schwierigkeiten bereiten wird.	kaum hinderlich		sehr hinderlich
Ich mache mir bewusst, dass ich gar keine andere Wahl habe, als mich durchzubeißen.	kaum hinderlich		sehr hinderlich
Ich mache mir einen Zeitplan.	kaum hinderlich		sehr hinderlich
Ich überlege mir, was ich brauche, um gut arbeiten zu können.	kaum hinderlich		sehr hinderlich
Ich stelle an mich selbst realistische Ansprüche.	kaum hinderlich		sehr hinderlich

2 Endseite

Vielen Dank für Ihre Unterstützung bei der inhaltlichen Validierung unseres Instrumentes!

Appendix D

Codebook: Expert Rating

Sehr geehrte Damen und Herren

Wir wenden uns an Sie als renommierte Expertin/ renommierten Experten im Feld des Selbstregulierten Lernens im tertiären Bildungsbereich. Wir möchten Sie darum ersuchen, uns bei der inhaltlichen Validierung unseres Instruments zur Messung von SRL-Kompetenzen im Tertiären Bildungsbereich zu unterstützen.

Dem Instrument liegt die Annahme zugrunde, dass SRL-Strategien (Kognitive Strategien, Metakognitive Strategien, Ressourcenmanagement) jeweils in hohem Maße situationsabhängig sind. Demzufolge wird die Anwendung sämtlicher SRL-Strategien nicht in allen Lernkontexten als günstig angesehen. In zwei Vorstudien wurden anhand von ExpertInneninterviews (mit Lehrpersonen sowie besonders erfolgreichen Studierenden) spezifische Lernsituationen ermittelt, in welchen die Anwendung bestimmter SRL-Strategien als besonders relevant erachtet wird. Dabei ergaben sich folgende Kombinationen: Schreiben einer kleineren wissenschaftlichen Arbeit x Metakognitive Strategien; Schreiben einer großen wissenschaftlichen Arbeit x Frustrationsregulationsstrategien; Prüfungsvorbereitung x Kognitive Strategien; Vorlesungsbesuch x Langeweileregulationsstrategien.

Wir möchten Sie nun um ihre Einschätzung als Expertin/Experte bitten: Bitte bewerten Sie auf Basis Ihres theoretischen und praktischen Wissens, ob die im Folgenden präsentierten SRL-Strategien für die erfolgreiche Bewältigung der jeweiligen Lernsituation günstig oder ungünstig sind. Zudem werden Sie jeweils im Anschluss daran gebeten, mittels eines Schiebereglers einzuschätzen, in welchem Ausmaß die jeweilige SRL-Strategie für die erfolgreiche Bewältigung der spezifischen Situation förderlich beziehungsweise hinderlich ist.

Sie können das ExpertInnenrating unter folgendem Link abrufen:

https://ww3.unipark.de/uc/expertsges/

Vielen Dank für Ihre Unterstützung!
Das PRO-SRL Projektteam

Frustration

FRUSTRATION BEI EINER GROSSEN WISSENSCHAFTLICHEN ABSCHLUSSARBEIT

Denken Sie an das Erstellen einer **großen wissenschaftlichen Abschlussarbeit** im Studium (Diplomarbeit, Masterarbeit, eventuell auch Bachelorarbeit). Dazu muss eine umfangreiche Thematik selbstständig über mehrere Monate hinweg unter Betreuung eines/einer Dozierenden bearbeitet werden.

Eine Befragung von Studierenden hat ergeben, dass hierbei vor allem die **Regulation von Frustration** von Bedeutung ist.

Frustration deklarativ

Welche der folgenden Strategien sind aus Ihrer Sicht **günstig, um beim Erstellen einer großen wissenschaftlichen Abschlussarbeit das Entstehen von Frustration zu verhindern?**

Item	Distraktor	Antwortformat	Quelle	Variablenname	Wertelabel
Ich mache mir einen Zeitplan.		dichotom (*ja – nein*)	angelehnt an Survivalguide Schreiben: Eine dezidierte Planung (Zeit, Ort usw.) entscheidet mit über den Erfolg von Schreibprojekten.	frdek1g frdek1a	1 = Ja 2 = Nein
Ich überlege mir, was ich brauche, um gut arbeiten zu können.		dichotom (*ja – nein*)	Eigenkonstruktion (Planung in Bezug auf externe Ressourcen)	frdek2g frdek2a	1 = Ja 2 = Nein
Ich mache mir bewusst, wie viel von der Arbeit abhängt.	x	dichotom (*ja – nein*)	Distraktor angelehnt an SVF-KJ: Emotionsregulierende Bewältigung: Bagatellisierung (BAG): Die Stärke, Dauer oder Gewichtung einer Belastung wird abgewertet.	frdek3g frdek3a	1 = Ja 2 = Nein
Ich stelle an mich Eigenkonstruktion realistische Ansprüche.		dichotom (*ja – nein*)	adaptiert nach Handbuch Lernstrategien, S. 234	frdek4g frdek4a	1 = Ja 2 = Nein

Ich verschiebe die Arbeit immer weiter nach hinten.	x	dichotom (*ja – nein*)	Handbuch Lernstrategien, S. 234	frdek5g frdek5a	1 = Ja 2 = Nein
Ich lasse die Dinge auf mich zukommen.	x	dichotom (*ja – nein*)	Disktraktor angelehnt an Pekrun (2006) Control-Value Theory of Achievement Emotions: geringe Kontrolle	frdek6g frdek6a	1 = Ja 2 = Nein
Ich vertraue darauf, dass mir die Aufgabe keine Schwierigkeiten bereiten wird.	x	dichotom (*ja – nein*)	Disktraktor angelehnt an SVF-KJ: Problemlösende Bewältigung: Situationskontrolle (STK): Die Situation analysieren, Handlungen zur Kontrolle/Problemlösung planen und ausführen	frdek7g frdek7a	1 = Ja 2 = Nein
Ich mache mir bewusst, dass ich gar keine andere Wahl habe, als mich durchzubeißen.	x	dichotom (*ja – nein*)	Distraktor mit Bezug auf SVF-KJ: Emotionsregulierende Bewältigung: Bagatellisierung (BAG): Die Stärke, Dauer oder Gewichtung einer Belastung wird abgewertet. Pekrun (2006) Control-Value Theory of Achievement Emotions: geringe Kontrolle	frdek8g frdek8a	1 = Ja 2 = Nein

Sie haben angegeben, dass folgende Strategien günstig sind, um Frustration beim Verfassen einer großen wissenschaftlichen Abschlussarbeit zu regulieren.

Wir bitten Sie nun einzuschätzen, **wie relevant** die einzelnen Strategien **bei der Regulation von Frustration** für das **erfolgreiche Schreiben einer großen wissenschaftlichen Abschlussarbeit** sind.

Item	Ungünstige Strategie	Antwortformat	Quelle	Variablenname	Wertelabel
Ich mache mir einen Zeitplan.		Analogskala (1= kaum relevant bis 101= sehr relevant)	angelehnt an Survivalguide Schreiben: Eine dezidierte Planung (Zeit, Ort usw.) entscheidet mit über den Erfolg von Schreibprojekten.	frdek1R	keine

Ich überlege mir, was ich brauche, um gut arbeiten zu können.		Analogskala (1= kaum relevant bis 101= sehr relevant)	Eigenkonstruktion (Planung in Bezug auf externe Ressourcen)	frdek2R	keine
Ich mache mir bewusst, wie viel von der Arbeit abhängt.	x	Analogskala (1= kaum relevant bis 101= sehr relevant)	Ungünstige Strategie angelehnt an SVF-KJ: Emotionsregulierende Bewältigung: Bagatellisierung (BAG): Die Stärke, Dauer oder Gewichtung einer Belastung wird abgewertet.	frdek3R	keine
Ich stelle an mich Eigenkonstruktion realistische Ansprüche.		Analogskala (1= kaum relevant bis 101= sehr relevant)	adaptiert nach Handbuch Lernstrategien, S. 234	frdek4R	keine
Ich verschiebe die Arbeit immer weiter nach hinten.	x	Analogskala (1= kaum relevant bis 101= sehr relevant)	Handbuch Lernstrategien, S. 234	frdek5R	keine
Ich lasse die Dinge auf mich zukommen.	x	Analogskala (1= kaum relevant bis 101= sehr relevant)	Disktraktor angelehnt an Pekrun (2006) Control-Value Theory of Achievement Emotions: geringe Kontrolle	frdek6R	keine
Ich vertraue darauf, dass mir die Aufgabe keine Schwierigkeiten bereiten wird.	x	Analogskala (1= kaum relevant bis 101= sehr relevant)	Disktraktor angelehnt an SVF-KJ: Problemlösende Bewältigung: Situationskontrolle (STK): Die Situation analysieren, Handlungen zur Kontrolle/Problemlösung planen und ausführen	frdek7R	keine
Ich mache mir bewusst, dass ich gar keine andere Wahl habe, als mich durchzubeißen.	x	Analogskala (1= kaum relevant bis 101= sehr relevant)	Ungünstige Strategie mit Bezug auf SVF-KJ: Emotionsregulierende Bewältigung: Bagatellisierung (BAG): Die Stärke, Dauer oder Gewichtung einer Belastung wird abgewertet. Pekrun (2006) Control-Value Theory of Achievement Emotions: geringe Kontrolle	frdek8R	keine

Sie haben angegeben, dass folgende Strategien nicht günstig sind, um Frustration beim Verfassen einer großen wissenschaftlichen Abschlussarbeit zu regulieren.

Wir bitten Sie nun einzuschätzen, **wie hinderlich** die einzelnen Strategien **bei der Regulation von Frustration** für das **erfolgreiche Schreiben einer großen wissenschaftlichen Abschlussarbeit** sind.

Item	Ungünstige Strategie	Antwortformat	Quelle	Variablenname	Wertelabel
Ich mache mir einen Zeitplan.		Analogskala (1= kaum hinderlich bis 101= sehr hinderlich)	angelehnt an Survivalguide Schreiben: Eine dezidierte Planung (Zeit, Ort usw.) entscheidet mit über den Erfolg von Schreibprojekten.	frdek1H	keine
Ich überlege mir, was ich brauche, um gut arbeiten zu können.		Analogskala (1= kaum hinderlich bis 101= sehr hinderlich)	Eigenkonstruktion (Planung in Bezug auf externe Ressourcen)	frdek2H	keine
Ich mache mir bewusst, wie viel von der Arbeit abhängt.	x	Analogskala (1= kaum hinderlich bis 101= sehr hinderlich)	Ungünstige Strategie angelehnt an SVF-KJ: Emotionsregulierende Bewältigung: Bagatellisierung (BAG): Die Stärke, Dauer oder Gewichtung einer Belastung wird abgewertet.	frdek3H	keine
Ich stelle an mich Eigenkonstruktion realistische Ansprüche.		Analogskala (1= kaum hinderlich bis 101= sehr hinderlich)	adaptiert nach Handbuch Lernstrategien, S. 234	frdek4H	keine
Ich verschiebe die Arbeit immer weiter nach hinten.	x	Analogskala (1= kaum hinderlich bis 101= sehr hinderlich)	Handbuch Lernstrategien, S. 234	frdek5H	keine
Ich lasse die Dinge auf mich zukommen.	x	Analogskala (1= kaum hinderlich bis 101= sehr hinderlich)	Disktraktor angelehnt an Pekrun (2006) Control-Value Theory of Achievement Emotions: geringe Kontrolle	frdek6H	keine

Ich vertraue darauf, dass mir die Aufgabe keine Schwierigkeiten bereiten wird.	x	Analogskala (1= kaum hinderlich bis 101= sehr hinderlich)	Disktraktor angelehnt an SVF-KJ: Problemlösende Bewältigung: Situationskontrolle (STK): Die Situation analysieren, Handlungen zur Kontrolle/Problemlösung planen und ausführen	frdek7H	keine
Ich mache mir bewusst, dass ich gar keine andere Wahl habe, als mich durchzubeißen.	x	Analogskala (1= kaum hinderlich bis 101= sehr hinderlich)	Ungünstige Strategie mit Bezug auf SVF-KJ: Emotionsregulierende Bewältigung: Bagatellisierung (BAG): Die Stärke, Dauer oder Gewichtung einer Belastung wird abgewertet. Pekrun (2006) Control-Value Theory of Achievement Emotions: geringe Kontrolle	frdek8H	keine

Metakognition

SCHREIBEN EINER KLEINEREN WISSENSCHAFTLICHEN ARBEIT

Denken Sie an das Erstellen einer **kleineren wissenschaftlichen Arbeit** im Studium **(z.B. Seminararbeit, schriftliche Ausarbeitung etc.).** Damit ist die Erstellung eines wissenschaftlichen Textes gemeint, für den selbstständig Literatur recherchiert und aufgearbeitet werden muss und für den es einen festgelegten Abgabetermin gibt.

Eine Befragung von Studierenden hat ergeben, dass es hierbei vor allem wichtig ist, **metakognitive Strategien anzuwenden.**

Metakognition deklarativ

Welche der folgenden Strategien sind aus Ihrer Sicht **für das erfolgreiche Erstellen einer kleineren wissenschaftlichen Arbeit günstig?**

Item	Distraktor	Antwortformat	Quelle	Variablenname	Wertelabel
Bevor ich zu arbeiten beginne, überlege ich mir einen Zeitplan.		dichotom (*ja – nein*)	angelehnt an Wild & Schiefele (1994), Planung von Lernschritten	metdek1g metdek1a	1 = Ja 2 = Nein
Ich lasse die Dinge auf mich zukommen.	x	dichotom (*ja – nein*)	Eigenkonstruktion	metdek2g metdek2a	1 = Ja 2 = Nein
Ich denke nicht viel über die Arbeit nach, bevor ich beginne.	x	dichotom (*ja – nein*)	Distraktor angelehnt an Streblow & Schiefele (2006), Auswahl geeigneter Strategien	metdek3g metdek3a	1 = Ja 2 = Nein
Ich überlege mir vorab, welche Mittel (z.B. Lernmaterial, Hilfe, etc.) ich für meine Arbeit brauche.		dichotom (*ja – nein*)	Eigenkonstruktion? (Planung externer Ressourcen)	metdek4g metdek4a	1 = Ja 2 = Nein
Ich plane vorab, wie ich am besten beim Schreiben der wissenschaftlichen Arbeit vorgehen will.		dichotom (*ja – nein*)	angelehnt an Streblow & Schiefele (2006), Auswahl geeigneter Strategien	metdek5g metdek5a	1 = Ja 2 = Nein
Ich mache keine Zeitpläne, da man sich im Endeffekt ja doch nie daran hält.	x	dichotom (*ja – nein*)	Distraktor angelehnt an Wild & Schiefele (1994), Planung von Lernschritten	metdek6g metdek6a	1 = Ja 2 = Nein

Mein einziges Ziel ist es, die Arbeit abzugeben.	x	dichotom (*ja – nein*)	Eigenkonstruktion?	metdek7g metdek7a	1 = Ja 2 = Nein
Während des Schreibens frage ich mich immer wieder, ob mein Vorgehen gerade sinnvoll ist.		dichotom (*ja – nein*)	Eigenkonstruktion	metdek8g metdek8a	1 = Ja 2 = Nein
Ich überprüfe immer wieder, ob ich das, was ich mir vorgenommen habe, auch erreiche		dichotom (*ja – nein*)	angelehnt an Wild & Schiefele (1994) Überprüfung eigener Lernfortschritte; Einschätzung der eigenen Lernziele	metdek9g metdek9a	1 = Ja 2 = Nein
Während des Arbeitens verlasse ich mich darauf, dass mein ursprünglicher Zeitplan aufgeht.	x	dichotom (*ja – nein*)	Disktraktor angelehnt an Wild & Schiefele (1994) Überprüfung eigener Lernfortschritte; Einschätzung der eigenen Lernziele	metdekAg metdekAa	1 = Ja 2 = Nein
Wenn ich merke, dass ich nicht weiterkomme, ändere ich meine Vorgehensweise.		dichotom (*ja – nein*)	angelehnt an Wild & Schiefele (1994) adaptive Regulation durch Änderung des Lernverhaltens	metdekBg metdekBa	1 = Ja 2 = Nein
Wenn ich merke, dass meine Ziele unrealistisch sind, überdenke ich sie nochmal.		dichotom (*ja – nein*)	Eigenkonstruktion	metdekCg metdekCa	1 = Ja 2 = Nein
Ich halte unter allen Umständen an meinen ursprünglichen Zielen fest.	x	dichotom (*ja – nein*)	Distraktor Eigenkonstruktion	metdekDg metdekDa	1 = Ja 2 = Nein
Ich überlege mir vor dem Schreiben, welche Ziele ich für die Arbeit habe.		dichotom (*ja – nein*)	angelehnt an Streblow & Schiefele (2006), Formulierung von Lernzielen	metdekEg metdekEa	1 = Ja 2 = Nein

Sie haben angegeben, dass folgende Strategien für das erfolgreiche Erstellen einer kleineren wissenschaftlichen Arbeit günstig sind.

Wir bitten Sie nun einzuschätzen, **wie relevant** die einzelnen Strategien **für das erfolgreiche Erstellen einer kleineren wissenschaftlichen Arbeit** sind.

Item	Ungünstige Strategie	Antwortformat	Quelle	Variablenname	Wertelabel
Bevor ich zu arbeiten beginne, überlege ich mir einen Zeitplan.		Analogskala (1= kaum relevant bis 101= sehr relevant)	angelehnt an Wild & Schiefele (1994), Planung von Lernschritten	metdek1R	keine
Ich lasse die Dinge auf mich zukommen.	x	Analogskala (1= kaum relevant bis 101= sehr relevant)	Eigenkonstruktion	metdek2R	keine
Ich denke nicht viel über die Arbeit nach, bevor ich beginne.	x	Analogskala (1= kaum relevant bis 101= sehr relevant)	Ungünstige Strategie angelehnt an Streblow & Schiefele (2006), Auswahl geeigneter Strategien	metdek3R	keine
Ich überlege mir vorab, welche Mittel (z.B. Lernmaterial, Hilfe, etc.) ich für meine Arbeit brauche.		Analogskala (1= kaum relevant bis 101= sehr relevant)	Eigenkonstruktion? (Planung externer Ressourcen)	metdek4R	keine
Ich plane vorab, wie ich am besten beim Schreiben der wissenschaftlichen Arbeit vorgehen will.		Analogskala (1= kaum relevant bis 101= sehr relevant)	angelehnt an Streblow & Schiefele (2006), Auswahl geeigneter Strategien	metdek5R	keine
Ich mache keine Zeitpläne, da man sich im Endeffekt ja doch nie daran hält.	x	Analogskala (1= kaum relevant bis 101= sehr relevant)	Ungünstige Strategie angelehnt an Wild & Schiefele (1994), Planung von Lernschritten	metdek6R	keine
Mein einziges Ziel ist es, die Arbeit abzugeben.	x	Analogskala (1= kaum relevant bis 101= sehr relevant)	Eigenkonstruktion?	metdek7R	keine

Während des Schreibens frage ich mich immer wieder, ob mein Vorgehen gerade sinnvoll ist.		Analogskala (1= kaum relevant bis 101= sehr relevant)	Eigenkonstruktion	metdek8R	keine
Ich überprüfe immer wieder, ob ich das, was ich mir vorgenommen habe, auch erreiche		Analogskala (1= kaum relevant bis 101= sehr relevant)	angelehnt an Wild & Schiefele (1994) Überprüfung eigener Lernfortschritte; Einschätzung der eigenen Lernziele	metdek9R	keine
Während des Arbeitens verlasse ich mich darauf, dass mein ursprünglicher Zeitplan aufgeht.	x	Analogskala (1= kaum relevant bis 101= sehr relevant)	Disktraktor angelehnt an Wild & Schiefele (1994) Überprüfung eigener Lernfortschritte; Einschätzung der eigenen Lernziele	metdekAR	keine
Wenn ich merke, dass ich nicht weiterkomme, ändere ich meine Vorgehensweise.		Analogskala (1= kaum relevant bis 101= sehr relevant)	angelehnt an Wild & Schiefele (1994) adaptive Regulation durch Änderung des Lernverhaltens	metdekBR	keine
Wenn ich merke, dass meine Ziele unrealistisch sind, überdenke ich sie nochmal.		Analogskala (1= kaum relevant bis 101= sehr relevant)	Eigenkonstruktion	metdekCR	keine
Ich halte unter allen Umständen an meinen ursprünglichen Zielen fest.	x	Analogskala (1= kaum relevant bis 101= sehr relevant)	Ungünstige Strategie Eigenkonstruktion	metdekDR	keine
Ich überlege mir vor dem Schreiben, welche Ziele ich für die Arbeit habe.		Analogskala (1= kaum relevant bis 101= sehr relevant)	angelehnt an Streblow & Schiefele (2006), Formulierung von Lernzielen	metdekER	keine

Sie haben angegeben, dass folgende Strategien nicht günstig sind, um beim Schreiben einer kleineren wissenschaftlichen Arbeit erfolgreich zu sein.

Wir bitten Sie nun einzuschätzen, **wie hinderlich** die einzelnen Strategien **für das erfolgreiche Erstellen einer kleineren wissenschaftlichen Arbeit** sind.

Item	Ungünstige Strategie	Antwortformat	Quelle	Variablenname	Wertelabel
Bevor ich zu arbeiten beginne, überlege ich mir einen Zeitplan.		Analogskala (1= kaum hinderlich bis 101= sehr hinderlich)	angelehnt an Wild & Schiefele (1994), Planung von Lernschritten	metdek1H	keine
Ich lasse die Dinge auf mich zukommen.	x	Analogskala (1= kaum hinderlich bis 101= sehr hinderlich)	Eigenkonstruktion	metdek2H	keine
Ich denke nicht viel über die Arbeit nach, bevor ich beginne.	x	Analogskala (1= kaum hinderlich bis 101= sehr hinderlich)	Ungünstige Strategie angelehnt an Streblow & Schiefele (2006), Auswahl geeigneter Strategien	metdek3H	keine
Ich überlege mir vorab, welche Mittel (z.B. Lernmaterial, Hilfe, etc.) ich für meine Arbeit brauche.		Analogskala (1= kaum hinderlich bis 101= sehr hinderlich)	Eigenkonstruktion? (Planung externer Ressourcen)	metdek4H	keine
Ich plane vorab, wie ich am besten beim Schreiben der wissenschaftlichen Arbeit vorgehen will.		Analogskala (1= kaum hinderlich bis 101= sehr hinderlich)	angelehnt an Streblow & Schiefele (2006), Auswahl geeigneter Strategien	metdek5H	keine
Ich mache keine Zeitpläne, da man sich im Endeffekt ja doch nie daran hält.	x	Analogskala (1= kaum hinderlich bis 101= sehr hinderlich)	Ungünstige Strategie angelehnt an Wild & Schiefele (1994), Planung von Lernschritten	metdek6H	keine
Mein einziges Ziel ist es, die Arbeit abzugeben.	x	Analogskala (1= kaum hinderlich bis 101= sehr hinderlich)	Eigenkonstruktion?	metdek7H	keine

Während des Schreibens frage ich mich immer wieder, ob mein Vorgehen gerade sinnvoll ist.		Analogskala (1= kaum hinderlich bis 101= sehr hinderlich)	Eigenkonstruktion	metdek8H	keine
Ich überprüfe immer wieder, ob ich das, was ich mir vorgenommen habe, auch erreiche		Analogskala (1= kaum hinderlich bis 101= sehr hinderlich)	angelehnt an Wild & Schiefele (1994) Überprüfung eigener Lernfortschritte; Einschätzung der eigenen Lernziele	metdek9H	keine
Während des Arbeitens verlasse ich mich darauf, dass mein ursprünglicher Zeitplan aufgeht.	x	Analogskala (1= kaum hinderlich bis 101= sehr hinderlich)	Disktraktor angelehnt an Wild & Schiefele (1994) Überprüfung eigener Lernfortschritte; Einschätzung der eigenen Lernziele	metdekAH	keine
Wenn ich merke, dass ich nicht weiterkomme, ändere ich meine Vorgehensweise.		Analogskala (1= kaum hinderlich bis 101= sehr hinderlich)	angelehnt an Wild & Schiefele (1994) adaptive Regulation durch Änderung des Lernverhaltens	metdekBH	keine
Wenn ich merke, dass meine Ziele unrealistisch sind, überdenke ich sie nochmal.		Analogskala (1= kaum hinderlich bis 101= sehr hinderlich)	Eigenkonstruktion	metdekCH	keine
Ich halte unter allen Umständen an meinen ursprünglichen Zielen fest.	x	Analogskala (1= kaum hinderlich bis 101= sehr hinderlich)	Ungünstige Strategie Eigenkonstruktion	metdekDH	keine
Ich überlege mir vor dem Schreiben, welche Ziele ich für die Arbeit habe.		Analogskala (1= kaum hinderlich bis 101= sehr hinderlich)	angelehnt an Streblow & Schiefele (2006), Formulierung von Lernzielen	metdekEH	keine

Kognition

VORBEREITUNG AUF EINE PRÜFUNG

Denken Sie an eine **Prüfungsvorbereitung.** Dabei müssen Studierende große Stoffmengen selbständig aufbereiten und lernen, wobei das Verständnis des Lernstoffs wesentlich ist.

Eine Befragung von Studierenden hat ergeben, dass hierbei vor allem **kognitive Lernstrategien** wichtig sind.

Welche der folgenden Strategien sind aus Ihrer Sicht **für die erfolgreiche Vorbereitung auf eine Prüfung günstig?**

Item	Ungünstige Strategie	Antwortformat	Quelle	Variablenname	Wertelabel
Ich strukturiere den zu lernenden Stoff.		dichotom (*ja – nein*)	angelehnt an Streblow & Schiefele (2006): Kernaussagen aus Texten herausarbeiten Inhalte neu strukturieren LIST: „ Für größere Stoffmengen fertige ich eine Gliederung an, die die Struktur des Stoffs am besten wiedergibt.“	kodek1g kodek1a	1 = Ja 2 = Nein
Ich beschäftige mich tiefgehend und ausführlich mit dem Stoff.		dichotom (*ja – nein*)	Tiefenstrategien – Zitat allgemein	kodek2g kodek2a	1 = Ja 2 = Nein
Ich fange direkt an zu lernen, der Rest ergibt sich.	x	dichotom (*ja – nein*)	Ungünstige Strategie Eigenkonstruktion	kodek3g kodek3a	1 = Ja 2 = Nein
Ich wiederhole den Stoff.		dichotom (*ja – nein*)	angelehnt an Wild & Schiefele (1994) aktives Wiederholen einzelner Fakten, wiederholte Durcharbeitung und Streblow & Schiefele (2006): Lernstoff immer wieder laut oder leise aufsagen	kodek4g kodek4a	1 = Ja 2 = Nein

Ich lerne eine Zusammenfassung.	x	dichotom (*ja – nein*)	Ungünstige Strategie Eigenkonstruktion	kodek5g kodek5a	1 = Ja 2 = Nein
Ich lese mir nur einmal das gesamte Skript durch.	x	dichotom (*ja – nein*)	Ungünstige Strategie Eigenkonstruktion	kodek6g kodek6a	1 = Ja 2 = Nein
Ich lerne den Stoff auswendig.	x	dichotom (*ja – nein*)	Ungünstige Strategie Eigenkonstruktion	kodek7g kodek7a	1 = Ja 2 = Nein

Sie haben angegeben, dass folgende Strategien für die erfolgreiche Vorbereitung auf eine Prüfung günstig sind. Wir bitten Sie nun einzuschätzen, **wie relevant** die einzelnen Strategien **für die erfolgreiche Vorbereitung auf eine Prüfung** sind.

Item	Ungünstige Strategie	Antwortformat	Quelle	Variablenname	Wertelabel
Ich strukturiere den zu lernenden Stoff.		Analogskala (1= kaum relevant bis 101= sehr relevant)	angelehnt an Streblow & Schiefele (2006): Kernaussagen aus Texten herausarbeiten Inhalte neu strukturieren LIST: „ Für größere Stoffmengen fertige ich eine Gliederung an, die die Struktur des Stoffs am besten wiedergibt.“	kodek1R	keine
Ich beschäftige mich tiefgehend und ausführlich mit dem Stoff.		Analogskala (1= kaum relevant bis 101= sehr relevant)	Tiefenstrategien – Zitat allgemein	kodek2R	keine
Ich fange direkt an zu lernen, der Rest ergibt sich.	x	Analogskala (1= kaum relevant bis 101= sehr relevant)	Ungünstige Strategie Eigenkonstruktion	kodek3R	keine
Ich wiederhole den Stoff.		Analogskala (1= kaum relevant bis 101= sehr relevant)	angelehnt an Wild & Schiefele (1994) aktives Wiederholen einzelner Fakten, wiederholte Durcharbeitung und Streblow & Schiefele (2006): Lernstoff immer wieder laut oder leise aufsagen	kodek4R	keine
Ich lerne eine Zusammenfassung.	x	Analogskala (1= kaum relevant bis 101= sehr relevant)	Ungünstige Strategie Eigenkonstruktion	kodek5R	keine

Ich lese mir nur einmal das gesamte Skript durch.	x	Analogskala (1= kaum relevant bis 101= sehr relevant)	Ungünstige Strategie Eigenkonstruktion	kodek6R	keine
Ich lerne den Stoff auswendig.	x	Analogskala (1= kaum relevant bis 101= sehr relevant)	Ungünstige Strategie Eigenkonstruktion	kodek7R	keine

Sie haben angegeben, dass folgende Strategien nicht günstig für die erfolgreiche Vorbereitung auf eine Prüfung sind.

Wir bitten Sie nun einzuschätzen, **wie hinderlich** die einzelnen Strategien **für die erfolgreiche Vorbereitung auf eine Prüfung** sind.

Item	Ungünstige Strategie	Antwortformat	Quelle	Variablenname	Wertelabel
Ich strukturiere den zu lernenden Stoff.		Analogskala (1= kaum hinderlich bis 101= sehr hinderlich)	angelehnt an Streblow & Schiefele (2006): Kernaussagen aus Texten herausarbeiten Inhalte neu strukturieren LIST: „ Für größere Stoffmengen fertige ich eine Gliederung an, die die Struktur des Stoffs am besten wiedergibt.“	kodek1H	keine
Ich beschäftige mich tiefgehend und ausführlich mit dem Stoff.		Analogskala (1= kaum hinderlich bis 101= sehr hinderlich)	Tiefenstrategien – Zitat allgemein	kodek2H	keine
Ich fange direkt an zu lernen, der Rest ergibt sich.	x	Analogskala (1= kaum hinderlich bis 101= sehr hinderlich)	Ungünstige Strategie Eigenkonstruktion	kodek3H	keine
Ich wiederhole den Stoff.		Analogskala (1= kaum hinderlich bis 101= sehr hinderlich)	angelehnt an Wild & Schiefele (1994) aktives Wiederholen einzelner Fakten, wiederholte Durcharbeitung und Streblow & Schiefele (2006): Lernstoff immer wieder laut oder leise aufsagen	kodek4H	keine

Ich lerne eine Zusammenfassung.	x	Analogskala (1= kaum hinderlich bis 101= sehr hinderlich)	Ungünstige Strategie Eigenkonstruktion	kodek5H	keine
Ich lese mir nur einmal das gesamte Skript durch.	x	Analogskala (1= kaum hinderlich bis 101= sehr hinderlich)	Ungünstige Strategie Eigenkonstruktion	kodek6H	keine
Ich lerne den Stoff auswendig.	x	Analogskala (1= kaum hinderlich bis 101= sehr hinderlich)	Ungünstige Strategie Eigenkonstruktion	kodek7H	keine

Langeweile

LANGEWEILE IN EINER VORLESUNG

Denken Sie an den Besuch einer **Vorlesung.** Damit ist eine Lehrveranstaltung mit einer größeren Anzahl von Studierenden gemeint, in der große Stoffmengen durch eine/n Dozierende/n vorgetragen werden.

Eine Befragung von Studierenden hat ergeben, dass hierbei vor allem die **Regulation von Langeweile** von Bedeutung ist.

Langeweile deklarativ

Welche der folgenden Strategien sind aus Ihrer Sicht bei der **Regulation von Langeweile für einen ergiebigen Vorlesungsbesuch günstig?**

Item	Ungünstige Strategie	Antwortformat	Quelle	Variablenname	Wertelabel
Ich mache mir bewusst, welchen Nutzen ich aus der Vorlesung ziehen kann.		dichotom (*ja* – *nein*)	Subjektiv höher extrinsisch valent uminterpretieren; Götz (2007)	ladek1g ladek1a	1 = Ja 2 = Nein
Ich bereite mich auf die Vorlesung vor.		dichotom (*ja* – *nein*)	Eigenkonstruktion angelehnt an Götz (2007): Aktivierung metakognitiver Strategie; Coping	ladek2g ladek2a	1 = Ja 2 = Nein
Ich mache mir bewusst, dass ich als Studierende/r Vorlesungen besuchen muss.	x	dichotom (*ja* – *nein*)	Eigenkonstruktion nach Pekrun (2006) Control-Value Theory of Achievement Emotions: geringer Wert	ladek3g ladek3a	1 = Ja 2 = Nein
Ich beteilige mich aktiv an der Vorlesung.		dichotom (*ja* – *nein*)	adaptiert nach Widulle (2009) Frontalunterricht aktiv nutzen	ladek4g ladek4a	1 = Ja 2 = Nein
Ich lenke mich während der Vorlesung ab.	x	dichotom (*ja* – *nein*)	Avoidance, Nett (2010)	ladek5g ladek5a	1 = Ja 2 = Nein
Ich gehe ohne große Erwartungen in die Vorlesung.	x	dichotom (*ja* – *nein*)	Eigenkonstruktion nach Pekrun (2006) Control-Value Theory of Achievement Emotions: geringer Wert	ladek6g ladek6a	1 = Ja 2 = Nein

Sie haben angegeben, dass folgende Strategien günstig sind, um Langeweile bei einer Vorlesung zu regulieren. Wir bitten Sie nun einzuschätzen, **wie relevant** die einzelnen Strategien **bei der Regulation von Langeweile** für einen **ergiebigen Vorlesungsbesuch** sind.

Item	Ungünstige Strategie	Antwortformat	Quelle	Variablenname	Wertelabel
Ich mache mir bewusst, welchen Nutzen ich aus der Vorlesung ziehen kann.		Analogskala (1= kaum relevant bis 101= sehr relevant)	Subjektiv höher extrinsisch valent uminterpretieren; Götz (2007)	ladek1R	keine
Ich bereite mich auf die Vorlesung vor.		Analogskala (1= kaum relevant bis 101= sehr relevant)	Eigenkonstruktion angelehnt an Götz (2007): Aktivierung metakognitiver Strategie; Coping	ladek2R	keine
Ich mache mir bewusst, dass ich als Studierende/r Vorlesungen besuchen muss.	x	Analogskala (1= kaum relevant bis 101= sehr relevant)	Eigenkonstruktion nach Pekrun (2006) Control-Value Theory of Achievement Emotions: geringer Wert	ladek4R	keine
Ich beteilige mich aktiv an der Vorlesung.		Analogskala (1= kaum relevant bis 101= sehr relevant)	adaptiert nach Widulle (2009) Frontalunterricht aktiv nutzen	ladek5R	keine
Ich lenke mich während der Vorlesung ab.	x	Analogskala (1= kaum relevant bis 101= sehr relevant)	Avoidance, Nett (2010)	ladek6R	keine
Ich gehe ohne große Erwartungen in die Vorlesung.	x	Analogskala (1= kaum relevant bis 101= sehr relevant)	Eigenkonstruktion nach Pekrun (2006) Control-Value Theory of Achievement Emotions: geringer Wert	ladek7R	keine

Sie haben angegeben, dass folgende Strategien nicht günstig sind, um Langeweile bei einer Vorlesung zu regulieren.

Wir bitten Sie nun einzuschätzen, **wie hinderlich** die einzelnen Strategien **bei der Regulation von Langeweile** für einen **ergiebigen Vorlesungsbesuch** sind.

Item	Ungünstige Strategie	Antwortformat	Quelle	Variablenname	Wertelabel
Ich mache mir bewusst, welchen Nutzen ich aus der Vorlesung ziehen kann.		Analogskala (1= kaum hinderlich bis 101= sehr hinderlich)	Subjektiv höher extrinsisch valent uminterpretieren; Götz (2007)	ladek1H	keine
Ich bereite mich auf die Vorlesung vor.		Analogskala (1= kaum hinderlich bis 101= sehr hinderlich)	Eigenkonstruktion angelehnt an Götz (2007): Aktivierung metakognitiver Strategie; Coping	ladek2H	keine
Ich mache mir bewusst, dass ich als Studierende/r Vorlesungen besuchen muss.	x	Analogskala (1= kaum hinderlich bis 101= sehr hinderlich)	Eigenkonstruktion nach Pekrun (2006) Control-Value Theory of Achievement Emotions: geringer Wert	ladek4H	keine
Ich beteilige mich aktiv an der Vorlesung.		Analogskala (1= kaum hinderlich bis 101= sehr hinderlich)	adaptiert nach Widulle (2009) Frontalunterricht aktiv nutzen	ladek5H	keine
Ich lenke mich während der Vorlesung ab.	x	Analogskala (1= kaum hinderlich bis 101= sehr hinderlich)	Avoidance, Nett (2010)	ladek6H	keine
Ich gehe ohne große Erwartungen in die Vorlesung.	x	Analogskala (1= kaum hinderlich bis 101= sehr hinderlich)	Eigenkonstruktion nach Pekrun (2006) Control-Value Theory of Achievement Emotions: geringer Wert	ladek7H	keine

Literature (see Appendix B)

Appendix E

Table E

Frequencies of rating per item and category in rating

Items	Categories *beneficial*	*adverse*	P_{ij}
Metacognitive strategies in writing a minor scientific work			
Before I begin to work, I set a timetable.	10	1	0.818
I take things as they come.	0	11	1.000
I don't think a lot about the work before I begin.	0	11	1.000
I think in advance about which recourses (e.g. learning material, help, etc.) I need for my work.	9	2	0.673
I plan in advance about how to proceed best in writing the scientific work.	10	1	0.818
I don‘t set timetables because eventually, I won‘t adhere to it anyway.	0	11	1.000
My only goal is to hand in the work.	0	11	1.000
During writing I regularly ask myself if my proceeding is useful.	9	2	0.673
I regularly check if I achieve what I had intended.	11	0	1.000
During working I count on my initial timetable to work out.	0	11	1.000
When I notice that I don't proceed, I change my approach.	11	0	1.000
When I notice that my goals are unrealistic, I reconsider them.	10	1	0.818
I stick to my initial goals by all means.	0	11	1.000
Before writing I consider which goals I have for the work.	11	0	1.000
Cognitive strategies in exam preparation			
I structure the study matter.	10	1	0.818
I follow up the subject matter thoroughly.	11	0	1.000
I start with learning directly. Everything else will follow.	0	11	1.000
I repeat the subject matter.	10	1	0.818
I study a summary.	2	9	0.673
I read through the lecture notes only once.	0	11	1.000
I learn the subject matter by heart.	3	8	0.564
Regulation of boredom in lecture attendance			
I make myself aware of the benefits I can extract from the lecture.	11	0	1.000
I prepare for the lecture.	9	2	0.673
I make myself aware that as a student, I have to attend to lectures.	2	9	0.673
I engage actively in the lecture.	11	0	1.000
I distract myself during the lecture.	1	10	0.818
I go to lectures without high expectations.	1	10	0.818
Regulation of frustration in writing a major scientific work			
I set a timetable.	11	0	1.000
I consider what I need to work well.	10	1	0.818
I make myself aware of how much is depending on the work.	1	10	0.818
I set realistic aspirations.	10	1	0.818
I postpone the work again and again.	0	11	1.000
I take things as they come.	0	11	1.000
I have confidence in mastering the task without difficulties.	3	8	0.564
I make myself aware of the fact that I don't have a choice other than being tenacious.	2	9	0.673
Total	189	196	0.867

Note. N = number of items = 35; n = number of raters = 11; k = number of categories = 2; κ = .73; SE(κ) = 0.023; κ/SE(κ) = .73/.023 = 23.10; CI(κ) = [.69 - .78].

Appendix F

Table F1

Factor loadings from CFA for metacognitive strategies in writing a minor scientific work

Items	λ	λ^2	θ
Before I begin to work, I set a timetable.	.483	.233	.767
I take things as they come.	.369	.136	.864
I don't think a lot about the work before I begin.	.557	.310	.690
I think in advance about which recourses (e.g. learning material, help, etc.) I need for my work.	.112	.013	.987
I plan in advance about how to proceed best in writing the scientific work.	.581	.338	.662
I don‘t set timetables because eventually, I won‘t adhere to it anyway.	.744	.554	.446
My only goal is to hand in the work.	.247	.061	.939
During writing I regularly ask myself if my proceeding is useful.	.191	.036	.964
I regularly check if I achieve what I had intended.	.297	.088	.912
During working I count on my initial timetable to work	.148	.022	.978
When I notice that I don't proceed, I change my	.178	.032	.968
When I notice that my goals are unrealistic, I reconsider them.	.340	.116	.884
I stick to my initial goals by all means.	.036	.001	.999
Before writing I consider which goals I have for the work.	.508	.258	.742

Note. N= 158; λ = standardized factor loading; θ = unique variance of item; CR = composite reliability = .66; VE = Variance extracted = .16.

Table F2

Factor loadings from CFA for cognitive strategies in exam preparation (all items)

Items	λ	λ^2	θ
I structure the study matter.	.483	.233	.767
I follow up the subject matter thoroughly.	.368	.135	.865
I start with learning directly. Everything else will	.377	.142	.858
I repeat the subject matter.	.966	.933	.067
I study a summary.	.157	.025	.975
I read through the lecture notes only once.	.604	.365	.635
I learn the subject matter by heart.	-.245	.060	.940

Note. N= 144; λ = standardized factor loading; θ = unique variance of item; CR = composite reliability = .59; VE = Variance extracted = .27.

Table F3

Factor loadings from CFA for cognitive strategies in exam preparation (item 6 exluded)

Items	λ	λ^2	θ
I structure the study matter.	.483	.233	.767
I follow up the subject matter thoroughly.	.358	.128	.872
I start with learning directly. Everything else will follow.	.436	.190	.810
I repeat the subject matter.	.924	.854	.146
I study a summary.	.125	.016	.984
I read through the lecture notes only once.	.608	.370	.630

Note. N = 144; λ = standardized factor loading; θ = unique variance of item; CR = composite reliability = .67; VE = Variance extracted = .30.

Table F4

Factor loadings from CFA for boredom regulation strategies in lecture attendance

Items	λ	λ^2	θ
I make myself aware of the benefits I can extract from the lecture.	.483	.233	.767
I prepare for the lecture.	.198	.039	.961
I make myself aware that as a student, I have to attend to lectures.	.221	.049	.951
I engage actively in the lecture.	.372	.138	.862
I distract myself during the lecture.	.781	.610	.390
I go to lectures without high expectations.	.664	.441	.559

Note. *N* = 109; λ = standardized factor loading; θ = unique variance of item; CR = composite reliability = .62; VE = Variance extracted = .25.

Table F5

Factor loadings from CFA for frustration regulation strategies in lecture attendance

Items	λ	λ^2	θ
I set a timetable.	.483	.233	.767
I consider what I need to work well.	.299	.089	.911
I make myself aware of how much is depending on the work.	.166	.028	.972
I set realistic aspirations.	.194	.038	.962
I postpone the work again and again.	.740	.548	.452
I take things as they come.	.791	.626	.374
I have confidence in mastering the task without	.396	.157	.843
I make myself aware of the fact that I don't have a choice other than being tenacious.	.483	.233	.767

Note. *N* = 76; λ = standardized factor loading; θ = unique variance of item; CR = composite reliability = .68; VE = Variance extracted = .24.

Curriculum **V**itae

Armin Schmolmüller, BSc

born on September 15th 1988

Education

09/1995 – 06/1999	Elementary School Kefermarkt
09/1999 – 06/2007	Grammar School (Bundesgymnasium/Bundesrealgymnasium) Freistadt
11/2007 – 07/2008	Civil Service at Volkshilfe Freistadt
since 10/2008	University of Vienna: Diploma Degree Study of Psychology
09/2011 – 07/2012	Universidad de Sevilla: Erasmus
20.02.2013	University of Vienna: Bachelor of Science (BSc) in Psychology
since 10/2013	University of Vienna : Master course in Applied Psychology
since 03/2014	University of Vienna : Master course in German Philology
02/2015 – 06/2015	6-weeks Internship at the department of Educational Psychology and Evaluation (University of Vienna)
03/2015 – 01/2016	Research assistance at the department of Educational Psychology and Evaluation (University of Vienna)
07/2015 – 09/2015	Research assistance in the project PRO-SRL
07/2015 – 09/2015	Research assistance in the project “Gendersensible Pädagogik im Elementarbereich”

Eidestattliche Erklärung

Hiermit erkläre ich, dass ich die vorliegende Masterarbeit selbständig und ohne Verwendung unerlaubter Hilfsmittel verfasst habe. Alle Inhalte, die direkt oder indirekt aus fremden Quellen entnommen wurden, sind durch entsprechende Quellenangaben gekennzeichnet.

Wien, am 21.04.2016 Armin Schmolmüller