

Secondary Publication



Pfost, Maximilian; Hübner, Vanessa

Assessment of Strategic Knowledge of Learning From Errors in Higher Education

Date of secondary publication: 20.05.2025

Version of Record (Published Version), Article

Persistent identifier: urn:nbn:de:bvb:473-irb-1083360

Primary publication

Pfost, Maximilian; Hübner, Vanessa (2025): Assessment of Strategic Knowledge of Learning From Errors in Higher Education, in: Diagnostica : Zeitschrift für psychologische Diagnostik und differentielle Psychologie, zugleich Informationsorgan über psychologische Tests und Untersuchungsmethoden, Göttingen: Hogrefe Verlag, Vol. 71, Nr. 2, pp. 53–63, doi: 10.1026/0012-1924/a000341.

Legal Notice

This work is protected by copyright and/or the indication of a licence. You are free to use this work in any way permitted by the copyright and/or the licence that applies to your usage. For other uses, you must obtain permission from the rights-holders.

This document is made available under a Creative Commons license.



The license information is available online:

<https://creativecommons.org/licenses/by/4.0/legalcode>



Assessment of Strategic Knowledge of Learning From Errors in Higher Education

Maximilian Pfost  and Vanessa Hübner 

Department of Educational Research, Faculty of Human Sciences and Education, University of Bamberg, Germany

Abstract: In higher education settings, students are confronted with learning materials and topics of great difficulty. Therefore, learning from errors and dealing with them in an adaptive way tends to be important. Besides facets such as error culture, students' learning in such situations might differ due to individual differences in strategic knowledge of dealing with errors while learning. In this study, we developed a scenario-based test that aims to measure strategic knowledge of learning from errors in higher education settings. Besides cognitive-behavioral strategies, affective-motivational regulation strategies were considered and a standard of comparison was derived through expert rating. The article addresses the evaluation of the psychometric properties of this test in German. Analyses relied on a sample of 178 university students. Findings showed good signs of unidimensionality. Construct reliability and test difficulty tended to be lower than optimal. Relationships with students' tendencies to deal adaptively with errors in learning were also explored. Finally, research desiderata and areas of application are discussed.

Keywords: learning from errors, strategic knowledge, higher education, test development

Entwicklung eines Messverfahrens zur Erfassung des strategischen Wissens von Studierenden im Umgang mit Fehlern beim Lernen

Zusammenfassung: Gerade für das Hochschulstudium ist die Vermittlung komplexer und schwieriger Inhalte kennzeichnend, weshalb ein lernförderlicher Umgang mit Fehlern für Studierende bedeutsam erscheint. Neben Aspekten wie beispielsweise einer Fehlerkultur ist anzunehmen, dass auch das strategische Wissen im Umgang mit Fehlern bedeutsam ist. Um dies messbar zu machen, haben wir ein szenariobasiertes Testverfahren entwickelt, das auf die Erfassung des strategischen Wissens Studierender im Umgang mit Fehlern beim Lernen abzielt. Die Strategien umfassen dabei nicht nur kognitiv-behaviorale, sondern auch affektiv-motivationale Aspekte. Die Qualität der Strategien wurde über ein Expertenrating erfasst. Ziel des Beitrags ist die Analyse psychometrischer Eigenschaften dieses deutschsprachigen Testverfahrens. Die Ergebnisse, basierend auf den Daten von 178 Studierenden, verweisen auf die Eindimensionalität des Testverfahrens bei eingeschränkt akzeptabler Konstruktreliabilität. Der Test zeigte sich von relativ leichter Bearbeitungsschwierigkeit. In der Studie wurden ferner Zusammenhänge mit Skalen zum adaptiven Umgang mit Fehlern bei Lernen exploriert. Schließlich wird ein Ausblick auf mögliche Anwendungen in der Forschung gegeben.

Schlüsselwörter: Lernen aus Fehlern, Lernstrategiewissen, Hochschulbildung, Testentwicklung

Learners, especially in higher education settings, often are required or they even actively seek to leave their comfort zone of assured knowledge and to concentrate on tasks that are of great difficulty in lectures and seminars (Hübner & Pfost, 2023). Therefore, the occurrence of errors while learning is not only expected but tends to be omnipresent. In this paper, we consider errors as unintended deficient deviations from plans and goals or desired outcomes of a learning task (cf., Frese & Keith, 2015). For example, in the field of higher education, a student may have received a poor grade on a term paper due to an omission of a central theoretical framework. From a perspective of the regulation of learning tasks, errors are often accompanied by negative feedback, which

induces the learner to take regulatory actions with regard to their learning behavior, motivation, and learning goals (Lipnevich et al., 2016; Tulis et al., 2016).

Prior studies have provided evidence that engaging in error-prone learning tasks and being confronted with such negative feedback have the potential to improve one's learning (Ivancic & Hesketh, 2000; Metcalfe, 2017). When errors are detected, they offer feedback information important for the regulation of one's learning, which ultimately may contribute to students' learning and skill development (Zhang & Fiorella, 2023). However, learning success may depend on various task and learner characteristics. For example, higher learning success was observed when errors were more severe and made by the

learner themselves (Horvath et al., 2021; Keith et al., 2020). In addition, the social context seems important, as errors often occur in interaction with instructors and peers (Spychiger et al., 2006; Steuer & Dresel, 2015). Finally, individuals require specific strategies for handling such error-prone learning situations. This will be discussed next.

Metacognition and Strategy Knowledge

Complex and challenging learning tasks require a situation-adequate use of learning strategies. Learning strategies refer to the orderly sequence of learning actions used to meet specific learning demands (Artelt et al., 2010). Empirical research has shown that a higher self-reported use of learning strategies such as elaboration, metacognition, or effort regulation tends to be positively related to learning outcomes (Pintrich et al., 1993; Richardson et al., 2012). Nevertheless, doubts have been raised about this simplified perspective of “the more, the better” (sometimes labeled the *maximum view*), especially when operationalized via situation-unspecific standardized self-report measures of frequency of use (Artelt & Neuenhaus, 2010; Spörer & Brunstein, 2006). In other words, in order to be successful, strategies need to be selected and applied dependent on further characteristics such as the learner’s goals and specific task requirements (*optimum view*; Wirth & Leutner, 2008). This means that learners need to coordinate different bases of knowledge for successful strategic learning: Besides the fact that students need to be aware of the existence of a specific strategy (declarative strategy knowledge), they also need to know how to apply this specific strategy in a given situation (procedural strategy knowledge). Finally, sufficient knowledge of the effectivity of a specific strategy depending on situation-specific task affordances, of its rationale for execution, and of the effectivity of alternatives within this learning situation is important (conditional strategy knowledge; Paris et al., 1983). Therefore, measures that apply a qualitative standard to evaluate the appropriateness of specific strategies for successfully completing a particular learning task are required. In order to reach this aim, the following procedure is often applied (Karlen, 2017; Wirth & Leutner, 2008): First, a group of scenarios that relate to a specific domain or learning situation (e.g., the emergence of an error while learning) is specified. In addition, a list of different strategies is developed for each scenario, which vary in their usefulness based on the specific learning requirements of the situation. Whether a strategy is considered appropriate or not is based on an expert rating. To summarize such ratings for a standard of comparison, the individual expert rating of each strategy in a specific scenario is compared with the rating of another strategy.

Consequently, each strategy can be rated as being more useful or less useful than another strategy. In the second step, the students are asked to rate the usefulness of the strategies in each scenario, and again each strategy is evaluated as being superior or inferior to another strategy. In sum, by using the expert rating as a standard of comparison, it is possible to represent the quality of a learner’s individual strategy knowledge numerically: The higher the correspondence between the learner’s and the experts’ ratings, the higher the learner’s strategy knowledge score. The use of such a clear-to-interpret benchmark can be seen as a substantial advantage in comparison with other measures, for example, questionnaires or self-reports that are available for the assessment of metacognitive knowledge (Händel et al., 2013). To date, such tests have been developed with a focus on assessing children’s general metacognitive strategy knowledge at the primary and secondary school level (e.g., Händel et al., 2013; Lockl et al., 2016) or, for example, domain-specific metacognitive strategy knowledge such as reading (e.g., Artelt & Schneider, 2015). For learners in higher education settings, tests are available for metacognitive strategy knowledge in academic writing (Karlen, 2017) or for foreign students in linguistically difficult situations (Seeger et al., 2021). To our knowledge, tests for the strategic handling of errors in learning situations have not yet been developed.

Strategies and Learning From Errors

Within a process framework of self-regulated learning (e.g., Zimmerman & Moylan, 2009), reasoning about errors may be located in the self-reflection phase. Intended learning goals from the forethought phase have not been reached by learning actions in the performance phase. Therefore, processes of reflection and evaluation that consciously and unconsciously address foregoing as well as future goal setting, planning, and action are expected to occur. Furthermore, with regard to component models of self-regulated learning (e.g., Boekaerts, 1999), errors may be a starting point not just for regulation of cognitive and metacognitive processes, but also to encompass processes of regulation of internal and external resources such as motivation, affect, and social support (Dresel et al., 2015; Grassinger et al., 2018; Tulis et al., 2016). In short, motivation regulation may be defined as activities that intend to initiate, maintain, or change the learner’s willingness to engage in a specific goal-oriented activity. This often occurs along with the necessity of emotion regulation, which describes processes of how individuals influence the experience and expression of their personal emotions (cf., Gross, 1998). In

their framework of learning from errors, Tulis et al. (2016) assume that the perception of an error while learning, based on primary appraisal, leads to emotions that range from a neutral value to a negative value such as surprise, frustration, or anger. Due to secondary appraisal, including aspects such as controllability and personal resources, these primary emotional reactions might attenuate, intensify, or change. Hence, to sustain the student's motivation and facilitate persistence in the learning task despite the negative emotions arising from learning with errors, it is essential to apply adaptive functional strategies for motivation and emotion regulation. Although it depends on the specific learning situation, functional emotion and motivation strategies include, for example, distraction from rather than rumination on negative emotions or a reframing of the cognitive evaluation of the event, such as seeing an error as an opportunity for learning. However, dysfunctional strategies may be characterized by elements such as ignorance of the error or maladaptive distractions such as drinking alcohol, which impair further actions of instrumental coping with the situation (Gross, 1998; Leroy et al., 2012; Reindl et al., 2020).

Besides the regulation of motivation and emotion, learners also have to deal with this situation on a behavioral level (Dresel et al., 2013; Grassinger et al., 2018). This aligns with the idea of differentiation between emotion-focused coping and problem-focused coping in the stress literature (Carver et al., 1989; Folkman & Lazarus, 1980): Whereas emotion-focused coping aims to deal with the emotional distress associated with a stressful situation, problem-focused coping aims to alter the person-environment relationship and to solve the source of the stressful situation itself. From the process perspective of self-regulated learning, the detection of an error while learning may take place in the self-reflection phase: Learners compare and reflect on realized learning outcomes in relation to intended and planned learning goals. Therefore, at first, adaptive strategies for handling errors during learning may aim to better understand the causes of the appearance of this particular learning outcome by, for example, exploring contextual and task conditions, questioning one's learning actions, or updating the situation model of the current learning problem (Dresel et al., 2013). Furthermore, learners may request support from peers or instructors, such as asking a fellow student to explain the task and to illustrate a potential solution (Santor et al., 2020). Second, learners might consider this updated knowledge while reinitiating a new learning cycle, for example, by adapting their cognitive learning strategies to this updated task model.

We assume that such strategic knowledge for handling errors during learning develops, among other things, through practical confrontation with errors during learn-

ing, through experimentation, or through training (Dimitrova et al., 2017; Frese & Keith, 2015). Furthermore, knowledge gained from confrontation in a specific error situation may generalize and support the handling of future errors. Since we are not aware of any studies that address the question of the dimensionality of this abstract knowledge base, our first assumption is that this expertise is unidimensional.

Aims of the Study

Prior research has provided strong support that making errors is a good starting point for learning and knowledge acquisition (Frese & Keith, 2015; Metcalfe, 2017). To date, empirical studies have focused on the assessment of general adaptive reactions and tendencies in dealing with errors while learning or on the assessment of error culture (e.g., Dresel et al., 2013; Santor et al., 2020; Spychiger et al., 2006). In these studies, the learners' knowledge regarding the strategic handling of specific error situations has not been considered. However, similar to studies on metacognitive strategy knowledge (e.g., Artelt & Schneider, 2015), it can be assumed that a corresponding declarative and conditional knowledge base is a prerequisite for an effective handling of errors in learning. In this study, we addressed the research question of analyzing the psychometric properties of a scenario-based test procedure that aims to assess error-related conditional strategy knowledge in higher education. This error-related conditional strategy knowledge encompasses knowledge about the functionality of cognitive-behavioral and emotion-motivational strategies that are used to meet the specific demands of learning situations characterized by the occurrence of errors. First, although the scenarios include different error-prone situations and the described learning errors differ in the severity of their consequences, we tested whether a measurement model that assumes a unidimensional structure of scores at the scenario level adequately represents the empirical data (Hypothesis 1). Unidimensionality was tested within the framework of confirmatory factor analysis and evaluated by common model fit criteria (Schermelleh-Engel et al., 2003). In addition to unidimensionality (model of τ -congeneric tests; Steyer & Eid, 2001), the constraint of equal factor loadings was explored (model of essentially τ -equivalent tests). Second, on the basis of these measurement models, we tested for acceptable construct reliability (McDonald's ω). General rules of thumb (e.g., Schermelleh-Engel & Gädde, 2020) were applied for evaluation. Third, the relationship with an external criterion was explored (nomological validity). On the one hand, students' general adaptive cognitive-behavioral and affec-

tive-motivational error handling tendencies while learning and students' strategy knowledge of learning with errors are two clearly discernible constructs; therefore, correlations should be below $r = .5$ (large effect; Cohen, 1992). On the other hand, comparable to the discussion of the relationship between self-reported strategy use and strategy knowledge (e.g., Artelt & Schneider, 2015; Karlen & Compagnoni, 2017; but see also discussion above), higher strategy knowledge of learning with errors should be associated with more adaptive error handling tendencies, which results in a positive correlation of the two constructs. Therefore, a minor to medium positive correlation might be expected (Hypothesis 2). However, due to a lack of prior studies in the field of learning with errors, this was only a weak expectation. Finally, we explored options for a shorter test version.

Method

Design and Sample Description

The study was designed as an online survey and addressed students enrolled in German universities and colleges. Data were collected between May and August 2023. Students were made aware of the study in person during courses (these students received a bar of chocolate together with printed study information on the package) and via mailing lists. No additional monetary incentives were provided. We counted 304 visits to the survey website. After providing informed consent, students were asked for sociodemographic data before responding to the error learning strategy test. Afterward, additional questions were asked about the error culture and study skills etc. The error learning strategy test was completed by 185 students. However, after the first inspection of the data, two participants were excluded due to fragmentary responses and test page skipping or having a uniform response pattern. Furthermore, three students had partial missing data on the error learning strategy test. We excluded these students from all analyses, as this study focused on test properties and specific analyses require complete data. Furthermore, due to the pairwise comparison scoring method that was applied, missing data on one item may affect several comparisons. Finally, after test scoring (see below), we identified two students with unusually low test scores (> 5 standard deviations below the mean). Further inspection revealed unusual short survey processing times for these two students, which led us to conclude that the test and questionnaire were probably not answered in due form. Both students were excluded. Therefore, a total of 178 students were con-

sidered for the analyses (74.7% female). The students were enrolled in two universities (79.2%), a university of education (6.7%), and a university of applied sciences (14.0%), all located in southern and central Germany (Bavaria, Baden-Württemberg, Hesse). The students were on average (median) 23 years old (P25–P75: 21–25 years) and in their fifth study semester (P25–P75: second–eighth semester). Concerning family education background, 47.9% of the students responding to this question were first-generation students. A majority of students reported studying either in the broad field of education and psychology (28.7%) or teacher training (41.6%). However, a substantial number of students from the fields of natural sciences (9.0%), technical sciences (4.5%), or computer sciences (5.1%) were also present.

Materials

Error Learning Strategy Test

Comparable to design procedures for tests of metacognitive strategy knowledge (e.g., Karlen, 2017; Lockl et al., 2016), we first constructed a pool of six scenarios that address different error learning situations. Thereby, the learning situations should cover a broad range of demands that students typically face in higher education settings (e.g., to give an incorrect description of a term due to insufficient literature work, to provide an incorrect answer in a group discussion). Furthermore, as the severity of the consequences of errors has been identified to be a crucial dimension affecting learning (Horvath et al., 2021; Keith et al., 2020), the constructed error learning situations encompass scenarios characterized by mild consequences, such as an error within a study group preparing for an exam or an error that remains unnoticed by others, to more severe error consequences, such as a failure on a written exam or a failure on a term paper. Subsequently, for every scenario, we collected a list of strategies in order to handle the error situation. Besides our own formulations, we asked students (between three and six students, depending on the scenario) through cognitive interviews for possible strategies. The strategies we collected varied in their functionality of dealing with the specific situational requirements. Furthermore, the strategies encompassed action-based and cognitive regulation strategies as well as emotion and motivation regulation strategies. Then, in the next step, we selected eight strategies per scenario according to the following principles: First, strategies should cover a broad spectrum of functionality. Second, within the spectrum of action-based and cognitive regulation strategies, strategies that aim for a better understanding of the error situation as well as strategies that address planning and future learning ac-

tions should be represented. Third, within each scenario, strategies that address the regulation of motivation and emotion should be considered. In sum, six scenarios and a total of 48 strategies were available in German (see Electronic Supplementary Material [ESM] 1; an English translation of the test material is provided in ESM 2; the English translation was not validated by this study). The quality of each strategy could be rated on a 5-point grade scale from 1 = *very good* to 5 = *poor*.

General Adaptive Error Handling Tendencies

In order to measure students' adaptive habits and tendencies of handling errors in learning situations, two scales were used. First, the subscale "learning orientation" from the student questionnaire on error culture in the classroom by Spychiger et al. (2006) was used. In this scale, students have to rate eight statements on attitudes and the cognitive-behavioral handling of errors while learning (e.g., "I reconsider wrong solutions in assignments several times" or "Mistakes in my studies help me to do better afterward"). As the original scale was constructed for use in schools, items were adapted for use in higher education settings. Agreement with these items was rated on a 5-point scale from 1 = *not true at all* to 5 = *accurate*. Internal consistency was $\alpha = .76$. Second, four statements formulated by the authors about students' adaptive *affect* and *motivational* regulation tendencies in dealing with errors were provided. The items follow the general suggestion of Dresel et al. (2013) to consider affective-motivational responses to errors during learning and were adapted to the university setting. Three items were formulated in a negative way ("If I make an error in a course, I try less hard afterward"; "If I make an error in my studies, my interest in the subject decreases"; "The prospect of making an error makes my motivation to study drop") and one item was formulated in a positive way ("Even if I make errors in my studies, I don't lose my joy in doing so afterwards"). A 5-point response scale was used and negative items were inverted. Internal consistency was $\alpha = .76$. For each scale, manifest scale scores were estimated by taking the arithmetic mean.

Expert Ratings

In order to get a benchmark for the quality of a specific strategy with regard to the demands illustrated in the corresponding scenario, we contacted 27 experts (university professors and research associates, currently and formerly active) in the field of learning with errors and self-

regulated learning. Overall, 20 experts provided a rating¹. Strategies in scenarios were rated on the same 5-point grade scale as described in the previous section. We had one missing rating on the first scenario, and thus contrary to the other scenarios, the expert rating benchmark of Scenario 1 relies on 19 expert ratings only. On the basis of the expert ratings, we derived a standard of evaluation for learners' responses using a pairwise comparison scoring method. Therefore, the quality of a strategy is regarded exclusively in comparison with the quality of other strategies in the same scenario, independent of its absolute quality rating. Strategy A is assumed superior to strategy B if 75% or more of the experts provided a better (= lower) grade to strategy A in comparison with strategy B (cf., Händel et al., 2013; Karlen, 2017; Lockl et al., 2016, for comparable benchmark criteria). This procedure enables the derivation of between 15 and 19 valid comparisons per scenario and a total number of 105 valid comparisons. Strategies that were rated by experts as being of equal quality were not considered. Finally, these pairwise comparisons were used for a scoring of students' responses. Students received one point if their rating of a pair of strategies was in line with the experts' standard of comparison and zero points if their rating deviated.

Procedure

First, for each scenario and based on the scoring method described earlier, we estimated the absolute number of correctly rated comparisons of strategies for all students. We tested the unidimensionality of the ratings of the six scenarios using confirmatory factor analysis in Mplus (Muthén & Muthén, 1998–2017). Two models were tested: First, a model with a single latent variable and without the restriction of equal factor loadings was estimated (model of τ -congeneric tests; Steyer & Eid, 2001). If this assumption fits the empirical data, unidimensionality of the different test scenarios may be accepted. Second, a model with the restriction of equal factor loadings (model of essentially τ -equivalent tests; Steyer & Eid, 2001) was estimated. If this model can be accepted, units of differences between scenarios may relate equally to the latent variable. Furthermore, construct reliability (McDonald's ω) was estimated in the different measurement models. Next, a correlation of the error strategy knowledge test with students' general adaptive cognitive-behavioral and affective-motivational error handling tendency while learning was estimated. Both variables were estimated as latent variables. Students' general adaptive

¹ On an individual strategy level, interrater agreement (Krippendorff's α , ordinal scale) across all strategies and raters was .74.

tendency in dealing with errors was measured by two manifest scale scores, learning orientation as well as adaptive affect and motivation regulation tendency, and assuming equal factor loadings. Finally, psychometric properties of a shortened test version were explored. In the latent variable models, standard errors were corrected for non-normality using the MLR estimator and missing data were treated via FIML (see <https://doi.org/10.23668/psycharchives.15240> for Mplus-output files).

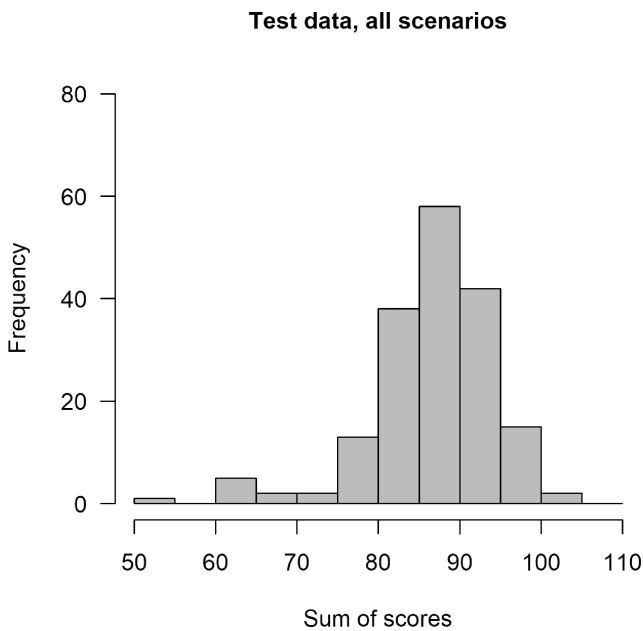


Figure 1. Distribution of test scores. $n = 178$ participants.

Results

Descriptive Statistics

Across all six error scenarios, students' ratings were compared with the 105 valid pair comparisons from the expert ratings. Students received one point if their rating of a pair of strategies was in line with the experts' standard of comparison. On average, students received a test score of $M = 87.30$ ($SD = 7.63$; min. = 54; max. = 101; skewness = -1.42 ; kurtosis = 3.39) points (see Figure 1). With the exception of the comparison of Strategy 2 against Strategy 6 in the first scenario, all items showed positive item-total correlations. Item difficulty varied between .22 and .99, with an average of .83 (see ESM 3). Separate findings on psychometric properties in each scenario as well as correlations between the six error scenarios are shown in Table 1. It is evident that descriptive data for Scenario 1 show the lowest standard deviation of test scores. Furthermore, results indicate positive correlations of scores between scenarios with two exceptions: Test scores in Scenario 1 did not show positive significant correlations with test scores in Scenarios 2, 4, and 6, as did test scores between Scenario 2 and Scenario 6.

Regarding students' adaptive error handling tendencies, we observed a mean score of $M = 3.55$ ($SD = 0.56$; min. = 2.25; max. = 4.75; skewness = -0.14 ; kurtosis = -0.48) for the cognitive-behavioral subscale and a mean score of $M = 3.58$ ($SD = 0.75$; min. = 1.75; max. = 5.00; skewness = -0.08 ; kurtosis = -0.55) for the affect and motivation regulation subscale.

Table 1. Descriptive statistics and correlations on scenario level

		α	Descriptives		Correlations				
			Max.	M (SD)	2 r	3 r	4 r	5 r	6 r
1	Scenario 1	.30	15	12.87 (1.30)	.13	.15*	.13	.25**	.05
2	Scenario 2	.63	19	14.49 (2.47)		.26**	.29**	.18*	.14
3	Scenario 3	.79	15	13.10 (2.45)			.25**	.30**	.17*
4	Scenario 4	.57	18	14.17 (2.33)				.23**	.21**
5	Scenario 5	.61	19	16.38 (2.12)					.22**
6	Scenario 6	.66	19	16.29 (2.34)					

Note. α = Cronbach's α (estimates might be biased due to local stochastic dependencies; see ESM 4); Max. = maximum test score/number of comparisons. * $p \leq .05$; ** $p \leq .01$

Table 2. Factor loadings and residual variances of the estimated measurement models

	All scenarios		Scenarios 2–6	
	τ -congeneric Est. (Std.)	e. τ -equivalent Est. (Std.)	τ -congeneric Est. (Std.)	e. τ -equivalent Est. (Std.)
Factor loadings				
Scenario 1	0.41 (.32)	0.83 (.59)		
Scenario 2	1.11 (.45)	0.83 (.34)	1.14 (.46)	1.11 (.45)
Scenario 3	1.30 (.53)	0.83 (.35)	1.31 (.54)	1.11 (.46)
Scenario 4	1.17 (.50)	0.83 (.36)	1.21 (.52)	1.11 (.48)
Scenario 5	1.11 (.53)	0.83 (.40)	1.03 (.49)	1.11 (.52)
Scenario 6	0.82 (.35)	0.83 (.35)	0.86 (.37)	1.11 (.46)
Residual variances				
Scenario 1	1.51 (.90)	1.29 (.65)		
Scenario 2	4.81 (.80)	5.18 (.88)	4.75 (.79)	4.82 (.80)
Scenario 3	4.28 (.72)	4.96 (.88)	4.26 (.71)	4.54 (.79)
Scenario 4	4.04 (.75)	4.50 (.87)	3.96 (.73)	4.08 (.77)
Scenario 5	3.24 (.72)	3.50 (.84)	3.42 (.76)	3.32 (.73)
Scenario 6	4.77 (.88)	4.84 (.88)	4.71 (.87)	4.48 (.79)
Construct reliability				
ω	.609	.503	.594	.590

Note. τ -congeneric = model of τ -congeneric tests; e. τ -equivalent = model of essentially τ -equivalent tests; Est. = estimated value unstandardized solution; Std. = estimated value standardized solution.

Confirmatory Factor Analyses

In order to test for unidimensionality of the error learning strategy test, a τ -congeneric measurement model with one latent variable was assumed. Fit indices showed signs of a satisfactory fit of the empirical variance-covariance structure with the theoretically implied variance-covariance structure ($\chi^2 = 5.66$, $df = 9$, $p = .774$; root mean square error of approximation [RMSEA] = .00; comparative fit index [CFI] = 1.00; Tucker-Lewis index [TLI] = 1.00). When taking a closer look at the estimated factor loadings, the lowest estimated factor loading was found for Scenario 1 and differed from loadings of the other scenarios (see Table 2). A restriction of equal factor loadings resulted in a significant deterioration of the model fit (essentially τ -equivalent measurement model: $\chi^2 = 21.06$, $df = 14$, $p = .100$; RMSEA = .05; CFI = .88; TLI = .87; Satorra-Bentler scaled χ^2 difference test: $\chi^2 = 13.44$, $df = 5$, $p = .020$). Therefore, we explored a second latent variable model without taking Scenario 1 into account. Again, fit indices of the τ -congeneric measurement model showed signs of excellent fit ($\chi^2 = 2.40$, $df = 5$, $p = .791$; RMSEA = .00; CFI = 1.00; TLI = 1.00) as did the essentially τ -equivalent measurement model with equal factor loadings ($\chi^2 = 3.73$, $df = 9$, $p = .928$; RMSEA = .00; CFI = 1.00;

TLI = 1.00; Satorra-Bentler scaled χ^2 difference test: $\chi^2 = 1.39$, $df = 4$, $p = .846$). In this last estimated model, standardized factor loadings varied between $\lambda = 0.45$ and 0.52. Estimates of construct reliability in the different measurement models varied between $\omega = .503$ and .609.²

Correlations With Students' General Adaptive Error Handling Tendencies

On the manifest level, estimated correlations between the error learning strategy test (all scenarios) and the cognitive-behavioral error learning tendency subscale of learning orientation were $r = .12$ ($p = .102$; $n = 174$), and $r = .09$ ($p = .223$) for the error handling tendency subscale of affect and motivation regulation. For the reduced error learning strategy test (exclusively encompassing Scenarios 2–6), correlations were $r = .15$ ($p = .054$; learning orientation) and $r = .09$ ($p = .257$; affect and motivation regulation), respectively. In order to estimate correlations while controlling for measurement error, a latent variable was assumed for the error learning strategy test as well as for the two general adaptive error handling tendency scales. Model fit was satisfactory (error learning test with all six scenarios and applying a τ -congeneric measure-

² Concerning estimates of internal consistency (Cronbach's α), please refer to the Electronic Supplementary Material (ESM 4).

ment model: $\chi^2 = 19.33$, $df = 20$, $p = .501$; RMSEA = .00; CFI = 1.00; TLI = 1.00; error learning test based on Scenarios 2–6 and applying an essentially τ -equivalent measurement model: $\chi^2 = 10.60$, $df = 18$, $p = .911$; RMSEA = .00; CFI = 1.00; TLI = 1.00). The results indicate a correlation of the latent error learning strategy test based on all scenarios with the latent adaptive error handling tendency variable of $r = .22$, which did not reach statistical significance ($p = .079$). For the reduced test version, exclusively encompassing the error learning Scenarios 2–6, the correlation between the latent error learning strategy test and the latent adaptive error handling tendency variable increased to $r = .24$ ($p = .048$; see Figure 2).

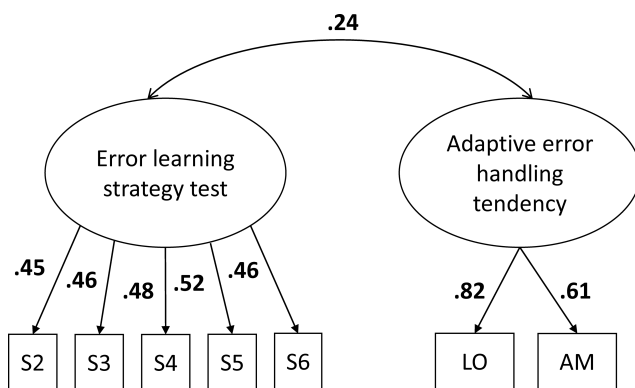


Figure 2. Latent variable model (standardized solution). The error learning strategy test was measured by five scenarios (Scenarios 2–6) and under the assumption of essentially τ -equivalence. LO = learning orientation; AM = affect and motivation regulation. Bold numbers indicate statistical significance ($p \leq .05$).

Short Test Version

After disregarding Scenario 1 and still applying an essentially τ -equivalent measurement model, the lowest loss in construct reliability was observed when further disregarding Scenario 6. Fit statistics of the error learning test based on Scenarios 2–5 were: $\chi^2 = 2.85$, $df = 5$, $p = .723$; RMSEA = .00; CFI = 1.00; TLI = 1.00. Construct reliability of the four scenarios was $\omega = .567$. The correlation between the latent error learning strategy test (Scenarios 2–5) and the latent adaptive error handling tendency variable was $r = .28$ ($p = .018$).

Discussion

When students face errors while being engaged in learning, different options and strategies of handling the situation may be present. This encompasses aspects of cog-

nitive-behavioral as well as affective-motivational regulation (Dresel et al., 2013; Tulis et al., 2016). Furthermore, strategies differ with regard to the adaptivity of handling the error learning situation (Dresel et al., 2013; Santor et al., 2020). We assume that besides general usage tendencies, conditional knowledge of the adaptivity of a specific cognitive-behavioral and affective-motivational regulation strategy when faced with a specific error situation while learning is crucial for an effective handling of this learning situation. Therefore, we developed a scenario-based test that aims to make such individual knowledge differences visible. From a total of 105 available pair comparisons, the students correctly reproduced about 83% of the pair comparisons on average. Therefore, the test tends to be relatively easy for students in higher education settings. This is not surprising, as former studies on metacognitive strategy knowledge in higher education faced comparable challenges (e.g., Händel et al., 2012). Students attending universities tend to be experienced and efficient learners. Nonetheless, our test shows substantial variability in test scores and the arithmetic mean was more than two standard deviations below the theoretical test maximum.

Confirmatory factor analyses with sum scores in scenarios as manifest indicators showed good signs of unidimensionality, which clearly support our expectations. However, the factor loading of the first scenario differed from the factor loadings of the other five scenarios. Therefore, a measurement model that assumes comparable units of difference, i.e., equal factor loadings (model of essentially τ -equivalent tests; Steyer & Eid, 2001), has been shown to be appropriate if the first scenario is skipped. Nonetheless, besides the finding of unidimensionality of the test, substantial unexplained variance on the individual scenario indicators was found. This is evident in the low standardized factor loadings, only reaching values between .4 and .5, and it is further reflected in the construct reliability measures, which only reached values of about .6. Therefore, reliability fell short of expectations (Schermelleh-Engel & Gädde, 2020). Besides the possibility of a relatively homogeneous group of study participants, which reduces true score variance that can be explained by the test, another possible explanation might be the aforementioned low difficulty of the test scenarios, which reduces the variability and reliability of the test scores. However, non-satisfactory reliabilities are not uncommon for this type of test and in this target population (e.g., Händel et al., 2012). Furthermore, we should mention that individual learning goals are not addressed in the scenarios, which is another source of error variance due to interpretation differences by test takers. Consequently, instead of using manifest scale scores, the test might be used well in a latent variable

modeling approach, for example, in structural equation modeling, which tends to separate true score variance from error variance. On the basis of the results discussed here, we would recommend that future test users use the test version with five scenarios (Scenarios 2–6) in their studies. If the time available for testing is very limited, a reduction of the test to four scenarios (Scenarios 2–5) would also be conceivable, although this reduction proposal should be replicated in a new sample. The test should not be used as a diagnostic tool for individual persons. With regard to the dimensional structure of the test, it should be noted that the assumption of unidimensionality lacks robust theoretical support. Theoretical models of learning with errors in higher education from a strategic perspective are still fragmentary and need to be developed, for example, with regard to the components or causes of individual development and interindividual differences. Furthermore, due to the procedure applied of evaluating the functionality of eight different strategies in one scenario, multidimensionality in the functionality of strategies in individual scenarios may remain undetected (see Waldeyer et al., 2019, for a different approach).

Finally, correlations with students' general adaptive error handling tendency in learning situations were explored. Findings of structural equation models, which treat knowledge of error learning strategies as well as adaptive regulation tendencies of dealing with errors as latent variables, show low to medium-high correlations of .22–.28 between the two constructs. The correlation between the latent variables reached significance when the first error learning scenario was not taken into account. When considering error learning Scenario 1, the correlation of the latent variables was not significant. In line with our expectations, these analyses show that the aspect of conditional knowledge of error learning strategies needs to be differentiated from the aspect of general, situation-unspecific, adaptive tendency in dealing with errors while learning, encompassing affect and motivation as well as cognitive-behavioral components. However, higher knowledge tends to go hand in hand with more adaptive regulation tendencies of dealing with errors, at least on a low to medium-high effect size level. In addition, we might assume that good knowledge of strategies for handling errors while learning impacts the effectivity of using specific cognitive and behavioral actions as well as affect-motivation regulation strategies when confronted with errors (see Artelt & Schneider, 2015, for such findings regarding reading comprehension). This hypothesis should be tested in future research.

Limitations and Outlook

According to our empirical findings, the test was relatively easy for university students. Therefore, developing further strategy item comparisons that reach sufficient agreement by experts while being of great difficulty for university students seems desirable. A greater difficulty of strategy comparisons might further contribute to improving the reliability of the test, which is a second desideratum. In addition, the test may be expanded to include more typical scenarios from disciplines outside of humanities and social sciences, such as engineering problems, and the errors presented in the scenarios may vary more based on cognitive demands. Thirdly, we should keep in mind that we are only dealing with hypothetical situations that require cognitive and affective perspective-taking skills. This could have a particular impact on affect regulation measurement. Furthermore, the severity of consequences is not an objective fact and its interpretation may vary among students. Moreover, all the test material was presented in German, and the use of the test material in other languages, such as English, requires separate validation. Finally, studies that address further signs of validity of our measures are desirable. For example, longitudinal studies might address the development of strategic knowledge in learning with errors in relation to specific learning incidents. This might also encompass students from further disciplines. Besides analyzing variables that affect the development of strategic knowledge of learning with errors, future research might also address the interesting topic of how such knowledge contributes to actual learning behavior, for example, within an experimental observation study. This may include facets such as self-efficacy beliefs and seeking out learning situations that are associated with specific academic risks.

Electronic Supplementary Material

The electronic supplementary material is available with the online version of the article at <https://doi.org/10.1026/0012-1924/a000341>

- ESM 1.** Error learning strategy test [in German language].
- ESM 2.** Error learning strategy test [English translation, non-validated].
- ESM 3.** Evaluation key and descriptive item parameters.
- ESM 4.** Local stochastic independence and internal consistency.

References

- Artelt, C., Naumann, J., & Schneider, W. (2010). Lesemotivation und Lernstrategien. In E. Klieme, C. Artelt, J. Hartig, N. Jude, O. Köller, M. Prenzel, W. Schneider, & P. Stanat (Eds.), *PISA 2009. Bilanz nach einem Jahrzehnt* [PISA 2009: Taking stock after a decade] (pp. 73–112). Waxmann.
- Artelt, C., & Neuenhaus, N. (2010). Metakognition und Leistung. In W. Bos, E. Klieme, & O. Köller (Eds.), *Schulische Lerngelegenheiten und Kompetenzentwicklung: Festschrift für Jürgen Baumert* [School learning opportunities and competence development: Festschrift for Jürgen Baumert] (pp. 127–146). Waxmann.
- Artelt, C., & Schneider, W. (2015). Cross-country generalizability of the role of metacognitive knowledge in students' strategy use and reading competence. *Teachers College Record*, *117*, 1–32. <https://doi.org/10.1177/016146811511700109>
- Boekaerts, M. (1999). Self-regulated learning: Where we are today. *International Journal of Educational Research*, *31*, 445–457. [https://doi.org/10.1016/S0883-0355\(99\)00014-2](https://doi.org/10.1016/S0883-0355(99)00014-2)
- Carver, C. S., Scheier, M. F., & Weintraub, J. K. (1989). Assessing coping strategies: A theoretically based approach. *Journal of Personality and Social Psychology*, *56*, 267–283.
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*, 155–159.
- Dimitrova, N. G., van Hooft, E. A. J., van Dyck, C., & Groenewegen, P. (2017). Behind the wheel: What drives the effects of error handling? *The Journal of Social Psychology*, *157*, 658–672. <https://doi.org/10.1080/00224545.2016.1270891>
- Dresel, M., Schmitz, B., Schober, B., Spiel, C., Ziegler, A., Engelschalk, T., et al. (2015). Competencies for successful self-regulated learning in higher education: Structural model and indications drawn from expert interviews. *Studies in Higher Education*, *40*, 454–470. <https://doi.org/10.1080/03075079.2015.1004236>
- Dresel, M., Schober, B., Ziegler, A., Grassinger, R., & Steuer, G. (2013). Affektiv-motivationale adaptive und handlungsadaptive Reaktionen auf Fehler im Lernprozess [Affective-motivationale adaptive and action adaptive reactions on errors in learning processes]. *Zeitschrift für Pädagogische Psychologie*, *27*, 255–271. <https://doi.org/10.1024/1010-0652/a000111>
- Folkman, S., & Lazarus, R. S. (1980). An analysis of coping in a middle-aged community sample. *Journal of Health and Social Behavior*, *21*, 219–239. <https://doi.org/10.2307/2136617>
- Frese, M., & Keith, N. (2015). Action errors, error management, and learning in organizations. *Annual Review of Psychology*, *66*, 661–687. <https://doi.org/10.1146/annurev-psych-010814-015205>
- Grassinger, R., Scheunpflug, A., Zeinz, H., & Dresel, M. (2018). Smart is who makes lots of errors? The relevance of adaptive reactions to errors and a positive error climate for academic achievement. *High Ability Studies*, *29*, 37–49. <https://doi.org/10.1080/13598139.2018.1459294>
- Gross, J. J. (1998). The emerging field of emotion regulation: An integrative review. *Review of General Psychology*, *2*, 271–299. <https://doi.org/10.1037/1089-2680.2.3.271>
- Händel, M., Artelt, C., & Weinert, S. (2013). Assessing metacognitive knowledge: Development and evaluation of a test instrument. *Journal for Educational Research Online*, *5*, 162–188. <https://doi.org/10.25656/01:8429>
- Händel, M., Tupac-Yupanqui, A., & Lockl, K. (2012). *Metakognitives Wissen und der Einsatz von Lernstrategien bei Studierenden* (NEPS Working Paper No. 20) [Metacognitive knowledge and the use of learning strategies by students]. Otto-Friedrich-Universität, Nationales Bildungspanel.
- Horvath, D., Klamar, A., Keith, N., & Frese, M. (2021). Are all errors created equal? Testing the effect of error characteristics on learning from errors in three countries. *European Journal of Work and Organizational Psychology*, *30*, 110–124. <https://doi.org/10.1080/1359432X.2020.1839420>
- Hübner, V., & Pfost, M. (2023). Operationalization of academic risk taking in university students. *Journal for Educational Research Online*, *15*, 74–94. <https://doi.org/10.31244/jero.2023.01.04>
- Ivancic, K., & Hesketh, B. (2000). Learning from errors in a driving simulation: Effects on driving skill and self-confidence. *Ergonomics*, *43*, 1966–1984. <https://doi.org/10.1080/00140130050201427>
- Karlen, Y. (2017). The development of a new instrument to assess metacognitive strategy knowledge about academic writing and its relation to self-regulated writing and writing performance. *Journal of Writing Research*, *9*, 61–86. <https://doi.org/10.17239/jowr-2017.09.01.03>
- Karlen, Y., & Compagnoni, M. (2017). Implicit theory of writing ability: Relationship to metacognitive strategy knowledge and strategy use in academic writing. *Psychology Learning & Teaching*, *16*, 47–63. <https://doi.org/10.1177/1475725716682887>
- Keith, N., Horvath, D., & Klamar, A. (2020). The more severe the merrier: Severity of error consequences stimulates learning from error. *Journal of Occupational and Organizational Psychology*, *93*, 712–737. <https://doi.org/10.1111/joop.12312>
- Leroy, V., Grégoire, J., Magen, E., Gross, J. J., & Mikolajczak, M. (2012). Resisting the sirens of temptation while studying: Using reappraisal to increase focus, enthusiasm, and performance. *Learning and Individual Differences*, *22*, 263–268. <https://doi.org/10.1016/j.lindif.2011.10.003>
- Lipnevich, A. A., Berg, D. A., & Smith, J. K. (2016). Toward a model of student response to feedback. In G. T. L. Brown & L. T. Harris (Eds.), *Handbook of human and social conditions in assessment* (pp. 169–185). Routledge.
- Lockl, K., Händel, M., Haberkorn, K., & Weinert, S. (2016). Metacognitive knowledge in young children: Development of a new test procedure for first graders. In H.-P. Blossfeld, J. von Maurice, M. Bayer, & J. Skopek (Eds.), *Methodological issues of longitudinal surveys: The example of the National Educational Panel Study* (pp. 465–484). Springer. https://doi.org/10.1007/978-3-658-11994-2_26
- Metcalfe, J. (2017). Learning from errors. *Annual Review of Psychology*, *68*, 465–489. <https://doi.org/10.1146/annurev-psych-010416-044022>
- Muthén, L. K., & Muthén, B. O. (1998–2017). *Mplus user's guide*. Eighth edition. Muthén & Muthén.
- Paris, S. G., Lipson, M. Y., & Wixson, K. K. (1983). Becoming a strategic reader. *Contemporary Educational Psychology*, *8*, 293–316. [https://doi.org/10.1016/0361-476X\(83\)90018-8](https://doi.org/10.1016/0361-476X(83)90018-8)
- 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*, 801–813. <https://doi.org/10.1177/0013164493053003024>
- Reindl, M., Tulis, M., & Dresel, M. (2020). Profiles of emotional and motivational self-regulation following errors: Associations with learning. *Learning and Individual Differences*, *77*, 101806. <https://doi.org/10.1016/j.lindif.2019.101806>
- Richardson, M., Abraham, C., & Bond, R. (2012). Psychological correlates of university students' academic performance: A systematic review and meta-analysis. *Psychological Bulletin*, *138*, 353–387. <https://doi.org/10.1037/a0026838>
- Santor, D. A., Colvin, E., & Sinclair, A. (2020). Coping with failure. *Educational Review*, *72*, 635–649. <https://doi.org/10.1080/00131911.2018.1524855>

- Schermelleh-Engel, K., & Gåde, J. C. (2020). Modellbasierte Methoden der Reliabilitätsschätzung. In H. Moosbrugger & A. Kellava (Eds.), *Testtheorie und Fragebogenkonstruktion* [Test theory and questionnaire construction] (pp. 335–368). Springer. https://doi.org/10.1007/978-3-662-61532-4_15
- Schermelleh-Engel, K., Moosbrugger, H., & Müller, H. (2003). Evaluating the fit of Structural Equation Models: Tests of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research Online*, 8, 23–74.
- Seeger, J., Lenhard, W., & Wisniewski, K. (2021). Metakognitives Strategiewissen in sprachbezogenen Situationen [Metacognitive strategy knowledge in language-related situations]. *Diagnostica*, 67, 189–199. <https://doi.org/10.1026/0012-1924/a000275>
- Spörer, N., & Brunstein, J. C. (2006). Erfassung selbstregulierten Lernens mit Selbstberichtsverfahren [Recording self-regulated learning with self-reporting methods]. *Zeitschrift für Pädagogische Psychologie*, 20, 147–160. <https://doi.org/10.1024/1010-0652.20.3.147>
- Spychiger, M., Kuster, R., & Oser, F. (2006). Dimensionen von Fehlerkultur in der Schule und deren Messung. Der Schülerfragebogen zur Fehlerkultur im Unterricht für Mittel- und Oberstufe [Dimensions of error culture at school and their measurement. The student questionnaire on error culture in the classroom for middle and high school]. *Schweizerische Zeitschrift für Bildungswissenschaften*, 28, 87–110.
- Steuer, G., & Dresel, M. (2015). A constructive error climate as an element of effective learning environments. *Psychological Test and Assessment Modeling*, 57, 262–275.
- Steyer, R., & Eid, M. (2001). *Messen und Testen* [Measuring and testing]. Springer.
- Tulis, M., Steuer, G., & Dresel, M. (2016). Learning from errors: A model of individual processes. *Frontline Learning Research*, 4, 12–26. <https://doi.org/10.14786/flr.v4i2.168>
- Waldeyer, J., Fleischer, J., Wirth, J., & Leutner, D. (2019). Entwicklung und erste Validierung eines Situational-Judgement-Instruments zur Erfassung von Kompetenzen im Bereich des Ressourcenmanagements (ReMI) [Development and initial validation of a situational judgment instrument to assess resource management skills (ReMI)]. *Diagnostica*, 65, 108–118. <https://doi.org/10.1026/0012-1924/a000217>
- Wirth, J., & Leutner, D. (2008). Self-regulated learning as a competence. *Zeitschrift für Psychologie / Journal of Psychology*, 216, 102–110. <https://doi.org/10.1027/0044-3409.216.2.102>
- Zhang, Q., & Fiorella, L. (2023). An integrated model of learning from errors. *Educational Psychologist*, 58, 18–34. <https://doi.org/10.1080/00461520.2022.2149525>
- Zimmerman, B. J., & Moylan, A. R. (2009). Self-regulation. Where metacognition and motivation intersect. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Handbook of metacognition in education* (pp. 299–315). Routledge.

Published online December 16, 2024

Acknowledgments

We would like to thank Friederike Karg (Otto-Friedrich-Universität Bamberg) for her support in conducting the study. We would also like to thank all the experts for their ratings and the students for their participation in the study.

Funding

Open access publication enabled by the University of Bamberg.

ORCID

Maximilian Pfof

 <https://orcid.org/0000-0002-7066-0456>

Vanessa Hübner

 <https://orcid.org/0000-0003-3568-7618>

Prof. Dr. Maximilian Pfof

Department of Educational Research
University of Bamberg
96045 Bamberg
Germany
maximilian.pfof@uni-bamberg.de