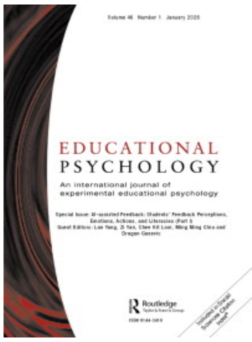


Routledge
Taylor & Francis Group

Educational Psychology

An International Journal of Experimental Educational Psychology

ISSN: 0144-3410 (Print) 1469-5820 (Online) Journal homepage: www.tandfonline.com/journals/cedp20

Predicting academic performance by self-assessed and objectively measured understanding and self-efficacy

Ursina E. Raemy, Natalie Borter, Mathias Mejeh & Stefan J. Troche

To cite this article: Ursina E. Raemy, Natalie Borter, Mathias Mejeh & Stefan J. Troche (13 Feb 2026): Predicting academic performance by self-assessed and objectively measured understanding and self-efficacy, Educational Psychology, DOI: 10.1080/01443410.2026.2625279

To link to this article: https://doi.org/10.1080/01443410.2026.2625279

EDUCATIONAL PSYCHOLOGY
https://doi.org/10.1080/01443410.2026.2625279

Predicting academic performance by self-assessed and objectively measured understanding and self-efficacy

Ursina E. Raemy[a], Natalie Borter[a], Mathias Mejeh[b] and Stefan J. Troche[a]

[a]Institute of Psychology, University of Bern, Bern, Switzerland; [b]Zurich University of Teacher Education, Zurich, Switzerland

ABSTRACT

An adequate assessment of what students have understood during the term is an essential precondition for self-regulated learning (SRL) as it allows for adapting learning strategies to enhance academic performance. This study compared self-assessed and objectively measured understanding of the learning material as predictors of academic performance and their interplay with general and academic self-efficacy. In a bachelor's Psychological Assessment course, 271 psychology students regularly completed online formative assessments (OFA) throughout the term. Before each OFA, they self-assessed how well they had understood the content of the previous lecture. Performance in the OFA was used as an objective measure of understanding. Furthermore, self-efficacy was assessed at the beginning of the term. Both self-assessed and objectively measured understanding were significant predictors of the grade in the final exam as a measure of academic performance. When considered concurrently in a latent regression analysis, only objectively measured understanding predicted academic performance. Academic self-efficacy predicted academic performance above both measures of understanding, maybe because it covers how students deal with the specific demands of the exam. According to the present results, students engaging in SRL should rely on objective indicators of their understanding instead of self-assessments to select appropriate learning strategies.

ARTICLE HISTORY
Received 17 April 2025
Accepted 27 January 2026

KEYWORDS
Self-assessed understanding; objectively measured understanding; self-efficacy; online formative assessments; academic performance

Introduction

In today's rapidly evolving and complex world, tertiary education plays a central role in fostering personal development and preparing individuals for societal and professional demands (Brewer & McEwan, 2010; OECD, 2024; Vierzigmann & Lehmann, 2022). Understanding the diverse factors shaping academic performance is essential for improving educational outcomes and empowering students to reach their full potential and contribute meaningfully to society. Unlike primary and secondary education,

CONTACT Stefan J. Troche stefan.troche@unibe.ch Institute of Psychology, University of Bern, Bern, Switzerland
Supplemental data for this article can be accessed online at https://doi.org/10.1080/01443410.2026.2625279.

which typically provide more structured guidance and support, tertiary education places greater demands on students' ability to self-regulate their own learning (Anthonysamy et al., 2020; Broadbent & Fuller-Tyszkiewicz, 2018; Vosniadou, 2020). These demands are reinforced by digital transformation, leading to new learning formats, such as blended or fully online courses (Cerezo et al., 2020). Consequently, self-regulated learning (SRL) has emerged as a key determinant of academic performance (Broadbent & Poon, 2015; Cheng et al., 2025; Ergen & Kanadli, 2017; Zhao et al., 2025).

Over the past decades, various models have been developed to describe SRL (e.g. Efklides, 2011; Pintrich, 2000; Winne & Hadwin, 1998; Zimmerman, 2000). Recently, Panadero (2017) attempted to find their common ground by identifying the most crucial characteristics of these models. One of them is the cyclical nature of SRL, typically comprising three phases: 1) a preparatory phase, involving task analysis, goal setting, and planning, 2) a performance phase, centred on executing the task while monitoring and controlling the process, and 3) an appraisal phase, focused on reflection as well as regulation and adaptation for future performance. These stages highlight the core of SRL with students taking ownership of their learning (Zimmerman & Schunk, 2011). A further common feature of SRL models in Panadero's overview is the incorporation of emotional (e.g. emotion regulation), motivational (e.g. self-motivation beliefs, goal setting), and (meta)cognitive aspects of learning, with models differing in the emphasis they place on these aspects. Metacognition plays a crucial role in SRL as it enables learners to monitor and regulate their learning (Panadero, 2017; Winne & Perry, 2000). Metacognitive regulation, often described as 'thinking about thinking,' can be further categorised into metacognitive knowledge and regulation (Dunlosky & Metcalfe, 2008). Metacognitive knowledge involves understanding oneself as a learner, recognising performance factors, and knowing appropriate strategies. Metacognitive regulation refers to applying these strategies, such as planning, monitoring comprehension and performance, and evaluating their effectiveness (Dunlosky & Metcalfe, 2008). A key element of this regulatory function is the ability to assess the own level of understanding accurately (Lai, 2011; Schwarz, 2015).

The level of understanding of the learning material has often been measured by asking students to indicate how well they have understood the learning material (Brown et al., 2015; Falchikov & Boud, 1989; Li & Zhang, 2021; Yan, 2020; Yan et al., 2023). However, such self-assessments are prone to biases and, therefore, can be assumed to provide only a rough estimate of the actual level of understanding (Kruger & Dunning, 1999; León et al., 2023; Neubauer & Hofer, 2021; Rammstedt & Rammsayer, 2002; Zell & Krizan, 2014). Technological advancements, such as online versions of formative assessments (OFAs), have enabled new ways of capturing students' level of understanding more directly (Ifenthaler & Drachsler, 2018; Kulasegaram & Rangachari, 2018; Robinson & Song, 2019; Tomasik et al., 2018). Although there is no uniform definition of formative assessments, a common element of most existing definitions is the collection of information on students' learning progress, e.g. by means of the number and kind of correctly/incorrectly solved tasks (Bennett, 2011; Black & Wiliam, 1998; Gikandi et al., 2011; McManus, 2008). Although formative assessments, in the sense of traditional homework, have long been part of tertiary education, they have the disadvantage that feedback on individual results can often only be given in a non-specific manner, especially in large student groups (e.g. Gibbs & Simpson, 2005). Their online versions, on the contrary, can be easily designed to offer automatic but still individualised and immediate feedback on the accuracy of each individual task and/or a summary of the number of tasks solved correctly (Gikandi et al., 2011). Thus, OFAs can provide students with objective information about their actual level of understanding, which, in turn, can be used to adapt the current learning strategy (Clark, 2012; Irons & Elkington, 2022).

Both self-assessed and objectively measured understanding have been shown to have positive associations with academic performance (Higgins et al., 2023; Pinquart & Ebeling, 2020; Richardson et al., 2012; Yan, 2020). However, direct comparisons of their relative predictive power remain scarce. Therefore, the major aim of the present study was to examine the interplay between the self-assessed and the objectively measured level of understanding in their prediction of academic performance.

Previous research indicated that self-assessments of the current level of understanding reflected motivational beliefs, confidence, and, especially, self-efficacy rather than the actual level of understanding (Andrade, 2019; Panadero et al., 2017; Pinquart & Ebeling, 2020; Yan & Brown, 2017; Zell & Krizan, 2014). From this perspective, the two measures of the level of understanding (self-assessed vs. objectively measured) may represent different constructs when self-assessed understanding is a measure of self-efficacy rather than the actual level of understanding. Self-efficacy is defined as the belief in one's capability to succeed in specific tasks or learning situations (Bandura, 1997). Self-efficacy can be considered at different levels of specificity, ranging from a general disposition to more domain-specific forms, and it has often been recommended to use context-specific measures of self-efficacy to predict outcomes in a given context (Bandura, 2012; Gallagher, 2012; Honicke & Broadbent, 2016; Schunk & DiBenedetto, 2016). Crucially, self-efficacy is based on students' perceived, rather than actual abilities (Waddington, 2023). In SRL models, self-efficacy is often embedded as a motivational driver that influences the use of self-regulatory strategies and significantly shapes how students approach learning (Panadero et al., 2017; Schunk, 1990; Schunk & DiBenedetto, 2016). Therefore, it is not surprising that self-efficacy has been shown to predict academic performance (Honicke & Broadbent, 2016). The relation between self-efficacy and learning, thereby, is considered to be reciprocal, with high self-efficacy shaping the approach to learning and thereby enhancing performance, which in turn leads to increasing self-efficacy (Talsma et al., 2018; Williams & Williams, 2010)

Nevertheless, the relationship between self-efficacy and the level of understanding may be artificially enhanced by response biases when both are measured by means of self-assessments (Paulhus & Vazire, 2007). For example, optimistic and hopeful students might report higher self-efficacy (Feldman & Kubota, 2015; Phan, 2016), but might also overestimate their level of understanding. Such an artificially enhanced influence of self-efficacy on the level of understanding should become obvious in the self-assessment but not in the objective measure of the level of understanding.

Therefore, in the present study, we assessed students' general and academic self-efficacy in addition to their self-assessed and objectively measured level of understanding as predictors of academic performance. Specifically, we investigated whether self-assessed and objectively measured understanding explain the same portion of variance in academic performance and, if not, whether a potential relationship between self-assessed understanding and academic performance can be better understood in terms of self-efficacy. Based on this design, the study addressed the following four research questions:

1. To what extent do self-assessed and objectively measured understanding predict academic performance uniquely and commonly?
2. How do general and academic self-efficacy predict academic performance?
3. To what extent do general and academic self-efficacy predict self-assessed and objectively measured understanding?
4. How are self-efficacy, self-assessed, and objectively measured understanding interrelated in predicting academic performance?

Method

Procedure

The course investigated in the present study was part of the bachelor's degree program in Psychology at the University of Bern. In the first week, students completed a prior knowledge test assessing general statistical knowledge (such as knowledge about central tendencies and measures of variability, covariances and correlations) and an online personality assessment using the survey software Qualtrics (https://www.qualtrics.com), which included the self-efficacy assessment. Students also responded to further items assessing students' attitudes towards mathematical tasks. The course was offered in a blended format with in-person lectures (concurrently recorded as podcasts) covering topics of the comprehensive psychological assessment process, including formulation of hypotheses and specific questions that can be the subject of psychological examinations, types of operationalisation (e.g. interviews, performance tests, questionnaires), classic test theory, quality criteria of tests and assessment tools (reliability, validity, norms, etc.), decision rules, decision strategies, the quality criteria of decisions (e.g. sensitivity, specificity, positive and negative predictor values) and a short introduction to the assessment report. Nine OFAs were developed collaboratively by the course professor and teaching assistants to closely align with lecture content. Each OFA began with students assessing their understanding of the topics covered by the previous lecture (or two previous lectures), followed by six to ten tasks. The OFAs were explicitly designed to reflect the key learning objectives of the course and included conceptual, applied, and calculation-based tasks directly derived from the course content. After each task, students received feedback explaining why their chosen solution was correct or incorrect. Three days after completing the OFAs, the tasks of the last OFA were discussed in class to clarify questions not answered by the feedback or to discuss how to apply the theoretical basics in practical diagnostics and assessments. Two weeks before the final exam, a mock exam covered the entire term's content, providing an opportunity to assess students' knowledge in advance of the final exam (see Figure 1 for the structure).

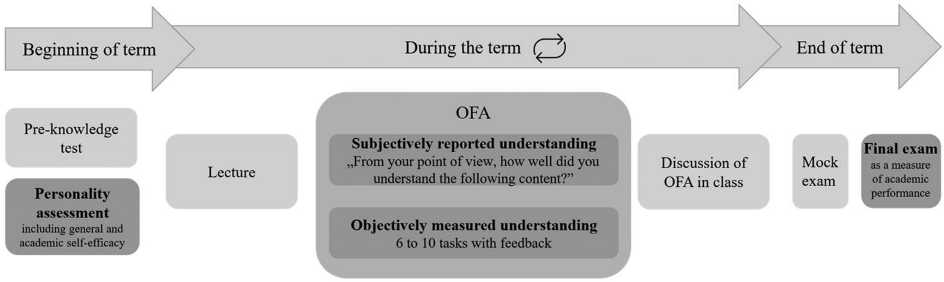


Figure 1. Structure of the term for the bachelor's lecture psychological assessment at the University of X.
Note. Variables considered in the present study are highlighted in bold and shaded darker.

Participants

Data from students across two years (2023 and 2024) were combined and analysed. Over the two years, a total of 578 students took the final exam of the course. The course structure and the content of the lectures and assessments remained consistent. In 2024, the accompanying case study was updated, with the example being revised, while the overall content and procedures remained unchanged. However, substantial content and structural revisions to the first OFA between the two years made its inclusion unfeasible, resulting in its exclusion and leaving eight OFAs for analysis. To justify combining data across cohorts and to ensure comparability of relationships between variables, we tested for at least metric measurement invariance, which is necessary when comparing relationships between variables across two or more groups (Thompson, 2016). Separate analyses were conducted per cohort to rule out systematic bias due to sample differences (see Supplementary Appendix A). The results of this comparison support the comparability of these two subsamples, with at least metric invariance for the measures of self-efficacy and self-assessed understanding, and partial metric invariance for objectively measured understanding, with only one out of eight measures (OFA 6) having a significantly different factor loading in the two cohorts. Schmitt and Ali (2015) emphasised that partial (metric) invariance can be considered of minor relevance as long as at least half of the manifest variables have the same factor loadings in the two groups to be compared or combined. Therefore, the level of comparability of the two groups could be assumed to be sufficient to combine them without biasing the relationships. Participation in the OFAs was a mandatory course component but was not actively monitored. At the end of the term, 401 students completed most of the OFAs with no more than two OFAs missing and provided consent for their data to be used for research purposes. To be included in the study, students must have completed the questionnaire battery at the beginning of the term, including the self-efficacy measures, and voluntarily submit their final exam grade under the same pseudonym they used for the OFAs. Based on these criteria, an additional 127 students were excluded. Three students were excluded from further analyses due to final exam grades of 2.0 or below (on a scale from 1.0 to 6.0, with 4.0 or higher indicating a passing grade). This exclusion aimed to minimise potential distortions caused by incomplete submissions from students who primarily attended to familiarise themselves with the exam format and prepare for a later exam, often submitting blank or random answers. This resulted in a final sample of 271 students (135 from 2023 and 136 from 2024).

To protect students' privacy, all data were encrypted using self-generated pseudonyms, which were not accessible to instructors or researchers at any time. No personal information (e.g. gender or age) was collected, ensuring complete anonymity. Participation in the study was voluntary; students who did not consent to share their data still had full access to all learning materials. The local ethics committee of the human science faculty of the University of Bern reviewed this study and had no objections to the planned procedure (Nr. 2021-12-00004).

Measures

Self-assessed understanding

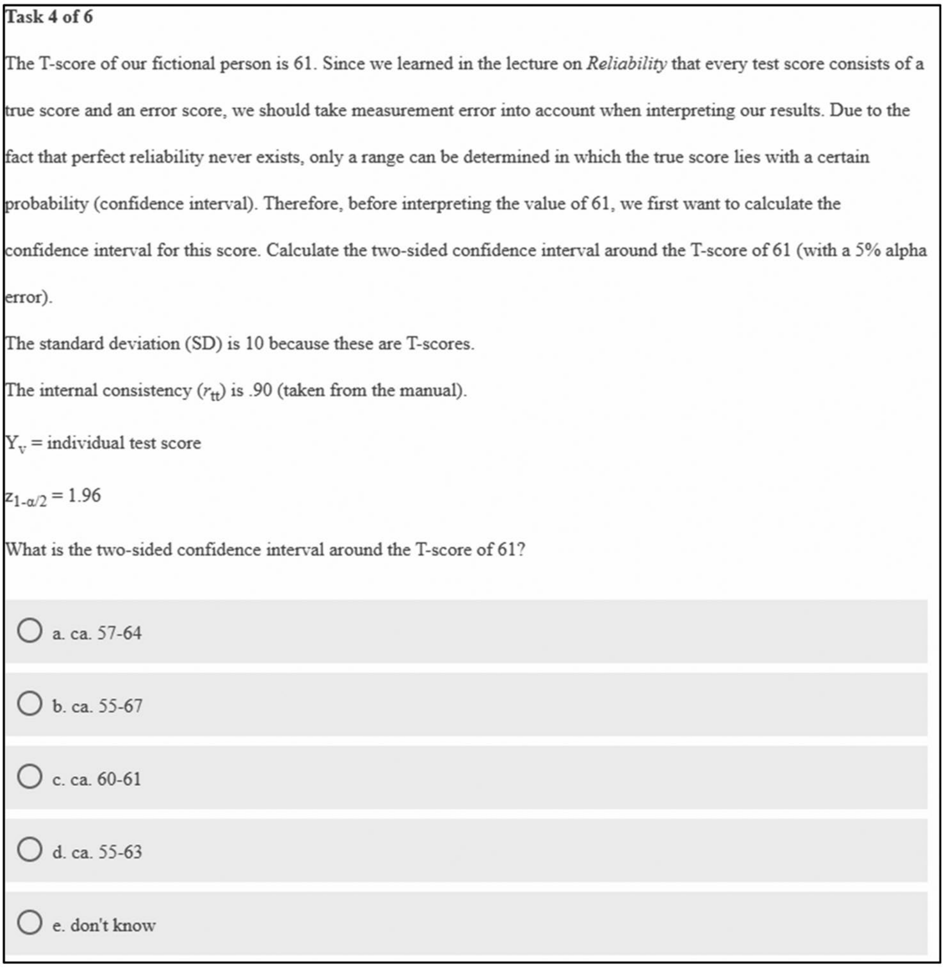
Before completing the tasks in each OFA, students were asked to rate their understanding of specific topics covered in the corresponding lecture. For example, they responded to questions such as, 'From your point of view, how well did you understand confidence intervals?' or '…how well did you understand internal consistency?' Responses were given on a percentage scale from 0 (no understanding) to 100 (complete understanding). Each OFA assessed one to five of these specific topic areas. To create a single score per OFA, an average was calculated across the self-ratings for all included topics, resulting in one self-assessed understanding score per OFA, ranging from 0 to 100.

Objectively measured understanding

Each OFA consisted of six to ten tasks. These tasks could have a single- or multiple-choice format (with one or more correct answers among four alternatives), a Kprim format (consisting of four statements per task to be judged as true or false), or be open-ended. Solving these tasks sometimes required performing calculations or conducting analyses in R (R Core Team, 2024). For illustrative examples, see Figures 2 and 3. Correct responses were awarded one point, while incorrect responses received zero points. In Kprim tasks, 0.25 points were given for each correctly judged statement. Each OFA's total score was calculated by summing the points across all tasks.

General self-efficacy

Jerusalem and Schwarzer's (2003) short-scale for general self-efficacy (WIRKALL-K) was used to assess general self-efficacy at the beginning of the term. The scale consisted of ten items, such as 'I am confident in my ability to handle unexpected events effectively.' and could be responded to on a four-point response scale ranging from 1 ('not agree') to 4 ('completely agree'). The full scale is openly accessible at https://www.testarchiv.eu/de/test/9001003. The short scale showed sufficient internal consistency (Cronbach's $\alpha = .82$ to .86; Jerusalem & Schwarzer, 2003). Furthermore, the scale demonstrated discriminant validity with related constructs such as self-esteem ($r = .49$ to .55), while strong negative correlations with general helplessness in student samples ($r = -.75$) provide evidence for convergent validity (Jerusalem & Schwarzer, 1986).



Task 4 of 6

The T-score of our fictional person is 61. Since we learned in the lecture on *Reliability* that every test score consists of a true score and an error score, we should take measurement error into account when interpreting our results. Due to the fact that perfect reliability never exists, only a range can be determined in which the true score lies with a certain probability (confidence interval). Therefore, before interpreting the value of 61, we first want to calculate the confidence interval for this score. Calculate the two-sided confidence interval around the T-score of 61 (with a 5% alpha error).

The standard deviation (SD) is 10 because these are T-scores.

The internal consistency (r_{tt}) is .90 (taken from the manual).

Y_v = individual test score

$z_{1-\alpha/2} = 1.96$

What is the two-sided confidence interval around the T-score of 61?

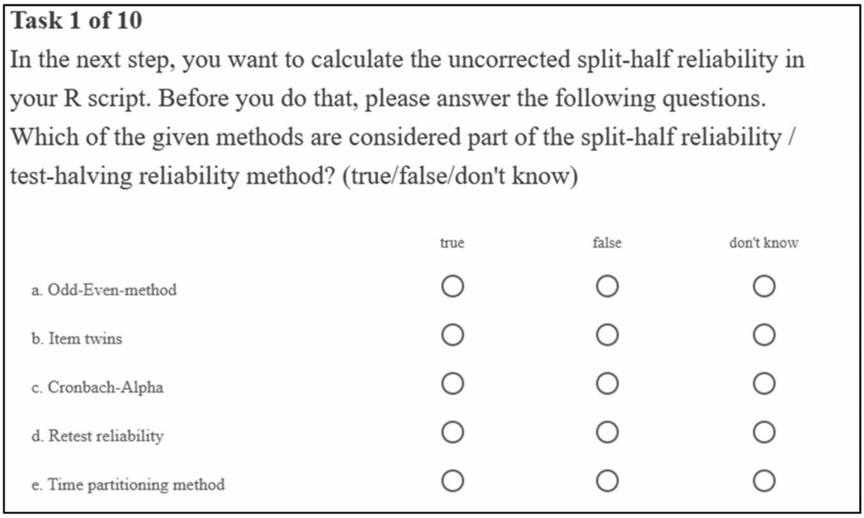
- a. ca. 57-64
- b. ca. 55-67
- c. ca. 60-61
- d. ca. 55-63
- e. don't know

Figure 2. Example of a single-choice task used in an OFA in the bachelor's psychological assessment course.

Moreover, the scale's relevance for academic contexts was supported by significant correlations between general self-efficacy and academic performance, underscoring its predictive value for learning outcomes (Luszczynska et al., 2005). For the descriptive statistics, the values were averaged over the ten items.

Academic self-efficacy

As a measure of academic self-efficacy, Jerusalem and Schwarzer's (2003) academic self-efficacy scale (WIRKSTUD) was applied, containing seven items (e.g. 'In my studies, I am always capable of achieving the required performance'). Responses were given on a four-point response scale ranging from 1 ('does not apply at all') to 4 ('totally applies'), with one item reverse-coded. The full scale is openly accessible at https://www.testarchiv.eu/de/test/9001003. The authors reported a good internal consistency (Cronbach's $\alpha = .87$; Jerusalem & Schwarzer, 2003). The scale is clearly differentiated from other constructs, such as self-esteem ($r = .37$) and anxiety in performance



Task 1 of 10

In the next step, you want to calculate the uncorrected split-half reliability in your R script. Before you do that, please answer the following questions.

Which of the given methods are considered part of the split-half reliability / test-halving reliability method? (true/false/don't know)

	true	false	don't know
a. Odd-Even-method	○	○	○
b. Item twins	○	○	○
c. Cronbach-Alpha	○	○	○
d. Retest reliability	○	○	○
e. Time partitioning method	○	○	○

Figure 3. Example of a Kprim task used in an OFA in the bachelor's psychological assessment course.

situations ($r = -.49$) (Jerusalem & Schwarzer, 1986). Additionally, significant relationships between academic self-efficacy and performance underscored the scale's relevance for academic contexts (Taşdemir, 2019). The test score was the average of the responses to the seven items.

Academic performance

The final exam grade at the end of the term was used to measure academic performance. The final exam covered the content of the lectures, the OFAs, and the mandatory literature. It included 15 single-choice tasks (one correct answer out of four alternatives) and 15 Kprim tasks (four statements per task to be judged for correctness). Each correctly answered single-choice task was awarded one point; incorrect answers received zero points. For Kprim tasks, all four statements had to be correct to earn one point. If three out of four statements were correct, 0.5 points were awarded; otherwise, no points were given. To pass the exam, students had to achieve at least 70% (21 points) of the maximum 30 points available, corresponding to a grade of 4.0. Grades ranged from 1.0 (lowest) to 6.0 (highest), with half-grade increments (e.g. 1.0, 1.5, 2.0, etc.). The two years did not differ in terms of mean grades or variance, $t(269) = -0.72$, $p = .474$. Nevertheless, the grades were z-standardised before combining data from the two years.

Data analyses

The data were analysed using R (version 4.4.1) and R-Studio (version 2024.12.0; R Core Team, 2024). To ensure a sufficient data basis for valid imputations, only students who have maximally missed two OFAs were included in the analyses (Enders, 2010). This supports the plausibility of imputed values under the Missing at Random assumption (Rubin, 1987), while balancing data quality, sample size, and the risk of bias due

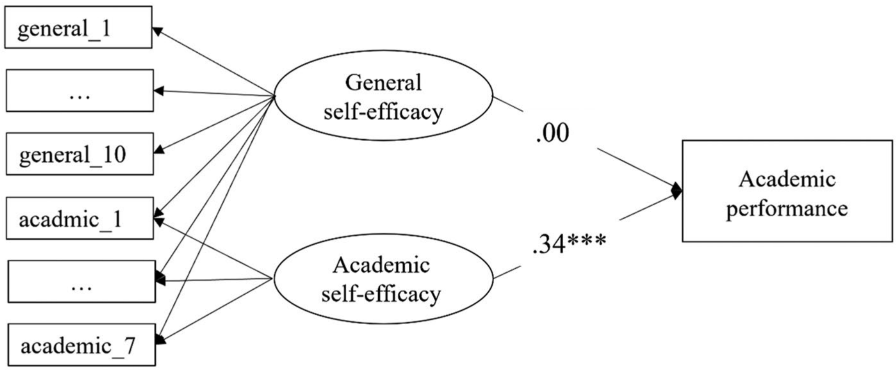


Figure 4. Latent regression model with bifactorial self-efficacy predicting later academic performance (N = 271).
Note. ***$p <$ .001.

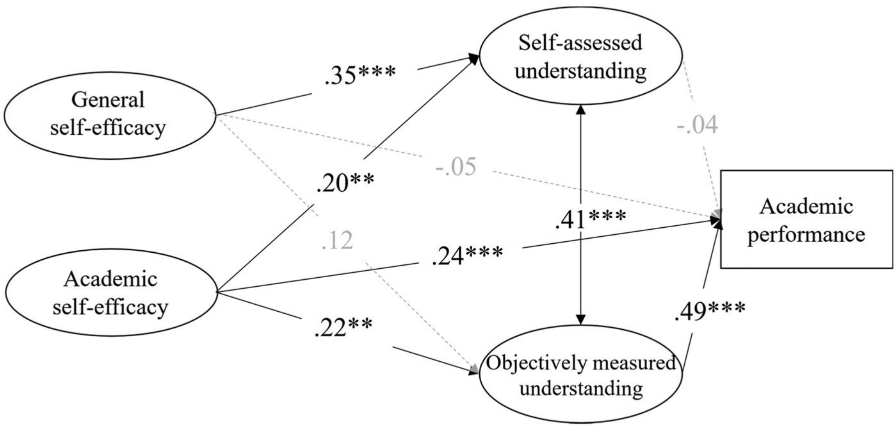


Figure 5. Latent regression model with general and academic self-efficacy, self-assessed and objectively measured understanding, predicting later academic performance (N = 271).
Note. *$p <$ .05, **$p <$ .01, ***$p <$ .001, one-tailed. Non-significant paths are depicted in grey and with dashed lines. Self-efficacy was modelled bi-factorially but presented in simplified form for clarity. The association between self-assessed and objectively measured understanding is modelled as a correlation.

to selective case exclusion. Of the 271 students, 27 (9.93%) missed one, and 11 (4.06%) missed two OFA(s). Little's (1988) MCAR test indicated that the missing OFA data can be considered Missing Completely at Random, $\chi^2(2069) = 2088.44$, $p = .378$. Missing data for the incomplete OFAs were imputed using multiple imputations by chained equations with the mice package (van Buuren & Groothuis-Oudshoorn, 2011). Each measure of interest was imputed separately using predictive mean matching (Little, 1988). Missing data on the self-efficacy measures at the beginning and the final exam grades at the end of the term could not be imputed, as these were one-time assessments with no additional data points available for imputation.

To verify the assumed unidimensionality of the constructs, confirmatory factor analyses (CFAs) were initially conducted using the *lavaan* package (Rosseel, 2012) in R. For each of the eight measures of self-assessed and objectively measured understanding, a separate single-factor model was tested. Additionally, the assumption of unidimensionality was examined separately for the two self-efficacy scales. CFAs were estimated using maximum likelihood with Satorra-Bentler correction (MLM). Mean scores were computed only if unidimensionality was confirmed. Based on these mean scores, descriptive statistics and correlations were calculated.

Given that self-efficacy was assessed at two nested levels of specificity (general and academic), we specified a bifactor model using MLM estimation to separate general self-efficacy from the measure of academic self-efficacy, following Gegenfurtner (2022). General self-efficacy was modelled as a general factor capturing variance common to all items, reflecting a broad sense of competence across domains. Academic self-efficacy was modelled with loadings only on its respective items, capturing variance unique to the academic level.

Self-assessed and objectively measured understanding, as well as the two levels of self-efficacy, were first analysed in separate latent regression analyses. In a final latent regression model, all predictors were included simultaneously, allowing for the estimation of their unique contributions to the explained variance in academic performance.

The model fits were evaluated using the comparative fit index (CFI), the root mean square error of approximation (RMSEA), and the standardised root mean square residual (SRMR) indices. A good (acceptable) model fit was indicated by CFI≥.95 (≥.90), SRMR < .08 (< .10), and RMSEA < .05 (< .08) (DiStefano, 2016; Schweizer, 2010). To ensure completeness, chi-square statistics were calculated but not interpreted due to their sensitivity to sample size (Hu & Bentler, 1999).

Results

Descriptive statistics and reliability of the measures of academic performance, the two levels of self-efficacy, and the self-assessed and objectively measured understanding (averaged across the eight OFAs during the term) are presented in Table 1. For the intercorrelations of the five variables on a manifest level, see Table 2. All variables except general self-efficacy were related to academic performance, whereby objectively measured understanding correlated more strongly with academic performance than self-assessed understanding, $Z = 3.384$, $p < .001$.

For both self-assessed understanding, $\chi^2 = 26.615$, df = 17, $p = .064$, CFI = .988, RMSEA = .058, SRMR = .028, and objectively measured understanding, $\chi^2 = 24.451$, df = 20, $p = .223$, CFI = .983, RMSEA = .032, SRMR = .042, a one-factor model with the eight measures across the term as manifest variables described the data well. It should be noted, however, that residual correlations had to be allowed for self-assessed understanding (see Supplementary Appendix B). Both variables predicted academic performance at the end of the term when submitted separately to a latent regression analysis. The regression coefficients were $\beta = .21$ and $\beta = .52$ for self-assessed and objectively measured understanding, respectively. When self-assessed and objectively measured understanding were used concurrently to predict academic performance, only objectively measured understanding remained a significant predictor, $\beta = .54$, $p < .001$. Self-assessed understanding was no longer a significant predictor, $\beta = -.03$, $p = .690$. Self-assessed and objectively measured understanding were latently correlated

Table 1. Descriptive statistics on a manifest level for academic performance, general and academic self-efficacy, and self-assessed and objectively measured understanding (N = 271).

	M	SD	Min.	Max.	Omega
1. Academic performance	4.43	0.81	2.50	6.00	–
2. General self-efficacy	2.82	0.38	1.30	3.90	.84
3. Academic self-efficacy	2.61	0.44	1.29	3.71	.78
4. Self-assessed understanding	64.59	13.24	10.34	92.19	.94
5. Objectively measured understanding	5.52	0.90	2.54	7.28	.81

Table 2. Pearson correlations on a manifest level for academic performance, general and academic self-efficacy, and self-assessed and objectively measured understanding.

	1.	2.	3.	4.
1. Academic performance	–			
2. General self-efficacy	.01	–		
3. Academic self-efficacy	.22***	.49***	–	
4. Self-assessed understanding	.19***	.30***	.31***	–
5. Objectively measured understanding	.45***	.09	.20***	.36***

Note. *$p < .05$; **$p < .01$; ***$p < .001$, one-tailed.

with $r = .44$, $p < .001$. The model described the data well, $\chi^2 = 153.976$, df = 114, $p = .008$, CFI = .972, RMSEA = .040, SRMR = .053, when considered together, and explained 27.5% of the variance in academic performance. Accordingly, only objectively measured understanding explained for a unique portion of variance in academic performance, whereas the predictive value of self-assessed understanding was fully captured by objectively measured understanding.

For the latent representation of self-efficacy, a bifactor model was computed where all items loaded on the latent variable *general self-efficacy*, while only more specific items were used to extract latent variables for *academic self-efficacy*. An alternative two-factor model yielded highly correlated factors; the bifactor model was therefore preferred, as it isolates the specific variance of academic self-efficacy while representing shared variance in a general factor. Notably, the pattern of results remained robust regardless of the choice of representation. This bifactor model described the students' responses well, $\chi^2 = 186.020$, df = 107, $p < .001$, CFI = .929, RMSEA = .052, SRMR = .055.

Regressing the two levels of self-efficacy on academic performance led to an acceptable to good model fit, $\chi^2 = 197.695$, df = 121, $p < .001$, CFI = .933, RMSEA = .051, SRMR = .054. Academic self-efficacy emerged as a significant predictor ($\beta = .34$, $p < .001$), whereas general self-efficacy showed no predictive value ($\beta = .00$, $p = .971$). Overall, 11.6% of the variance in academic performance was explained by self-efficacy (see Figure 4.).

For a comprehensive view, the bifactorial model for self-efficacy was integrated into a latent regression model. In this model, both self-assessed and objectively measured understanding were specified as latent variables that were allowed to correlate with each other. All four latent predictors, general and academic self-efficacy, as well as self-assessed and objectively measured understanding, were regressed onto (the manifest variable of) academic performance (see Figure 5.). The model described the data well, $\chi^2 = 650.671$, df = 501, $p < .001$, CFI = .949, RMSEA = .034, SRMR = .054.

About 33% of the variance in academic performance was explained. A substantial portion of the effects of the self-efficacy measures on academic performance were mediated by both self-assessed and objectively measured understanding. Only academic self-efficacy explained a unique portion of the variance in academic performance, with a decrease from $\beta = .34$ to $\beta = .24$ when controlling for the measures of understanding. Both self-efficacy measures, but especially general self-efficacy, were associated with self-assessed understanding. Only academic self-efficacy was associated with objectively measured understanding.

Given that prior knowledge has been often shown to be relevant in the context of self-efficacy and academic performance (Honicke & Broadbent, 2016; Sitzmann & Yeo, 2013), an additional analysis was conducted, including students' prior knowledge as a control variable. The overall pattern of relationships remains unchanged when prior knowledge, which captures general statistical knowledge at the beginning of the term, was included in the final model (see Supplementary Appendix C).

Discussion

The positive influence of SRL on academic performance, especially in tertiary education, has been extensively demonstrated (Broadbent & Poon, 2015; Ergen & Kanadli, 2017; Zhao et al., 2025). A core aspect of SRL is the ability to assess one's own understanding accurately, which informs subsequent learning decisions (Hadwin et al., 2025). This understanding can be either self-assessed or objectively measured. Here, we investigated whether these two kinds of assessment explain the same portion of variance in academic performance. Furthermore, we aimed to clarify the role of self-efficacy in the interplay with self-assessed and objectively measured understanding when examining their association with academic performance.

Separate regression analyses revealed that both self-assessed and objectively measured understanding in the OFAs during the term significantly predicted academic performance at the end of the term, with objectively measured understanding showing a stronger association. This became especially evident when both measures of understanding were entered into a common regression model, where objectively measured understanding fully explained the effect of self-assessed understanding. Among the measures of self-efficacy, only academic, but not general self-efficacy, shared unique variance with academic performance. Both general and academic self-efficacy predicted self-assessed understanding, but only academic self-efficacy was directly predictive of objectively measured understanding. Finally, when both levels of self-efficacy were examined together with self-assessed and objectively measured understanding in a latent regression model, only academic self-efficacy and objectively measured understanding were significantly associated with academic performance.

As in previous studies, both self-assessed and objectively measured understanding explained a reasonable portion of variance in academic performance (Richardson et al., 2012; Yan, 2020). However, their direct comparison in the present study revealed that the objective measure of understanding during the term is a much better predictor of the final grades at the end of the term compared to the self-assessed understanding during the term (Fuchs & Fuchs, 1986; Kuncel et al., 2001; Žuljević & Buljan, 2022). Consequently, students would be well-advised to consider objective measures rather than subjective self-assessments of understanding when adapting their learning strategies. However, this requires that students are able to actively process such objective insights as feedback (Gibbs & Simpson, 2005). To optimally support students in this process, educational approaches emphasise encouraging students to critically reflect on feedback and giving them the opportunities to practice (Gibbs & Simpson, 2005; Nicol & Macfarlane-Dick, 2004). Furthermore, feedback delivery can be facilitated through dashboards that visualise indicators such as the percentage of correctly solved tasks or comparisons with peers in the same class (Cavalcanti et al., 2021; Masiello et al., 2024; Rets et al., 2021). Importantly, to be effective, feedback should not only address content-related aspects but also be tailored to students' level of SRL and self-efficacy (Brown et al., 2016).

In addition to the level of understanding, self-efficacy at the beginning of the term predicted academic performance at the end of the term, as also reported in previous studies (Honicke & Broadbent, 2016). However, this effect was limited to academic self-efficacy, whereas general self-efficacy showed no meaningful association with the academic performance at the end of the term (neither in the correlational analyses nor in the latent regression models). These findings align with the need for specificity between predictor and outcome (Bandura, 2012; Choi, 2005; Wittmann, 1988). Academic self-efficacy refers to the demands of studying (such as preparation and organization) and, especially, focuses on dealing with the exam situation (such as the belief in the ability to perform under exam conditions). It might be this aspect of academic self-efficacy that is uniquely related to the grade in the final exam, given that the examination itself presents a challenge, requiring students to demonstrate their knowledge and manage stress, cope with test anxiety, and perform under pressure (Ostermann et al., 2022; von der Embse et al., 2018). In this light, academic self-efficacy may serve as a psychological buffer, influencing students' perceptions and their ability to perform in examination situations. Therefore, academic self-efficacy accounts for a unique share of variance in academic performance that is independent of both self-assessed and objectively measured understanding.

Regarding the interplay between self-efficacy and the level of understanding, both self-assessed and objectively measured understanding were similarly predicted by academic self-efficacy. This result indicates that there is a positive influence of academic self-efficacy on the level of understanding, which might reflect students' confidence in approaching learning. The nominally strongest relationship, however, was found between general self-efficacy and self-assessed understanding, which was not significant for objectively measured understanding. This differential result might suggest a bias in the self-assessments of the level of understanding. General self-efficacy has been reported to be closely related to general traits such as hope and optimism (Feldman & Kubota, 2015). Therefore, the relationship between self-assessment of the level of understanding and general self-efficacy might have been artificially inflated when hope or optimism affected both the self-assessment of the level of understanding as well as self-efficacy. This potential bias in the self-assessed understanding would be worth investigating in future studies.

Although we suspect the self-assessment of understanding to be biased in the present study, it should be noted that self-assessments remain important as they are substantial when focusing on individuals' internal perspectives, including perceived competence, task value, and personal interest or engagement (Eccles & Wigfield, 2002; Ryan & Deci, 2000; Zimmerman, 2000). Thus, self-assessments provide valuable insights into students' perceived understanding, but they represent only one part of the picture (Moskal, 2010). For a comprehensive assessment of students' learning and understanding, particularly when designing interventions or providing targeted support, objective measures are essential and should be taken into account.

Limitations

The present study has some important limitations that should be considered. The results offer new perspectives but cannot be easily generalised. The sample is highly specific, consisting solely of psychology students enrolled in the course Psychological Assessment. Even generalising to other psychology courses may be problematic, as the constellation of motivational variables and learning dynamics is likely to differ substantially depending on the course content. Due to data pseudonymization, no further demographics were available to protect participants' identities, although factors such as gender may still have an influence (Huang, 2013; Voyer & Voyer, 2014). Moreover, there is a considerable selection bias within this group, as the analyses only included students who completed the initial questionnaires, a substantial portion of the OFAs during the term, and provided their final grade at the end of the term. This, combined with the fact that participation was not enforced, also substantially reduced the sample size.

It should also be noted that the measure of objectively assessed understanding is not entirely free from potential confounds. Since the OFAs were completed remotely, testing conditions could not be controlled, such as whether students worked individually or in small groups, or whether they engaged in other activities simultaneously. As such, the resulting scores may not perfectly reflect their individual performance or actual level of understanding. Moreover, modelling objective understanding highlights a central tension in practice-oriented approaches: maintaining comparability across cohorts while allowing ongoing empirical refinement.

Finally, differences in students' learning strategies, depending on whether understanding was self-assessed or objectively measured, were not examined. Future studies could usefully explore how learning strategies adapt dynamically, particularly in response to feedback on current levels of understanding.

Conclusion

The present study provides converging evidence for the relevance of SRL and objectively measured understanding (e.g. OFAs) in predicting academic performance. Within a blended learning context, objectively measured understanding proved more predictive of academic performance than students' self-assessed understanding. While self-assessed understanding appeared to reflect both self-efficacy and actual understanding, it seems less suited for effectively monitoring the learning progress compared to objectively measured understanding. Furthermore, academic self-efficacy emerged as a direct and meaningful predictor of academic performance. These findings underscore the value of designing learning environments that integrate more objective, performance-based assessments with timely, content-specific feedback. Such environments enable students to gain a clearer picture of their actual understanding and adapt their learning strategies accordingly, thereby supporting accurate self-monitoring, fostering academic self-efficacy, and promoting SRL. Together, these factors may enhance the chance of success in tertiary education.

Ethical approval and informed consent statements

The ethics committee of the Faculty of Human Sciences of the University of Bern approved this study (submission number 2021-12-00004). All participants gave their informed consent for their data collected during the course to be used for research purposes.

Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this work, the authors used ChatGPT (version GPT-4o) to support linguistic revision. The tool was applied section by section to improve grammar, clarity, and academic tone. A typical prompt used was: *'Please revise the following section to improve grammar and academic style, without changing the meaning or scientific content: [insert text].'* After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

Disclosure statement

The authors have no conflicts of interest to declare.

Funding

This work was supported by a BeLEARN grant.

ORCID

Ursina E. Raemy http://orcid.org/0009-0005-6582-4144
Natalie Borter http://orcid.org/0009-0005-8742-4643
Mathias Mejeh http://orcid.org/0000-0003-4923-8936
Stefan J. Troche http://orcid.org/0000-0002-0961-1081

Data availability statement

https://osf.io/m6fd9

References

Andrade, H. L. (2019). A critical review of research on student self-assessment. *Frontiers in Education*, *4*, 87. https://doi.org/10.3389/feduc.2019.00087

Anthonysamy, L., Koo, A. C., & Hew, S. H. (2020). Self-regulated learning strategies in higher education: Fostering digital literacy for sustainable lifelong learning. *Education and Information Technologies*, *25*(4), 2393–2414. https://doi.org/10.1007/s10639-020-10201-8

Bandura, A. (1997). *Self-efficacy: The exercise of control.* Freeman and Company.
Bandura, A. (2012). On the functional properties of perceived self-efficacy revisited. *Journal of Management*, *38*(1), 9–44. https://doi.org/10.1177/0149206311410606
Bennett, R. E. (2011). Formative assessment: A critical review. *Assessment in Education: Principles, Policy & Practice*, *18*(1), 5–25. https://doi.org/10.1080/0969594X.2010.513678
Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*, *5*(1), 7–74. https://doi.org/10.1080/0969595980050102
Brewer, D. J., & McEwan, P. J. (2010). *Economics of education.* Elsevier.
Broadbent, J., & Fuller-Tyszkiewicz, M. (2018). Profiles in self-regulated learning and their correlates for online and blended learning students. *Educational Technology Research and Development*, *66*(6), 1435–1455. https://doi.org/10.1007/s11423-018-9595-9
Broadbent, J., & Poon, W. L. (2015). Self-regulated learning strategies & academic achievement in online higher education learning environments: A systematic review. *Internet and Higher Education*, *27*, 1–13. https://doi.org/10.1016/j.iheduc.2015.04.007
Brown, G. T. L., Andrade, H. L., & Chen, F. (2015). Accuracy in student self-assessment: Directions and cautions for research. *Assessment in Education: Principles, Policy & Practice*, *22*(4), 444–457. https://doi.org/10.1080/0969594X.2014.996523
Brown, G. T. L., Peterson, E. R., & Yao, E. S. (2016). Student conceptions of feedback: Impact on self-regulation, self-efficacy, and academic achievement. *British Journal of Educational Psychology*, *86*(4), 606–629. https://doi.org/10.1111/bjep.12126
Cavalcanti, A. P., Barbosa, A., Carvalho, R., Freitas, F., Tsai, Y., S., Gašević, D., & Mello, R. F. (2021). Automatic feedback in online learning environments: A systematic literature review. *Computers and Education: Artificial Intelligence*, *2*, 100027. https://doi.org/10.1016/j.caeai.2021.100027
Cerezo, R., Bogarín, A., Esteban, M., & Romero, C. (2020). Process mining for self-regulated learning assessment in e-learning. *Journal of Computing in Higher Education*, *32*(1), 74–88. https://doi.org/10.1007/s12528-019-09225-y
Cheng, Z., Zhang, Z., Xu, Q., Maeda, Y., & Gu, P. (2025). A meta-analysis addressing the relationship between self-regulated learning strategies and academic performance in online higher education. *Journal of Computing in Higher Education*, *37*(1), 195–224. https://doi.org/10.1007/s12528-023-09390-1
Choi, N. (2005). Self-efficacy and self-concept as predictors of college students' academic performance. *Psychology in the Schools*, *42*(2), 197–205. https://doi.org/10.1002/pits.20048
Clark, I. (2012). Formative assessment: Assessment is for self-regulated learning. *Educational Psychology Review*, *24*(2), 205–249. https://doi.org/10.1007/s10648-011-9191-6
DiStefano, C. (2016). Examining fit with structural equation models. In K. Schweizer & C. DiStefano (Eds.), *Principles and methods of test construction: Standards and recent advances* (pp. 166–193). Hogrefe Publishing.
Dunlosky, J., & Metcalfe, J. (2008). *Metacognition*. Sage Publications.
Eccles, J. S., & Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annual Review of Psychology*, *53*(1), 109–132. https://doi.org/10.1146/annurev.psych.53.100901.135153
Efklides, A. (2011). Interactions of metacognition with motivation and affect in self-regulated learning: The MASRL model. *Educational Psychologist*, *46*(1), 6–25. https://doi.org/10.1080/00461520.2011.538645
Enders, C. K. (2010). The imputation phase of multiple imputation. In C. K. Enders (Ed.), *Applied missing data analysis* (pp. 187–216). The Guilford Press.
Ergen, B., & Kanadli, S. (2017). The effect of self-regulated learning strategies on academic achievement: A meta-analysis study. *Eurasian Journal of Educational Research*, *17*(69), 55–74. https://doi.org/10.14689/ejer.2017.69.4
Falchikov, N., & Boud, D. (1989). Student self-assessment in higher education: A meta-analysis. *Review of Educational Research*, *59*(4), 395–430. https://doi.org/10.3102/00346543059004395
Feldman, D. B., & Kubota, M. (2015). Hope, self-efficacy, optimism, and academic achievement: Distinguishing constructs and levels of specificity in predicting college grade-point average. *Learning and Individual Differences*, *37*, 210–216. https://doi.org/10.1016/j.lindif.2014.11.022

Fuchs, L. S., & Fuchs, D. (1986). Effects of systematic formative evaluation: A meta-analysis. *Exceptional Children*, *53*(3), 199–208. https://doi.org/10.1177/001440298605300301
Gallagher, M. W. (2012). Self-efficacy. In V. S. Ramachandran (Ed.), *Encyclopedia of human behavior* (pp. 314–320). Elsevier. https://doi.org/10.1016/B978-0-12-375000-6.00312-8
Gegenfurtner, A. (2022). Bifactor exploratory structural equation modeling: A meta-analytic review of model fit. *Frontiers in Psychology*, *13*, 1037111. https://doi.org/10.3389/fpsyg.2022.1037111
Gibbs, G., & Simpson, C. (2005). Conditions under which assessment supports students' learning. *Learning and Teaching in Higher Education*, *1*, 3–31.
Gikandi, J. W., Morrow, D., & Davis, N. E. (2011). Online formative assessment in higher education: A review of the literature. *Computers & Education*, *57*(4), 2333–2351. https://doi.org/10.1016/j.compedu.2011.06.004
Hadwin, A. F., Rostampour, R., & Winne, P. H. (2025). Advancing self-reports of self-regulated learning: Validating new measures to assess students' beliefs, practices, and challenges. *Educational Psychology Review*, *37*, 1-61. https://doi.org/10.1007/s10648-024-09977-9
Higgins, N. L., Rathner, J. A., & Frankland, S. (2023). Development of self-regulated learning: A longitudinal study on academic performance in undergraduate science. *Research in Science & Technological Education*, *41*(4), 1242–1266. https://doi.org/10.1080/02635143.2021.1997978
Honicke, T., & Broadbent, J. (2016). The influence of academic self-efficacy on academic performance: A systematic review. *Educational Research Review*, *17*, 63–84. https://doi.org/10.1016/j.edurev.2015.11.002
Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1–55. https://doi.org/10.1080/10705519909540118
Huang, C. (2013). Gender differences in academic self-efficacy: A meta-analysis. *European Journal of Psychology of Education*, *28*(1), 1–35. https://doi.org/10.1007/s10212-011-0097-y
Ifenthaler, D., & Drachsler, H. (2018). Learning analytics. In H. Niegemann & A. Weinberger (Eds.), *Lernen mit Bildungstechnologien* (pp. 1–20). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-54373-3_42-1
Irons, A., & Elkington, S. (2022). Enhancing learning through formative assessment and feedback. In *Key guides for effective teaching in higher education series* (2nd ed.). Routledge. https://doi.org/10.4324/9781138610514
Jerusalem, M., & Schwarzer, R. (1986). Selbstwirksamkeit [Self-efficacy]. In R. Schwarzer (Ed.), *Skalen zur Befindlichkeit und Persönlichkeit* (pp. 15–28). Institut für Psychologie, Freie Universität.
Jerusalem, M., & Schwarzer, R. (2003). *SWE – Skala zur Allgemeinen Selbstwirksamkeitserwartung [General self-efficacy expectation scale]*. https://doi.org/10.23668/psycharchives.4515
Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology*, *77*(6), 1121–1134. https://doi.org/10.1037//0022-3514.77.6.1121
Kulasegaram, K., & Rangachari, P. K. (2018). Beyond "formative": Assessments to enrich student learning. *Advances in Physiology Education*, *42*(1), 5–14. https://doi.org/10.1152/advan.00122.2017
Kuncel, N. R., Hezlett, S. A., & Ones, D. S. (2001). A comprehensive meta-analysis of the predictive validity of the graduate record examinations: Implications for graduate student selection and performance. *Psychological Bulletin*, *127*(1), 162–181. https://doi.org/10.1037/0033-2909.127.1.162
Lai, E. R. (2011). *Metacognition: A literature review*. Pearson Assessments Research Reports.
León, S. P., Panadero, E., & García-Martínez, I. (2023). How accurate are our students? A meta-analytic systematic review on self-assessment scoring accuracy. *Educational Psychology Review*, *35*, 1–106. https://doi.org/10.1007/s10648-023-09819-0
Li, M., & Zhang, X. (2021). A meta-analysis of self-assessment and language performance in language testing and assessment. *Language Testing*, *38*(2), 189–218. https://doi.org/10.1177/0265532220932481
Little, R. J. A. (1988). Missing-data adjustments in large surveys. *Journal of Business & Economic Statistics*, *6*(3), 287–296. https://doi.org/10.1080/07350015.1988.10509663

Luszczynska, A., Gutiérrez-Doña, B., & Schwarzer, R. (2005). General self-efficacy in various domains of human functioning: Evidence from five countries. *International Journal of Psychology*, *40*(2), 80–89. https://doi.org/10.1080/00207590444000041

Masiello, I., Mohseni, Z., Palma, F., Nordmark, S., Augustsson, H., & Rundquist, R. (2024). A current overview of the use of learning analytics dashboards. *Education Sciences*, *14*(1), 82. https://doi.org/10.3390/educsci14010082

McManus, S. (2008). *Attributes of effective formative assessment*. Council of Chief State School. https://csaa.wested.org/resource/attributes-of-effective-formative-assessment/

Moskal, B. M. (2010). Self-assessments: What are their valid uses? *Academy of Management Learning & Education*, *9*(2), 314–320. https://doi.org/10.5465/AMLE.2010.51428553

Neubauer, A. C., & Hofer, G. (2021). Self-estimates of abilities are a better reflection of individuals' personality traits than of their abilities and are also strong predictors of professional interests. *Personality and Individual Differences*, *169*, 109850. https://doi.org/10.1016/j.paid.2020.109850

Nicol, D., & Macfarlane-Dick, D. (2004). Rethinking formative assessment in HE: A theoretical model and seven principles of good feedback practice. In B. C. Juwah, D. Macfarlane-Dick, B. Matthew, D. Nicol, & D. Smith (Eds.), *Enhancing student learning though effective formative feedback*. The Higher Education Academy.

OECD. (2024). *Education at a glance 2024*. https://doi.org/10.1787/c00cad36-en

Ostermann, T., Pawelkiwitz, M., & Cramer, H. (2022). The influence of mindfulness-based interventions on the academic performance of students measured by their GPA. A systematic review and meta-analysis. *Frontiers in Behavioral Neuroscience*, *16*, 961070. https://doi.org/10.3389/fnbeh.2022.961070

Panadero, E. (2017). A review of self-regulated learning: Six models and four directions for research. *Frontiers in Psychology*, *8*, 422. https://doi.org/10.3389/fpsyg.2017.00422

Panadero, E., Jonsson, A., & Botella, J. (2017). Effects of self-assessment on self-regulated learning and self-efficacy: Four meta-analyses. *Educational Research Review*, *22*, 74–98. https://doi.org/10.1016/j.edurev.2017.08.004

Paulhus, D. L., & Vazire, S. (2007). The self-report method. *Handbook of Research Methods in Personality Psychology*, *1*, 224–239.

Phan, H. P. (2016). Longitudinal examination of optimism, personal self-efficacy and student well-being: A path analysis. *Social Psychology of Education*, *19*(2), 403–426. https://doi.org/10.1007/s11218-015-9328-4

Pinquart, M., & Ebeling, M. (2020). Students' expected and actual academic achievement – A meta-analysis. *International Journal of Educational Research*, *100*, 101524. https://doi.org/10.1016/j.ijer.2019.101524

Pintrich, P. R. (2000). The role of goal orientation in self-regulated learning. In M. Boekaerts, P. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 451–502). Elsevier. https://doi.org/10.1016/B978-012109890-2/50043-3

R Core Team. (2024). *R: A language and environment for statistical computing (Version 4.4.1) [Computer software]*. R Foundation for Statistical Computing. https://www.R-project.org/

Rammstedt, B., & Rammsayer, T. H. (2002). Self-estimated intelligence. *European Psychologist*, *7*(4), 275–284. https://doi.org/10.1027//1016-9040.7.4.275

Rets, I., Herodotou, C., Bayer, V., Hlosta, M., & Rienties, B. (2021). Exploring critical factors of the perceived usefulness of a learning analytics dashboard for distance university students. *International Journal of Educational Technology in Higher Education*, *18*(1), 46. https://doi.org/10.1186/s41239-021-00284-9

Richardson, M., Abraham, C., & Bond, R. (2012). Psychological correlates of university students' academic performance: A systematic review and meta-analysis. *Psychological Bulletin*, *138*(2), 353–387. https://doi.org/10.1037/a0026838

Robinson, S. E., & Song, J. J. (2019). Student academic performance system: Quantitative approaches to evaluating and monitoring student progress. *International Journal of Quantitative Research in Education*, *4*(4), 332–353. https://doi.org/10.1504/IJQRE.2019.100170

Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, *48*(2), 1-36. https://doi.org/10.18637/jss.v048.i02

Rubin, D. B. (1987). *Multiple imputation for nonresponse in surveys*. Wiley. https://doi.org/10.1002/9780470316696

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, *55*(1), 68–78. https://doi.org/10.1037/0003-066X.55.1.68

Schmitt, N., & Ali, A. A. (2015). The practical importance of measurement invariance. In C. E. Lance & R. J. Vandenberg (Eds.), *More statistical and methodological myths and urban legends* (pp. 327–346). Routledge.

Schunk, D. H. (1990). Goal setting and self-efficacy during self-regulated learning. *Educational Psychologist*, *25*(1), 71–86. https://doi.org/10.1207/s15326985ep2501_6

Schunk, D. H., & DiBenedetto, M. K. (2016). Self-efficacy theory in education. In D. B. Miele & K. R. Wentzel (Eds.), *Educational psychology handbook series. Handbook of motivation at school* (2nd ed., pp. 34–54). Routledge.

Schwarz, N. (2015). Metacognition. In M. Mikulincer, P. R. Shaver, E. Borgida, & J. A. Bargh (Eds.), *APA handbook of personality and social psychology. Volume 1: Attitudes and social cognition* (pp. 203–229). American Psychological Association. https://doi.org/10.1037/14341-006

Schweizer, K. (2010). Some guidelines concerning the modeling of traits and abilities in test construction. *European Journal of Psychological Assessment*, *26*(1), 1–2. https://doi.org/10.1027/1015-5759/a000001

Sitzmann, T., & Yeo, G. (2013). A meta-analytic investigation of the within-person self-efficacy domain: Is self-efficacy a product of past performance or a driver of future performance? *Personnel Psychology*, *66*(3), 531–568. https://doi.org/10.1111/peps.12035

Talsma, K., Schüz, B., Schwarzer, R., & Norris, K. (2018). I believe, therefore I achieve (and vice versa): A meta-analytic cross-lagged panel analysis of self-efficacy and academic performance. *Learning and Individual Differences*, *61*, 136–150. https://doi.org/10.1016/j.lindif.2017.11.015

Taşdemir, C. (2019). An investigation of academic self-efficacy perceptions of primary mathematics teacher candidates. *Higher Education Studies*, *9*(2), 72. https://doi.org/10.5539/hes.v9n2p72

Thompson, M. S. (2016). Assessing measurement invariance of scales using multiple-group structural equation modeling. In K. Schweizer & C. DiStefano (Eds.), *Principles and methods of test construction* (pp. 218–244). Hogrefe.

Tomasik, M. J., Berger, S., & Moser, U. (2018). On the development of a computer-based tool for formative student assessment: Epistemological, methodological, and practical Issues. *Frontiers in Psychology*, *9*, 2245. https://doi.org/10.3389/fpsyg.2018.02245

van Buuren, S., & Groothuis-Oudshoorn, K. (2011). mice: Multivariate imputation by chained equations in R. *Journal of Statistical Software*, *45*(3), 1-67. https://doi.org/10.18637/jss.v045.i03

Vierzigmann, G., & Lehmann, B. (2022). Herausforderungen in Hochschulbildung und Hochschullehre [Challenges in higher education and university teaching]. In L. Kolhoff (Ed.), *Perspektiven Sozialwirtschaft und Sozialmanagement. Aktuelle Diskurse in der Sozialwirtschaft IV* (pp. 79–94). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-38677-1_5

von der Embse, N., Jester, D., Roy, D., & Post, J. (2018). Test anxiety effects, predictors, and correlates: A 30-year meta-analytic review. *Journal of Affective Disorders*, *227*, 483–493. https://doi.org/10.1016/j.jad.2017.11.048

Vosniadou, S. (2020). Bridging secondary and higher education. The importance of self-regulated learning. *European Review*, *28*(S1), S94–S103. https://doi.org/10.1017/S1062798720000939

Voyer, D., & Voyer, S. D. (2014). Gender differences in scholastic achievement: A meta-analysis. *Psychological Bulletin*, *140*(4), 1174–1204. https://doi.org/10.1037/a0036620

Waddington, J. (2023). Self-efficacy. *ELT Journal*, *77*(2), 237–240. https://doi.org/10.1093/elt/ccac046

Williams, T., & Williams, K. (2010). Self-efficacy and performance in mathematics: Reciprocal determinism in 33 nations. *Journal of Educational Psychology*, *102*(2), 453–466. https://doi.org/10.1037/a0017271

Winne, P. H., & Hadwin, A. F. (1998). Studying as self-regulated engagement in learning. In D. Hacker, J. Dunlosky, & A. Graesser (Eds.), *Metacognition in educational theory and practice* (pp. 277–304). Erlbaum.

Winne, P. H., & Perry, N. E. (2000). Measuring self-regulated learning. In M. Boekaerts, P. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 531–566). Elsevier. https://doi.org/10.1016/B978-012109890-2/50045-7

Wittmann, W. W. (1988). Multivariate reliability theory. In J. R. Nesselroade & R. B. Cattell (Eds.), *Handbook of multivariate experimental psychology. Perspectives on individual differences* (pp. 505–560). Springer. https://doi.org/10.1007/978-1-4613-0893-5_16

Yan, Z. (2020). Self-assessment in the process of self-regulated learning and its relationship with academic achievement. *Assessment & Evaluation in Higher Education*, *45*(2), 224–238. https://doi.org/10.1080/02602938.2019.1629390

Yan, Z., & Brown, G. T. (2017). A cyclical self-assessment process: Towards a model of how students engage in self-assessment. *Assessment & Evaluation in Higher Education*, *42*(8), 1247–1262. https://doi.org/10.1080/02602938.2016.1260091

Yan, Z., Wang, X., Boud, D., & Lao, H. (2023). The effect of self-assessment on academic performance and the role of explicitness: A meta-analysis. *Assessment & Evaluation in Higher Education*, *48*(1), 1–15. https://doi.org/10.1080/02602938.2021.2012644

Zell, E., & Krizan, Z. (2014). Do people have insight into their abilities? A metasynthesis. *Perspectives on Psychological Science: A Journal of the Association for Psychological Science*, *9*(2), 111–125. https://doi.org/10.1177/1745691613518075

Zhao, Y., Li, Y., Ma, S., Xu, Z., & Zhang, B. (2025). A meta-analysis of the correlation between self-regulated learning strategies and academic performance in online and blended learning environments. *Computers & Education*, *230*, 105279. https://doi.org/10.1016/j.compedu.2025.105279

Zimmerman, B. J. (2000). Attaining self-regulation. In M. Boekaerts. P. Pintrich, &. M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 13–39). Elsevier. https://doi.org/10.1016/B978-012109890-2/50031-7

Zimmerman, B. J., & Schunk, D. H. (2011). Self-regulated learning and performance: An introduction and an overview. In D. H. Schunk & B. J. Zimmerman (Eds.), *Handbook of self-regulation of learning and performance* (pp. 1–12). Routledge.

Žuljević, M. F., & Buljan, I. (2022). Academic and non-academic predictors of academic performance in medical school: An exploratory cohort study. *BMC Medical Education*, *22*(1), 366. https://doi.org/10.1186/s12909-022-03436-1