Studies in Educational Evaluation 75 (2022) 101206

Contents lists available at ScienceDirect

Studies in Educational Evaluation

journal homepage: www.elsevier.com/locate/stueduc

Short form of revised two-factor study process questionnaire: Development, validation, and cross-validation in two European countries

Yusuf F. Zakariya [a,*], Barattucci Massimiliano [b]

[a] Department of Mathematical Sciences, University of Agder, Kristiansand, Norway
[b] Dipartimento di Scienze Umane e Sociali, Università degli studi di Bergamo, Italy

A R T I C L E   I N F O

Keywords:
Approaches to learning
Psychometric properties
Exploratory factor analysis
Confirmatory factor analysis
Measurement invariance

A B S T R A C T

The utility of students' approaches to learning (AL) has evolved from a simple act of academic monitoring to a systemic process of continuous improvement of teaching and learning in higher education. Research on the careful development of measures for AL is consequently crucial. The purpose of this research is to develop and validate a brief measure of AL that is theoretically sound with good psychometric properties. To achieve this purpose, we combined three rounds of studies in two European countries consisting of 253, 196, and 440 undergraduate students. Multiple tests of validity and reliability unveil a new measure of AL with two correlated dimensions of deep (four items) and surface (four items) approaches. The psychometric analyses of the new 8-item measure of AL provide promising results, both in validity and reliability demonstrating its possible use in academic contexts.

1. Introduction

There is growing interest in students' approaches to learning partly due to their relevance that transcends students' learning outcomes to the overall teaching activity. The students' dispositions toward learning and the processes adopted by students before or during a learning activity are linked to the quality of teaching and learning activities in higher education (Barattucci, 2017; Biggs, 2012; Zakariya, Nilsen et al., 2020). Evidence shows that students' approaches to learning can be used as a basis to (a) assess the effectiveness of instructional interventions; (b) compare teaching-learning experiences across classes; (c) identify students with learning difficulties and (d) examine the external validity of student-related research measures in higher education (Biggs et al., 2001; Lahdenperä et al., 2019; Zakariya et al., 2019). These multi-dimensional usages of students' approaches to learning have put the construct in the spotlight within the higher education research community. As such, we contend that a disciplined investigation into the measurement of students' approaches in higher education is necessary.

Admittedly, there are genuine and rigorous attempts to measure students' approaches to learning in the literature. Arguably, the revised two-factor study process questionnaire (R-SPQ-2 F) that was developed by Biggs et al. (2001) is one of the most popular measures of students' approaches to learning in higher education. The R-SPQ-2 F has received wide attention among researchers and it has been translated/validated in several languages across the world (Justicia et al., 2008; Önder & Besoluk, 2010; Xie, 2014; Zakariya, Bjørkestøl et al., 2020). However, R-SPQ-2 F has been equally criticised for its lack of construct validity (i.e., the inability of R-SPQ-2 F to reflect the constructs – deep and surface approaches – it is purported to measure) when validated in different languages (Merino & Kumar, 2013; Stes et al., 2013; Zakariya, 2019). This lack of construct validity of R-SPQ-2 F has prompted some researchers (e.g., Socha & Sigler, 2014; Stes et al., 2013) to delete some problematic items from the original measure while other researchers (Immekus & Imbrie, 2010; López-Aguado & Gutiérrez-Provecho, 2018) have called for a revision of the measure.

In a response to researchers that called for a revision of R-SPQ-2 F, the purpose of this research is to develop and validate a short form of R-SPQ-2 F (SF-R-SPQ-2 F) that is theoretically sound and has good psychometric properties. As such, we attempt to address the questions of which and how many of the items of the R-SPQ-2 F support the underlying constructs of the measure. This research is a combination of three studies combined to provide a coherent argument for the development, validation, and cross-validation of SF-R-SPQ-2 F in two European countries. We contend that the SF-R-SPQ-2 F will reduce respondents' burden of completing a long questionnaire and minimise the contested construct validity and reliability of the original measure. Further, it is

* Corresponding author.
E-mail address: yusuf.zakariya@uia.no (Y.F. Zakariya).

https://doi.org/10.1016/j.stueduc.2022.101206
Received 26 January 2022; Received in revised form 20 July 2022; Accepted 25 August 2022
Available online 30 August 2022

envisaged that students' scores on the SF-R-SPQ-2 F will be easier to interpret than the original 20-item R-SPQ-2 F.

2. Theoretical background

2.1. Conceptualising approaches to learning

Student approaches to learning (SAL) theory as pursued by Marton and Booth (1997); Marton and Säljö (2005) provided a well-grounded theoretical foundation for the conceptualisation and operationalisation of approaches to learning. In the SAL tradition, approaches to learning connote students' adopted *processes (strategies)*, when presented with learning tasks, which are offshoots of their *distinctive intentions (motives)* in engaging with the tasks (Biggs et al., 2001; Marton & Booth, 1997). One may remark that approaches to learning are a blend of students' motives and strategies that sit at the nexus of a dynamic relationship between the students, the presented task, and the context. This is because the students' motives that reflect in their adopted strategies while engaging in a task could be intrinsic or extrinsic (Ryan & Deci, 2020). As such, approaches to learning in the SAL tradition are context-dependent that vary from one context to another (Biggs et al., 2001). Some of the contextual factors that influence students' approaches to learning are students' perceptions of the presented task difficulty, their conceptions of learning, student-teacher relationship, instructional methods, assessment criteria, and classroom climate conditions (Biggs et al., 2001; Marton & Säljö, 2005).

A central tenet of the SAL tradition is the categorisation of approaches to learning into deep and surface approaches (Marton & Booth, 1997; Marton & Säljö, 2005). Despite the distinctive students' motives, several strategies, and non-trivial contextual influences, there is an accumulation of evidence that suggests that deep and surface approaches appropriately characterise students' approaches to learning (Biggs, 1993; Entwistle, 2005; Marton & Säljö, 2005). The deep approaches to learning characterise the adopted processes or strategies used by students that are intrinsically motivated in engaging with the presented tasks. The motives of such students are to: develop a conceptual understanding of the tasks; engage with the tasks out of personal interest and task-related enjoyment; understand ideas in the presented tasks out of curiosity for knowledge. As such, students with deep approaches to learning used several strategies such as meaning maximisation, relating old ideas with new ones, searching for underlying principles, and getting actively involved in learning tasks, to actualise their motives for the tasks (Biggs et al., 2001; Marton & Säljö, 2005). In contrast, the surface approaches to learning characterise the adopted processes or strategies used by students that are extrinsically motivated in engaging with the presented tasks. By extrinsic motivation, we refer to instrumental values (Ryan & Deci, 2020) attached to the task in which the students engage in an activity for its instrumental values such as passing the course, coping with the course requirements, and perceived usefulness for future careers. As such, students with surface approaches to learning used several strategies such as routine memorisation of facts and procedures, failure to relate old ideas with new ones, and striving to pass the course with minimal work, to actualise their motives for the tasks (Biggs et al., 2001; Marton & Säljö, 2005; Zakariya et al., 2021).

Admittedly, SAL theory has been criticised for its lack of clarity on the philosophical underpinnings and the conceptualisations of deep and surface approaches to learning. Researchers (e.g., Haggis, 2003; Howie & Bagnall, 2013) have argued that the theory was widely accepted prematurely with several applications in higher education whilst some crucial constructs such as deep and surface have not been fully developed. We have addressed some of these criticisms by clarifying the meanings of deep and surface approaches, their context-dependent, and their underlying mechanisms of motives and strategies. Further, the relationships of approaches to learning with students' performance and the measurement of the constructs are discussed in subsequent paragraphs. Thus, we contend that the critiques of the SAL theory pose no challenge to the development of ideas in the present study.

2.2. Approaches to learning and students' performance

The relationship between specific approaches to learning (deep and surface approaches) and students' performance is a bit controversial. It is controversial because there is a lack of coherence in the literature on the strength of such relationships. On the one hand, there is an accumulation of evidence that shows that deep approaches to learning have a substantial positive relationship with students' performance while the effect of surface approaches on performance is not substantial (Cano et al., 2018; Guo et al., 2017; Maciejewski & Merchant, 2016). On the other hand, there is an accumulation of evidence that shows that surface approaches to learning have a substantial negative relationship with students' performance while the effect of deep approaches on performance is not substantial (Diseth et al., 2009; Nguyen, 2016; Zakariya et al., 2021). Between these two extreme hands, there are some studies that either show a substantial positive relationship between deep approaches to learning and performance, and negative relationships between surface approaches to learning and performance or no substantial relationship between the two approaches and performance (Gijbels et al., 2005; Herrmann et al., 2016; Mundia & Metussin, 2019). A plausible explanation for these incoherent or rather conflicting reports is the variation in the learning contexts. As we argued in earlier sections of this article that approaches to learning are context-dependent, the research instruments used for measuring the constructs only measure what is prevalent in the context of the studies. It is expected that a teaching and learning context that encourages the adoption of deep approaches to learning should reflect a substantial positive relationship between deep approaches to learning and performance. In contrast, a teaching and learning context that encourages the adoption of surface approaches to learning should reflect a substantial negative relationship between surface approaches to learning and performance. This observation brings us to the question of what do approaches to learning questionnaires measure?

2.3. Measurement of approaches to learning

There are numerous attempts to operationalise and measure students' approaches to learning in the literature. Some of these attempts include the development and validation of approaches and study skills inventory for students (ASSIST), revised approaches to studying inventory (RASI), and R-SPQ-2 F (Biggs et al., 2001; Tait et al., 1998). Theoretical and practical issues on the latent constructs, number of construct dimensions, number of items, and cultural sensitivity of such measuring instruments have been debated (Diseth, 2001; Lake et al., 2017; Valadas et al., 2010; Zakariya, 2019). Yet, there are some obvious advantages of R-SPQ-2 F such as its arguably diverse acceptability around the world, its relatively small number of items, and its small number of latent constructs when compared with ASSIST and RASI. Historically, R-SPQ-2 F was developed by Biggs et al. (2001) through multiple studies that involved conceptualisation, operationalisation, re-conceptualisation, re-operationalisation, and validation of students' approaches to learning with a focus mostly on higher education (Biggs, 1987, 1993; Biggs et al., 2001).

For Biggs et al. (2001), approaches to learning can be preferred which capture individual differences in a teaching-learning activity, ongoing which captures students' strategies to handle specific tasks, and contextual which captures differences in the teaching-learning activity. Following the SAL tradition, Biggs et al. (2001) argued that approaches to learning are appropriately described as deep and surface approaches. As such, R-SPQ-2 F contains twenty items with ten items on the deep and surface approaches subscale, respectively. Further, the ten items on each of the two subscales are equally subdivided into motives and strategies. The subdivisions follow the SAL tradition that conceptualises approaches to learning as a combination of motives and strategies (Biggs et al., 2001; Marton & Säljö, 2005). Therefore, R-SPQ-2 F has four subcategories of five items each measuring deep motive, surface motive, deep strategy, and surface strategy, respectively. Table 1 presents some sample items in each of these subcategories of R-SPQ-2 F. The students are to acknowledge the level at which the item statements are true of them on a five-point Likert scale *never* or *only rarely, sometimes, half the time, frequently,* and *always* or *almost always* (Biggs et al., 2001).

The R-SPQ-2 F is well-received by researchers in different parts of the world such as in Africa (e.g., Matoti, 2014), Asia (e.g., Xie, 2014), Europe (e.g., Zakariya, 2019), the Middle East (e.g., Shaik et al., 2017), and the United States of America (e.g., Immekus & Imbrie, 2010). However, there is an accumulation of evidence that suggests that a two-factor model (deep and surface approaches) is sufficient to describe the underlying constructs of the R-SPQ-2 F (López-Aguado & Gutiérrez-Provecho, 2018; Merino & Kumar, 2013; Önder & Besoluk, 2010; Xie, 2014; Zakariya, Bjørkestøl et al., 2020). This is contrary to the four-factor hypothesised model of R-SPQ-2 F by Biggs et al. (2001). Meanwhile, the two-factor model of R-SPQ-2 F also came at a cost of deleting some items from the original 20-item R-SPQ-2 F. Zakariya (2019) reports a systematic review of studies on the construct validity of R-SPQ-2 F including the number of items deleted in each of the reviewed studies to achieve an appropriate fit of the model. To address the intercultural disparity in what R-SPQ-2 F measures, we investigate items of the R-SPQ-2 F that support its underlying constructs using CFA.

3. Methodology

This article reports on three independent studies that are strategically tailored toward addressing the research question as raised in the introduction section. The three studies follow exploratory factor analysis, confirmatory factor analysis, re-validation, and cross-cultural validation analytic methods. We suppose that it is prudent to present the specific aims, methods, results, and discussion while describing each of the studies.

3.1. Study one

Research aim.

3.1.1. Research aim

The purpose of the first study on the development of SF-R-SPQ-2 F was to explore the factor structure of the 20-item R-SPQ-2 F. This explorative process avails us an opportunity to examine the pattern of factor loadings and decide on how reflective the 20 items measure the constructs that they are purported to measure.

3.1.2. Research method

We prepared both electronic and paper versions of R-SPQ-2 F and administered them to first-year undergraduate students in two Norwegian universities. A relatively high number of 253 engineering students (72 females) gave consent to take part in the study and returned the completed questionnaires. Only engineering students are targeted in this study because the researchers delimited the research scope to a convenient and accessible sample of STEM students. The generated data were examined for missingness, outliers, kurtosis, and skewness. The preliminary analysis showed that the data contained excess kurtosis and skewness (i.e., absolute values of both indices are greater than one for some items) but no outlier and missing values. As such, we used a polychoric correlation matrix instead of Pearson's correlation matrix for subsequent analysis. The choice of the correlation rests on the fact that the polychoric matrix is more robust to defects in both kurtosis and skewness than the Pearson matrix (Holgado–Tello et al., 2010).

As a first step in the development of the SF-R-SPQ-2 F, we used exploratory factor analysis to examine the factor structures of the 20-item R-SPQ-2 F. For factor extraction, we used minimum rank factor analysis due to its proven effectiveness in yielding optimal communalities of items (Shapiro & ten Berge, 2002). To determine the number of factors to retain, we used parallel analysis against the common scree plot and Kaiser's criterion of eigenvalues greater than one. This is because both the scree plot and Kaiser's criterion have been criticised and shown to be less efficient than the parallel analysis (Timmerman & Lorenzo-Seva, 2011). Thereafter, we used promin (Lorenzo-Seva, 1999) to rotate the factor loading matrix. Given that the extracted factors are meant to be correlated, we argue that promin (oblique rotation) is more appropriate than an orthogonal rotation. We performed the exploratory factor analysis using the FACTOR program version 10.8.04 (Lorenzo-Seva & Ferrando, 2013).

3.1.3. Results and discussion

The polychoric correlation matrix for the 20-item R-SPQ-2 F is presented in Appendix 1. The availability of the polychoric correlation matrix is crucial for replication and independent verification of the subsequent findings. The multicollinearity and the adequacy of the polychoric correlation matrix results show that the sample is sufficient for exploratory factor analysis with a significant Bartlett statistic ($N = 253$, $df = 190$) = 1037.7, $p = 0.00001$, a fair Kaiser-Meyer-Olkin (KMO) test = 0.79155, and the matrix determinant of greater than 0.00001 (Field, 2018). The result of the parallel analysis based on minimum rank factor analysis shows that two latent factors are sufficient to describe the correlations between the 20 items of the R-SPQ-2 F. The pattern of rotated factor loadings of the two-factor R-SPQ-2 F model is presented in Table 2 (factor loadings with an absolute value of less than 0.30 are suppressed).

The results reveal some interesting findings: first, Table 2 shows that Item 01 and Item 17 have factor loadings of less than.30 on both extracted and rotated factors of R-SPQ-2 F. Second, Table 2 shows that Item 03, Item 07, Item 08, and Item 11 have substantial cross-loadings on both extracted and rotated factors of R-SPQ-2 F. These results challenge the strength and appropriateness of these six items with the factors they are purported to measure; suggesting that Item 01 and Item 07 reflect a weak strength of the factors that they are purported to measure. Further, the substantial cross-loadings of Item 03, Item 07, Item 08, and Item 11 show that these items are not reflective of the factors that they purported to measure. As such, these six items are excluded from further analysis of the development and validation of SF-R-SPQ-2 F. Except for Item 11, the exclusion of these items corroborates previous research that

Table 1
Subcategories of its items and sample items of the R-SPQ-2F.

	Motive	Strategy
Deep approaches	Item 05: *"I feel that virtually any topic can be highly interesting once I get into it"*	Item 06: *"I find most new topics interesting and often spend extra time trying to obtain more information about them"*
Deep approaches	Item 13: *"I work hard at my studies because I find the material interesting"*	Item 10: *"I test myself on important topics until I understand them completely"*
Surface approaches	Item 15: *"I find it is not helpful to study topics in depth. It confuses and wastes time when all you need is a passing acquaintance with topics"*	Item 12: *"I generally restrict my study to what is specifically set as I think it is unnecessary to do anything extra"*
Surface approaches	Item 19: *"I see no point in learning material which is not likely to be in the examination"*	Item 16: *"I believe that lecturers shouldn't expect students to spend significant amounts of time studying material everyone knows won't be examined"*

Note. Table 1 is adapted from the 20-item R-SPQ-2 F that is provided by Biggs et al. (2001, p. 149).

Table 2
Rotated pattern matrix of the 20-item R-SPQ-2F (|factor loadings| ≤ 0.3 are excluded).

Item	Factor one	Factor two
01	–	–
02	–	0.412
03	0.336	-0.355
04	0.416	–
05	–	0.573
06	–	0.737
07	0.368	-0.466
08	0.475	0.392
09	–	0.580
10	–	0.616
11	0.666	0.324
12	0.436	–
13	–	0.786
14	–	0.611
15	0.518	–
16	0.498	–
17	–	–
18	–	0.487
19	0.607	–
20	0.438	–

deleted or recommended the deletion of these items to achieve the construct validity of R-SPQ-2 F (Immekus & Imbrie, 2010; Stes et al., 2013). Another crucial finding that is revealed in Table 2 is the clean separation of the remaining 14 items of R-SPQ-2 F into Factor 1 and Factor 2 which are consistent with the surface and deep approaches as initially hypothesised by Biggs et al. (2001), respectively. The finding shows that six items (Item 04, Item 12, Item 15, Item 16, Item 19, and Item 20) measure surface approaches to learning, and eight items (Item 02, Item 05, Item 06, Item 09, Item 10, Item 13, Item 14, Item 18) measure deep approaches to learning. The reduction of items of R-SPQ-2 F from 20 to 14 items set the stage for the second strand of the present research.

3.2. Study two

3.2.1. Research aim

As a follow-up to study one, the purpose of study two was to confirm the two-factor structure of the 14-item R-SPQ-2 F (eight items on deep and six items on surface approaches to learning) in an independent sample from that of study one. This confirmatory research avails the opportunity to examine the construct validity of the 14-item R-SPQ-2 F (Fig. 1) and serves as build-up evidence for the development of the SF-R-SPQ-2 F.

Fig. 1 shows the hypothesised model of the 14-item R-SPQ-2 F with deep and surface approaches to learning in big oval shapes and their items in boxes at the end of respective directed arrows. The singled-headed directed arrows from the big oval shapes to the boxes show that the items are reflective of the latent factors: deep and surface. The λ_i's are factor loadings and show the strength of the reflective relationship between the items and the corresponding factors that they are purported to measure. The small oval shapes with short arrows pointing toward the items are disturbances of the items. That is the variance of the items that are not explained by the latent factors. The standardised correlation (r) between deep and surface approaches to learning is represented by the double-headed arrow between the constructs. This standardised correlation is expected to be negative as respondents with high scores on deep approach items are expected to have low scores on surface approach items of the questionnaire and vice-versa.

3.2.2. Research method

As in study one, we prepared both the electronic and paper versions of R-SPQ-2 F and administered them to first-year undergraduate students in a Norwegian university. 196 undergraduate engineering students (34 females) with an average age range between 21 and 25 years gave consent to take part in the research and returned the completed questionnaires. Engineering students are targeted in study two, as well, because we delimited the research scope to a convenient and accessible sample of STEM students. We examined the generated data for missing values, outliers, kurtosis, and skewness and found that the data

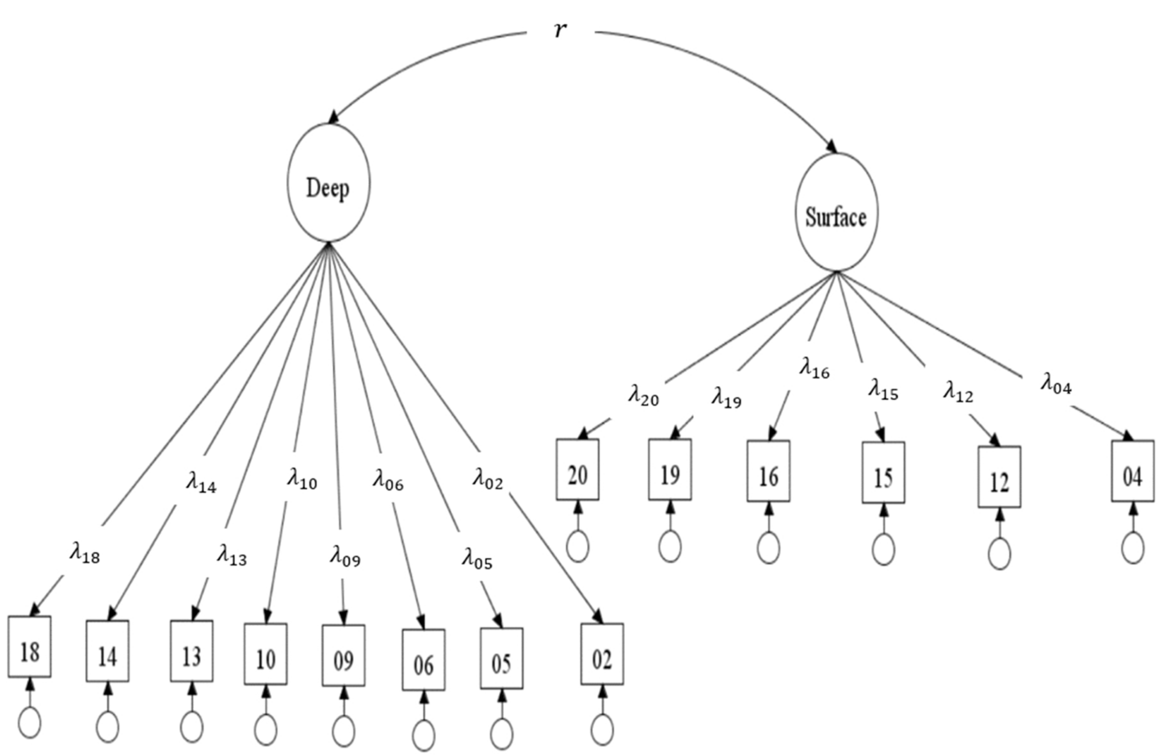

Fig. 1. Hypothesised model of the 14-item R-SPQ-2F.

contained excess kurtosis, skewness (i.e., absolute values of both indices are greater than one for some items), and contained neither missing values nor outliers. As such, we used a polychoric correlation matrix (Appendix 2) instead of a Pearson correlation matrix for subsequent analysis. The polychoric correlation matrix is available in Appendix 2 for possible replication and independent verification of the subsequent findings.

We analysed the generated data using confirmatory factor analysis (CFA) with a weighted least square mean and variance adjusted (WLSMV) estimator. We used WLSMV instead of the competing maximum likelihood estimator because of some defects in the normality distribution of the generated data and its categorical level of measurement. The logic of the CFA involves comparing the sample variance-covariance matrix with the predicted model-implied matrix for consistency between the hypothesised model (Fig. 1) and the generated data. We used a combination of criteria to judge the global and local fit of the generated data with the hypothesised model. A model exhibits an excellent global fit with the generated data if the Tucker-Lewis index (TLI) and the comparative fit index (CFI) are greater than or equal to 0.95 (Hu & Bentler, 1999), standardised root mean square residual (SRMR) is less than 0.6, root mean square error of approximation (RMSEA) is less than.08, and the ratio of the chi-square value to the degree of freedom is less than 3 (Brown, 2015; MacCallum et al., 1996). For the local fit of the model with the generated data, we used significant factor loadings (p-value of less than 0.05) to judge an excellent local fit. We performed all the analyses in Mplus 8.3 software.

3.2.3. Results and discussion

The results of the CFA are presented in Table 3. The table shows the goodness of fit statistics for three measurement models which are labelled Model 1, Model 2, and Model 3. Model 1 is the measurement model of 14-item R-SPQ-2 F as presented in Fig. 1. The consistency of Fig. 1 with the generated data was evaluated and the results are presented under the heading Model 1. Model 2 is an improvement on Model 1 while Model 3 is an improvement on Model 2.

The presented results in Table 3 (Model 1) show that the 14-item R-SPQ-2 F measurement model demonstrates a poor global fit of the generated data. That is, there is a lack of consistency between the measurement model and the generated data. Admittedly, the ratio of chi-square to the degree of freedom is less than 3, and the RMSEA and the SRMR values are within the recommended ranges. However, the CFI and the TLI values are below the minimum value of.95 for an acceptable model fit. As such, we fail to accept the 14-item measurement model of the R-SPQ-2 F. A further examination of the local fit statistics shows that Item 04 has a non-significant factor loading, Item 20 has a weak factor loading (0.203), and Item 14 and item 18 have weak R-squared values of.200 and.207, respectively. We removed these four items in the second round of the analysis. The remaining ten items are distributed such that there are six items on deep approaches to learning (Item 02, Item 05, Item 06, Item 09, Item 10, and Item 13) and four items on the surface approaches to learning (Item 12, Item 15, Item 16, and Item 19). The results of the CFA of the 10-item R-SPQ-2 F are presented in Table 3 (Model 2).

The presented results in Table 3 (Model 2) show that the 10-item R-SPQ-2 F measurement model demonstrates an excellent global fit of the generated data. That is, there is consistency between the measurement model and the generated data. We deduce the excellent global fit of the model from the fact that all the goodness of fit statistics are within the recommended ranges of an excellent fit (Hu & Bentler, 1999). More so, the model demonstrates an excellent local fit of the generated data because all the factor loadings are significant. Partly due to the quest to develop a short form of the R-SPQ-2 F with equal subscale items and partly due to evenness in the theoretical distribution of the scale items, we argue for the removal of two items on the deep approaches to learning subscale. These two items are Item 02 (I find that I have to do enough work on a topic so that I can form my own conclusions before I am satisfied) and item 09 (I find that studying academic topics can at times be as exciting as a good novel or movie). Theoretically, the six items of the deep approaches to learning subscale are such that there are three items on deep-motive (Item 05, Item 09, and Item 13) and three items on deep-strategy (Item 02, Item 06, and Item 10). We removed Item 09 from the deep-motive subcategory because it is relatively long and contains words like “novel” and “movie” which could distract the respondents who are not fond of reading novels or watching movies (Zakariya, Bjørkestøl et al., 2020a). Further, we removed Item 02 from the deep-strategy subcategory because has the least factor loading. To buttress the conceptual justification for removing Item 02 and Item 09 from the new scale we investigate the fit of the 8-item R-SPQ-2 F with the generated data. The results of the CFA of the 8-item R-SPQ-2 F are presented in Table 3 (Model 3).

The presented results in Table 3 (Model 3) confirm an excellent global model fit of the 8-item R-SPQ-2 F (henceforth, SF-R-SPQ-2 F) model with the generated data. That is, the SF-R-SPQ-2 F is consistent with generated data. The goodness of fit statistics that are all within the recommended ranges suggests excellent model fit. The SF-R-SPQ-2 F also demonstrates an excellent local fit of the generated data. Fig. 2 presents some local fit statistics of the SF-R-SPQ-2 F. The local and global fit statistics provide statistical corroborative evidence to support the researchers’ conceptual arguments for removing Item 02 and Item 09 from the new scale.

Fig. 2 shows the standardised factor loadings of the SF-R-SPQ-2 F with the least factor loading of.471 (Item 10) and the highest factor

Table 3
The goodness of fit statistics of the three measurement models of the 14-item R-SPQ-2F.

Global fit statistics	Model 1	Model 2	Model 3
Chi-square			
Estimate (χ^2)	140.107	58.722	33.659
Degrees of freedom (*df*)	76	34	19
χ^2/df	1.844	1.727	
SRMR	0.057	0.043	0.038
RMSEA			
Estimate	0.066	0.061	0.063
90% confidence interval	[0.048,0.082]	[0.033,0.087]	[0.025,0.097]
Probability RMSEA < =0.05	0.068	0.232	0.249
CFI/TLI			
CFI	0.926	0.968	0.976
TLI	0.911	0.958	0.964

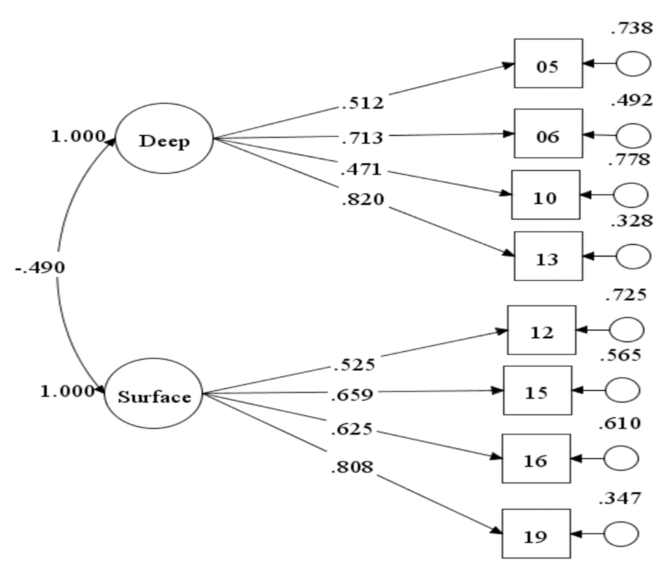

Fig. 2. Evaluated measurement model of the 8-item R-SPQ-2F.

loading of.820 (Item 13). These factor loadings show that each item of the SF-R-SPQ-2 F has a strong relationship with the factor that it is purported to measure. Fig. 2 also shows the standardised disturbance of each item of the SF-R-SPQ-2 F. For instance, Item 13 has the least standardised disturbance of.328. That is, 32.8% of its variance is due to disturbance of the item. On the other hand, Item 10 has the greatest standardised disturbance of.778. The standardised correlation between the deep and surface approaches to learning constructs is negative and significant ($r = -.490$) and their standardised variance is fixed to 1 for model identification. This negative correlation confirms our expectation and provides evidence of discriminant validity of the SF-R-SPQ-2 F between the two constructs that it is purported to measure. For a reconfirmation of the factor structure of SF-R-SPQ-2 F in independent samples, reliability, and measurement invariance, we proceed to another round of research.

3.3. Study three

3.3.1. Research aim

The purpose of study three was to reconfirm the factor structure of the SF-R-SPQ-2 F in independent samples and examine more psychometric properties such as reliability and measurement invariance. The measurement invariance involves a statistical examination of whether a measuring instrument measures what it purported to measure across multiple groups (Putnick & Bornstein, 2016). It is a criterion for judging the viability of using a research instrument for cross-cultural mean comparisons (Zakariya, 2021).

3.3.2. Research methods

We prepared both the electronic and paper versions of SF-R-SPQ-2 F and administered them to undergraduate students in two countries: Norway and Italy. The Norwegian sample was a homogeneous sample of 190 first-year undergraduate engineering students (157 males) with an average age range between 21 and 25 years who gave consent to take part in the study and returned the completed questionnaires. The Italian sample, on the other hand, was a heterogeneous sample of 250 university students (133 males, average age = 23.51 years) following different courses in sciences, engineering, social sciences, and humanities at different years of study. We examined the generated data for missing values, outliers, kurtosis, and skewness and found that the data did not contain excess kurtosis, skewness (i.e., absolute values of both indices are greater than one for some items), and contained neither missing values nor outliers. However, we used a polychoric correlation matrix instead of Pearson's correlation matrix because of the categorical level of measurement of the generated data. The polychoric correlation matrices for both the Norwegian and the Italian samples are presented in Appendix 3 and Appendix 4, respectively, for possible replication and independent verification of the subsequent findings.

For the re-validation and the cross-cultural validation of the factor structure of the SF-R-SPQ-2 F, we used multiple group CFA with the WLSMV estimator to analyse the generated datasets. Following the latent variable approach to modelling, we used coefficient omega (Dunn et al., 2014; McDonald, 2011) to compute the reliability coefficient of each dimension of the SF-R-SPQ-2 F. The latent variable for computing reliability outperforms the popular Cronbach alpha coefficient in estimating reliability coefficients, and it is more robust to violations of assumptions such as tau-equivalence and normal distribution (Dunn et al., 2014; Trizano-Hermosilla & Alvarado, 2016). We investigated the viability of using scores on SF-R-SPQ-2 F for cross-cultural mean comparison by examining the scale measurement invariance. The measurement invariance is at three levels depending on the restrictions placed on the parameters of the measuring instruments. The first level is the configural measurement invariance where the factor structure of a model is investigated across multiple groups. The second level is the metric measurement invariance where in addition to the factor structure the factor loadings are constrained to be equal across multiple groups.

The third level is the scalar measurement invariance where in addition to both the factor structure and equality of factor loadings, the item intercepts/thresholds are constrained to be equal across multiple groups (Brown, 2015; Putnick & Bornstein, 2016). Some researchers (e.g., Byrne, 2012) suggest that for a trustworthy cross-cultural mean comparison, scalar measurement invariance is required of a measuring instrument. However, others (e.g., Zakariya, 2021; Zakariya, Bjørkestøl, & Nilsen, 2020) have argued for and applied approximate measurement invariance for a cross-cultural mean comparison when the scalar invariance condition is violated. We judged the consistency of the models with generated data using the criteria set in study two.

3.3.3. Results and discussion

The first set of results concerns the re-confirmation of the factor structure of the SF-R-SPQ-2 F in both the Norwegian and the Italian samples. The Norwegian analysis was to examine the construct validity of the SF-R-SPQ-2 F in an independent sample from the one used to develop the measure. The Italian analysis, on the other hand, was to examine the construct validity of the SF-R-SPQ-2 F in a different cultural context from the one used to develop the measure. This Italian analysis was envisaged to address the question of whether SF-R-SPQ-2 F measure what it is purported to measure when it is translated to the Italian language. Table 4 presents the goodness of fit statistics for the evaluated SF-R-SPQ-2 F models in both the Norwegian and the Italian samples.

The presented results in Table 4 (Norwegian) show an excellent global fit of the SF-R-SPQ-2 F model with the Norwegian data. This is because the goodness of fit statistics are within the recommended ranges of an excellent fit. This finding confirms the factor structure of the SF-R-SPQ-2 F in an independent sample from the one that was used to develop the measuring instrument. More so, the presented results in Table 4 (Italian) show a very good global fit (Marsh et al., 2004) of the SF-R-SPQ-2 F model with the Italian data. The global fit statistics suggest that the fit of the model with the Italian data is not excellent according to the criteria for judging an excellent model. This is because the ratio of the chi-square value to the degree of freedom is slightly greater than 3. However, the global fit of the model is very good since the SRMR, CFI, and TLI statistics are within their recommended ranges of excellent global fit. More so, RMSEA is appropriate since its 90% confidence interval contains.80 (Marsh et al., 2004). Therefore, there is consistency between the hypothesised model of SF-R-SPQ-2 F with the generated Italian data which provides evidence of cross-cultural construct validity for SF-R-SPQ-2 F. This finding shows that SF-R-SPQ-2 F measures what it purported to measure when it is translated into the Italian language.

It is important to remark that despite the heterogeneous nature of the Italian sample, SF-R-SPQ-2 F managed to exhibit construct validity. This interesting observation shows that SF-R-SPQ-2 F measures what it is purported to measure regardless of the course of study and the year of study of the participating undergraduate students. Fig. 3 presents the local fit statistics of both the Norwegian and Italian models of SF-R-SPQ-2 F.

Table 4
Global fit statistics of the Norwegian and Italian validations of SF-R-SPQ-2F.

Global fit statistics	Norwegian	Italian
Chi-square		
Estimate (χ^2)	32.082	62.711
Degrees of freedom (*df*)	19	19
χ^2/df	1.689	3.300
SRMR	0.039	0.040
RMSEA		
Estimate	0.060	0.096
90% confidence interval	[0.019,0.095]	[0.070,0.123]
Probability RMSEA < =0.05	0.292	0.002
CFI/TLI		
CFI	0.978	0.966
TLI	0.968	0.950

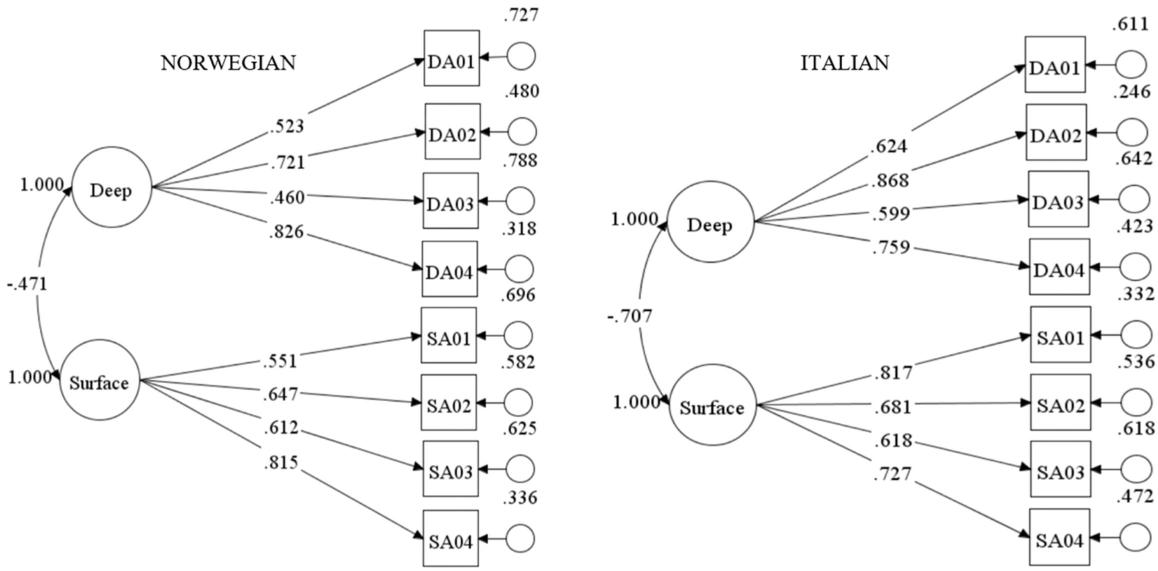

Fig. 3. Local fit statistics of SF-R-SPQ-2F models in both Norwegian and Italian samples.

The presented statistics in Fig. 3 show that the SF-R-SPQ-2 F model exhibits a substantial local fit of the generated data in both Norway and Italy. These local fit statistics corroborated the global fit statistics in providing evidence of the construct validity of SF-R-SPQ-2 F in both Norway and Italy. For unidimensionality and internal consistency of the two dimensions of SF-R-SPQ-2 F, we evaluated a four-item one-factor model for each dimension of SF-R-SPQ-2 F and computed its coefficient omega in both the Norwegian and Italian samples. Table 5 presents the local fit statistics and reliability coefficients of both dimensions for the Norwegian and Italian models of SF-R-SPQ-2 F.

The presented results in Table 5 coupled with an excellent global model fit of each dimension of the SF-R-SPQ-2 F confirm the unidimensionality of both dimensions in the Norwegian and Italian samples. The factor loadings and the residuals were used to compute omega coefficients for each dimension of SF-R-SPQ-2 F using the formula proposed by McDonald (2011). Table 5 shows that the reliability coefficients of each dimension of SF-R-SPQ-2 F are high (McDonald, 2011) and comparable in both the Norwegian and Italian samples. As such, one can infer that the consistency with which SF-R-SPQ-2 F measures what it purported to measure is high. Finally, the researchers investigate the measurement invariance of SF-R-SPQ-2 F, and the results are presented in Table 6.

The presented results in Table 6 show that the SF-R-SPQ-2 F model demonstrates an excellent global fit with the generated data for the configural and metric models. The results show that the SF-R-SPQ-2 F satisfies configural invariance, that is, the pattern of the factor structure of the SF-R-SPQ-2 F is preserved in both the Norwegian and Italian samples. Even though, the metric model demonstrates an excellent global fit with generated data we cannot conclude that the SF-R-SPQ-2 F model satisfies metric invariance. This is because the model comparison between the configural and the metric models is significant $\left(\Delta\chi^2 = 28.271, \Delta df = 6, p = .001\right)$ and the $\Delta CFI = .013$ exceeds the recommended.10 even though the $\Delta RMSEA = .01$ is within the recommended.015 (Chen, 2007). That is, the SF-R-SPQ-2 F model does not preserve its excellent model fit of the generated data when the factor loadings are constrained to be equal in both the Norwegian and Italian samples. We did not go further to evaluate the model scalar invariance because of the failure of the model to exhibit metric invariance. By implication, we advise caution when using scores of the SF-R-SPQ-2 F for cross-cultural mean comparisons of approaches to learning.

Table 5
Local fit statistics and reliability coefficients of SF-R-SPQ-2F dimensions.

	Loading		Residual		Omega coefficient	
	Norway	Italy	Norway	Italy	Norway	Italy
Deep						
DA01	0.572	0.688	0.673	0.527	0.739	0.813
DA02	0.755	0.759	0.430	0.424		
DA03	0.495	0.606	0.755	0.633		
DA04	0.736	0.823	0.458	0.323		
Surface						
SA01	0.557	0.747	0.690	0.442	0.755	0.807
SA02	0.643	0.694	0.587	0.518		
SA03	0.652	0.696	0.575	0.516		
SA04	0.779	0.722	0.393	0.479		

4. Conclusion

In this study, we attempt to develop, validate, and cross-validate a short form of an instrument that measures students' approaches to learning. The new measuring instrument (SF-R-SPQ-2 F) of approaches to learning contains eight items of the original 20-item R-SPQ-2 F with good psychometry properties. The eight items of SF-R-SPQ-2 F are distributed such that there are four items on deep approaches to learning and four items on surface approaches. These eight items match their respective theoretical factor structure (two items each on deep motive, deep strategy, surface motive, and surface strategy) as hypothesised by Biggs et al. (2001) in the original 20-item instrument. The respondents are to rate their agreement to the SF-R-SPQ-2 F items on a five-point Likert scale: never or only rarely (1), sometimes (2), half the time (3), frequently (4), and always or almost always (5). The full SF-R-SPQ-2 F is provided in Appendix 5. The respondents' scores on the items DA01 – DA04 can be added and interpreted as respondents' scores on deep approaches to learning. In contrast, the respondents' scores on the items SA01 – SA04 can be added and interpreted as respondents' scores on surface approaches to learning. Where possible, it is highly recommended to use the latent variable approach in computing the factor scores for a more precise estimation of the factor scores.

Evidence from this study suggests that SF-R-SPQ-2 F has robust construct validity which is transferable from a Scandinavian country to another European country. More so, evidence shows that SF-R-SPQ-2 F has a high-reliability coefficient. To the best of our knowledge, this research is the first attempt in the literature that develops a brief measure of approaches to learning with such psychometry properties. This novel contribution to literature will open research opportunities on the

Table 6
Multiple group configural and metric invariance of the SF-R-SPQ-2F model.

Model	χ^2	*df*	CFI	TLI	RMSEA	SRMR	Model comparison	$\Delta\chi^2$	Δdf	p
Configural (1)	91.157	38	0.973	0.959	0.080	0.040	–	–	–	–
Metric (2)	121.931	44	0.960	0.949	0.090	0.045	(2) vs. (1)	28.271	6	0.001

Note. $\Delta\chi^2$ means change in chi-square values with Satorra-Bentler correction, Δdf means change the in degree of freedom, and "vs." means versus. All the chi-square values are significant at $p < 0.05$.

construct, reduce respondents' burden in completing a long questionnaire, and ease scores interpretation for researchers. Despite the good psychometric properties of SF-R-SPQ-2 F, there are some reservations and limitations that are worth mentioning. Admittedly, SF-R-SPQ-2 F has construct validity. However, it fails to exhibit scalar measurement invariance. The source for the lack of scalar invariance can be ascribed to the heterogeneous sample used for the cross-validation of the instrument. Many factors such as the differences in students' study programmes, languages, and culture could serve as confounders that are responsible for the lack of scalar invariance. Thus, caution should be observed while using scores on the instrument for cross-cultural mean comparison. Also, we could neither investigate the predictive validity of SF-R-SPQ-2 F nor its relationship with other constructs such as performance. These limitations offer opportunities for further research on the instrument. Further, our exclusive use of engineering students in the development and validation of SF-R-SPQ-2 F can be problematic for the validity and reliability of the instrument. We detected this limitation early and tried to rectify it by using a heterogeneous sample in the Italian study. We acknowledge that our rectification of this limitation may not be enough and as such recommend further study with heterogeneous samples.

In conclusion, the interest of universities in measuring the strategies and approaches to learning of their students has gone from a simple act of academic monitoring to a systemic process of continuous improvement of teaching and the learning environment. Research concerning the development and implementation of tools for measuring learning is consequently crucial and with broad and understandable practical applications. This study aimed to develop a new tool that would overcome past problems such as weak construct validity and ambiguity in the interpretations of scale scores and provide indications for future adjustments. The psychometric analyses of the proposed 8-item measure of approaches to learning provide promising results, both in validity and reliability demonstrating its possible use in academic contexts.

Funding

No funding to declare.

CRediT authorship contribution statement

Yusuf F. Zakariya: Conceptualization, Methodology, Formal analysis, Software, Data curation, Investigation, Writing – original draft, Writing – review & editing. **Barattucci Massimiliano:** Data curation, Writing – review & editing.

Competing interests

No competing interest.

Data Availability

The data used for present research are available upon request from the corresponding author. Meanwhile, the appendix contains important matrices for possible independent replications of the results.

Acknowledgement

The corresponding author acknowledges the support received from the University of Agder library for funding the article processing charge.

Appendix A. Supplementary material

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.stueduc.2022.101206.

References

Barattucci, M. (2017). Approach to study as an indicator of the quality of teaching and of learning environment: The contribution of John Biggs. *Journal of e-Learning and Knowledge Society, 13*(2), 77–88. https://doi.org/10.20368/1971-8829/1311
Biggs, J. B. (1993). What do inventories of students' learning processes really measure? A theoretical review and clarification. *British Educational Research Journal, 63*, 3–19. https://doi.org/10.1111/j.2044-8279.1993.tb01038.x
Biggs, J. B. (2012). What the student does: Teaching for enhanced learning. *Higher Education Research & Development, 31*(1), 39–55. https://doi.org/10.1080/07294360.2012.642839
Biggs, J. B., Kember, D., & Leung, D. Y. P. (2001). The revised two-factor study process questionnaire: R-SPQ-2F. *British Journal of Educational Psychology, 71*(1), 133–149. https://doi.org/10.1348/000709901158433
Biggs, J. B. (1987). Student approaches to learning and studying. Australian Council for Educational Research.
Brown, T. A. (2015). *Confirmatory factor analysis for applied research*. The Guilford Press.
Byrne, B. M. (2012). Structural equation modeling with Mplus: Basic concepts, applications, and programming. Routledge, Taylor & Francis Group.
Cano, F., Martin, A. J., Ginns, P., & Berbén, A. B. G. (2018). Students' self-worth protection and approaches to learning in higher education: Predictors and consequences. *Higher Education, 76*(1), 163–181. https://doi.org/10.1007/s10734-017-0215-0
Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal, 14*(3), 464–504. https://doi.org/10.1080/10705510701301834
Diseth, Å. (2001). Validation of a Norwegian version of the approaches and study skills inventory for students (ASSIST): Application of structural equation modelling. *Scandinavian Journal of Educational Research, 45*(4), 381–394.
Diseth, Å., Pallesen, S., Brunborg, G. S., & Larsen, S. (2009). Academic achievement among first semester undergraduate psychology students: The role of course experience, effort, motives and learning strategies. *Higher Education, 59*(3), 335–352. https://doi.org/10.1007/s10734-009-9251-8
Dunn, T. J., Baguley, T., & Brunsden, V. (2014). From alpha to omega: A practical solution to the pervasive problem of internal consistency estimation. *British Journal of Psychology, 105*(3), 399–412. https://doi.org/10.1111/bjop.12046
Entwistle, N. (2005). Contrasting perspectives on learning. In F. Marton, D. Hounsell, & N. Entwistle (Eds.), *The experience of learning: Implications for teaching and studying in higher education* (3rd (Internet) ed., pp. 106–125). The University of Edinburgh, Centre for Teaching, Learning and Assessment.
Field, A. (2018). *Discovering statistics using IBM SPSS statistics* (5th ed.). SAGE Publication Ltd.
Gijbels, D., Van de Watering, G., Dochy, F., & Van den Bossche, P. (2005). The relationship between students' approaches to learning and the assessment of learning outcomes. *European Journal of Psychology of Education, 20*(4), 327–341. https://doi.org/10.1007/BF03173560
Guo, J., Yang, L., & Shi, Q. (2017). Effects of perceptions of the learning environment and approaches to learning on Chinese undergraduates' learning. *Studies in Educational Evaluation, 55*, 125–134. https://doi.org/10.1016/j.stueduc.2017.09.002
Haggis, T. (2003). Constructing images of ourselves? A critical investigation into 'approaches to learning' research in higher education. *British Educational Research Journal, 29*(1), 89–104. https://doi.org/10.1080/0141192032000057401
Herrmann, K. J., Bager-Elsborg, A., & McCune, V. (2016). Investigating the relationships between approaches to learning, learner identities and academic achievement in higher education. *Higher Education, 74*(3), 385–400. https://doi.org/10.1007/s10734-016-9999-6

Holgado-Tello, F. P., Chacón-Moscoso, S., Barbero-García, I., & Vila-Abad, E. (2010). Polychoric versus Pearson correlations in exploratory and confirmatory factor analysis of ordinal variables. *Quality & Quantity, 44*(1), 153–166. https://doi.org/10.1007/s11135-008-9190-y

Howie, P., & Bagnall, R. (2013). A critique of the deep and surface approaches to learning model. *Teaching in Higher Education, 18*(4), 389–400. https://doi.org/10.1080/13562517.2012.733689

Hu, L. T., & 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

Immekus, J. C., & Imbrie, P. K. (2010). A test and cross-validation of the revised two-factor study process questionnaire factor structure among western university students. *Educational and Psychological Measurement, 70*(3), 495–510. https://doi.org/10.1177/0013164409355685

Justicia, F., Pichardo, M. C., Cano, F., Berbén, A. B. G., & De la Fuente, J. (2008). The revised two-factor study process questionnaire (R-SPQ-2F): Exploratory and confirmatory factor analyses at item level. *European Journal of Psychology of Education, 23*(3), 355–372.

Lahdenperä, J., Postareff, L., & Rämö, J. (2019). Supporting quality of learning in university mathematics: A comparison of two instructional Designs. *International Journal of Research in Undergraduate Mathematics Education, 5*(1), 75–96. https://doi.org/10.1007/s40753-018-0080-y

Lake, W. W., Boyd, W. E., & Boyd, W. (2017). Understanding how students study: The genealogy and conceptual basis of a widely used pedagogical research tool, Biggs' study process questionnaire. *International Education Studies, 10*(5), 100–108. https://doi.org/10.5539/ies.v10n5p100

López-Aguado, M., & Gutiérrez-Provecho, L. (2018). Checking the underlying structure of R-SPQ-2F using covariance structure analysis / Comprobación de la estructura subyacente del R-SPQ-2F mediante análisis de estructura de covarianza. *Cultura York Educación/Culture and Education, 30*(1), 105–141. https://doi.org/10.1080/11356405.2017.1416787

Lorenzo-Seva, U. (1999). Promin: a method for oblique factor rotation. *Multivariate Behavioral Research, 34*(3), 347–365. https://doi.org/10.1207/s15327906mbr3403_3

Lorenzo-Seva, U., & Ferrando, P. J. (2013). FACTOR 9.2: A comprehensive program for fitting exploratory and semiconfirmatory factor analysis and IRT models. *Applied Psychological Measurement, 37*(6), 497–498.

MacCallum, R. C., Browne, M. W., & Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods, 1*(2), 130–149. https://doi.org/10.1037/1082-989X.1.2.130

Maciejewski, W., & Merchant, S. (2016). Mathematical tasks, study approaches, and course grades in undergraduate mathematics: A year-by-year analysis. *International Journal of Mathematical Education in Science and Technology, 47*(3), 373–387. https://doi.org/10.1080/0020739x.2015.1072881

Marsh, H. W., Hau, K.-T., & Wen, Z. (2004). In search of golden rules: Comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing Hu and Bentler's (1999) findings. *Structural Equation Modeling: A Multidisciplinary Journal, 11*(3), 320–341. https://doi.org/10.1207/s15328007sem1103_2

Marton, F., & Booth, S. A. (1997). *Learning and awareness*. Lawrence Erlbaum.

Marton, F., & Säljö, R. (2005). Approaches to learning. In F. Marton, D. Hounsell, & N. Entwistle (Eds.), *The experience of learning: Implications for teaching and studying in higher education* (3rd (Internet) ed., pp. 39–58). The University of Edinburgh, Centre for Teaching, Learning and Assessment.

Matoti, S. N. (2014). Approaches to learning by pre-service science students of an institution of higher learning: an exploratory study. *Journal of Psychology in Africa, 24*(4), 385–387. https://doi.org/10.1080/14330237.2014.980627

McDonald, R. P. (2011). *Test theory: A unified treatment*. Routlege. https://doi.org/10.4324/9781410601087

Merino, S. C., & Kumar, P. R. (2013). Validación estructural del R-SPQ-2F: Un análisis factorial confirmatorio. *Revista Digital Délelőtt Investigación Eno¯sis Docencia Universitaria, 7*, 113–127.

Mundia, L., & Metussin, H. (2019). Exploring factors that improve mathematics achievement in Brunei. *Studies in Educational Evaluation, 60*, 214–222. https://doi.org/10.1016/j.stueduc.2018.10.003

Nguyen, T. M. (2016). Learning approaches, demographic factors to predict academic outcomes. *International Journal of Educational Management, 30*(5), 653–667. https://doi.org/10.1108/ijem-06-2014-0085

Önder, I., & Besoluk, S. (2010). Adaptation of revised two factor study process questionnaire (R-SPQ-2F) to Turkish. *Education and Science, 35*(157), 55–67.

Putnick, D. L., & Bornstein, M. H. (2016). Measurement invariance conventions and reporting: The state of the art and future directions for psychological research. *Developmental Review, 41*, 71–90. https://doi.org/10.1016/j.dr.2016.06.004

Ryan, R. M., & Deci, E. L. (2020). Intrinsic and extrinsic motivation from a self-determination theory perspective: Definitions, theory, practices, and future directions. *Contemporary Educational Psychology*, *61*. https://doi.org/10.1016/j.cedpsych.2020.101860

Shaik, S. A., Almarzuqi, A., Almogheer, R., Alharbi, O., Jalal, A., & Alorainy, M. (2017). Assessing Saudi medical students learning approach using the revised two-factor study process questionnaire. *International Journal of Medical Education, 8*, 292–296. https://doi.org/10.5116/ijme.5974.7a06

Shapiro, A., & ten Berge, J. M. F. (2002). Statistical inference of minimum rank factor analysis. *Psychometrika, 67*(1), 79–94.

Socha, A., & Sigler, E. A. (2014). Exploring and "reconciling" the factor structure for the revised two-factor study process questionnaire. *Learning and Individual Differences, 31*, 43–50. https://doi.org/10.1016/j.lindif.2013.12.010

Stes, A., De Maeyer, S., & Van Petegem, P. (2013). Examining the cross-cultural sensitivity of the revised two-factor study process questionnaire (R-SPQ-2F) and validation of a Dutch version. *PLoS One, 8*(1). https://doi.org/10.1371/journal.pone.0054099

Tait, H., Entwistle, N. J., & McCune, V. (1998). ASSIST: A re-conceptualisation of the approaches to studying inventory. In C. Rust (Ed.), *Improving students as learners* (pp. 262–271). Oxford Brookes University.

Timmerman, M. E., & Lorenzo-Seva, U. (2011). Dimensionality assessment of ordered polytomous items with parallel analysis. *Psychological Methods, 16*(2), 209–220. https://doi.org/10.1037/a0023353

Trizano-Hermosilla, I., & Alvarado, J. M. (2016). Best alternatives to Cronbach's alpha reliability in realistic conditions: Congeneric and asymmetrical measurements. *Frontiers in Psychology*, *7*. https://doi.org/10.3389/fpsyg.2016.00769

Valadas, S., Gonçalves, F., & Faísca, L. (2010). Approaches to studying in higher education Portuguese students: A Portuguese version of the approaches and study skills inventory for students. *Higher Education, 59*, 259–275.

Xie, Q. (2014). Validating the revised two-factor study process questionnaire among Chinese university students. *The International Journal of Educational and Psychological Assessment, 16*(1).

Zakariya, Y. F. (2019). Study approaches in higher education mathematics: Investigating the statistical behaviour of an instrument translated into Norwegian. *Education Sciences, 9*(3). https://doi.org/10.3390/educsci9030191

Zakariya, Y. F. (2021). Application of an innovative alignment optimisation method to a cross-cultural mean comparison of teacher self-efficacy: A cross-country study. *Heliyon, 7*(10). https://doi.org/10.1016/j.heliyon.2021.e08212

Zakariya, Y. F., Bjørkestøl, K., & Nilsen, H. K. (2020). Teacher job satisfaction across 38 countries and economies: an alignment optimization approach to a cross-cultural mean comparison. *International Journal of Educational Research*, *101*. https://doi.org/10.1016/j.ijer.2020.101573

Zakariya, Y. F., Bjørkestøl, K., Nilsen, H. K., Goodchild, S., & Lorås, M. (2020). University students' learning approaches: an adaptation of the revised two-factor study process questionnaire to Norwegian. *Studies in Education Evaluation*. https://doi.org/10.1016/j.stueduc.2019.100816

Zakariya, Y. F., Goodchild, S., Bjørkestøl, K., & Nilsen, H. K. (2019). Calculus self-efficacy inventory: Its development and relationship with approaches to learning. *Education Sciences, 9*(3). https://doi.org/10.3390/educsci9030170

Zakariya, Y. F., Nilsen, H. K., Bjørkestøl, K., & Goodchild, S. (2021). Analysis of relationships between prior knowledge, approaches to learning, and mathematics performance among engineering students. *International Journal of Mathematical Education in Science and Technology*, 1–19. https://doi.org/10.1080/0020739x.2021.1984596

Zakariya, Y. F., Nilsen, H. K., Goodchild, S., & Bjørkestøl, K. (2020). Self-efficacy and approaches to learning mathematics among engineering students: Empirical evidence for potential causal relations. *International Journal of Mathematical Education in Science and Technology, 53*(4), 827–841. https://doi.org/10.1080/0020739X.2020.1783006