


Article

Study Approaches in Higher Education Mathematics: Investigating the Statistical Behaviour of an Instrument Translated into Norwegian

Yusuf F. Zakariya 

Department of mathematical sciences, University of Agder, 4630 Kristiansand, Norway; yusuf.zakariya@uia.no

Received: 31 May 2019; Accepted: 19 July 2019; Published: 20 July 2019



Abstract: The revised two-factor study process questionnaire (R-SPQ-2F) has widely been considered valid and reliable in many contexts for measuring students' learning approaches. However, its cultural specificity has generated considerable discussion, with inconclusive results when translated to different languages. This paper provides more insights into the construct validity of a Norwegian version of this instrument. The R-SPQ-2F is composed of ten items designed to expose deep learning approaches and 10 items designed to expose surface learning approaches. A survey research design involving a sample of 253 first year university students in a mathematics course was adopted. Ten hypothesized models were compared using a series of confirmatory factor analyses following the model proposals reported in the literature. A weighted least square mean and variance adjusted (WLSMV) estimator was used to enhance model parameter estimations under multiple violations of assumptions inherent in ordinal data. The results favored a two first-order factor model with ten items measuring the deep approach and nine items measuring the surface approach including a deletion of one item from this instrument. The findings of this study provide empirical evidence for the cultural specificity of the instrument that is consistent with the literature. The R-SPQ-2F is therefore recommended to assess students' approaches to learning, and further studies into its cultural specificity are recommended.

Keywords: university mathematics; deep learning; surface learning; multivariate statistics; confirmatory factor analysis

1. Introduction

Empirical evidence has shown that students' learning approaches contribute significantly to their academic success in higher education (e.g., [1,2]). Learning approaches could be conceived as an individual's adopted predispositions when dealing with tasks and strategies used to process learning materials which can be deep or surface in nature [3,4]. A deep approach to learning involves concentration on latent meanings of the material to be learned, while a surface approach entails memorization and less priority on the conveyed messages in the presented tasks. Deep learning has been an emerging focus of higher education studies in preparing future leaders for our ever-increasingly diverse society [2]. For many decades, educators have been challenged by the proper conceptualization and operationalization of students' approaches to learning (SAL).

The SAL theory of Marton and Säljö [5,6] uses phenomenography coupled with some constructivist perspectives of Biggs [7,8] and has provided theoretical frameworks for conceptualizing students' approaches to learning. This is evident in the way approaches to learning have been defined to include motives, predispositions, styles, strategies used in adopting a process of learning tasks. Moreover, the classification of approaches students adopted when learning into 'surface' and 'deep' has greatly influenced SAL measuring instruments. A widely studied instrument for measuring

SAL is the Biggs' [9] study process questionnaires [10]. This instrument has undergone several revisions and validations from its initial 72-item to the present 20-item two-factor revised study process questionnaire [9]. It has gained equally wide acceptance among educators, with many studies on its psychometric properties, and Cronbach's alpha ranges 0.57–0.85 have been reported as evidence of the item's internal consistency [9,11].

However, the cultural specificity of the two-factor revised study process questionnaires (R-SPQ-2F) has generated considerable discussion with inconclusive results when the instrument has been translated to different languages (e.g., [12,13]). Apart from the two models hypothesized by Biggs, Kember and Leung [9], several alternative models have been proposed, and some items were deleted to achieve modest fits in explaining the underlying factor structures of the instrument. This current study was framed with the sole aim of exploring and comparing alternative models that best explain the construct validity of the R-SPQ-2F when translated to the Norwegian context. This article is a continuation of a work reported in [14], where the Biggs' et al. [9] hypothesized models were investigated and found to poorly represent Norwegian data with non-admissible solutions. In the earlier work, a new model for the R-SPQ-2F was proposed and confirmed, and the model fits and scale reliability were investigated and reported. The purpose of this paper is to contribute to this body of research and expose some observable methodological weaknesses inherent in some reported hypothesized models in literature.

2. Literature Review

Studies on the cultural specificity of the R-SPQ-2F can broadly be classified into two major categories. The first category represents those that report first-order two-factor structures—the deep approach (DA) and the surface approach (SA)—as the best explanatory models for the instrument with ten items on each subscale [15–18]. This first category can further be divided into those that include error covariance—the presence of a systematic commonly shared variance—between indicators (e.g., [17]) and those that did not include the covariance (e.g., [16]). However, Biggs et al. [9] were the first to start a discussion on the factor structure of their then newly developed instrument, the R-SPQ-2F, by hypothesizing and testing two models. The first model was a first-order four-factor model containing 'deep motive,' 'deep strategy,' 'surface motive,' and 'surface strategy' measured by five items each. The first model was tested and found to fit their 495 data with a comparative fit index (CFI) of 0.904 and a standardized root mean squared residual (SRMR) of 0.058 [9]. Further, CFIs of 0.997, 0.998, 0.988, and 0.998 and SRMRs of 0.01, 0.02, 0.02 and 0.02 were also reported on 'deep motive,' 'deep strategy,' 'surface motive,' and 'surface strategy' subscales, respectively. The second model was a first-order two-factor model containing deep and surface approaches with two indicators each—motive and strategy—gotten by corresponding item parceling (adding scores on five items) in the first model. The results of the second model also suggest a good model fit with a CFI of 0.992 and an SRMR of 0.015, both of which are within the proposed cutoffs by Hu and Bentler [19].

These two hypothesized models of Biggs et al. [9] have steered heated debates among educators and methodologists when subjected to confirmatory analysis in an independent cultural context. For example, the two models were tested and found to fairly explain the factor structure of the R-SPQ-2F when translated to Spanish in a study involving 836 undergraduate students, out of which 314 were used for exploratory factor analysis and the remaining 522 were used for confirmatory factor analysis [15]. An alternative model of a first-order two-factor model was proposed and tested containing the deep and surface approaches measured by their corresponding ten items each as theorized in [9]. The results suggest a modest fit with a significant χ^2 -value (169) = 645.77, $p < 0.05$, goodness of fit index (GFI) = 0.95, SRMR = 0.09, root mean square error of approximation (RMSEA) = 0.07, non-normed fit index (NNFI) = 0.91, CFI = 0.92, parsimony normed fit index (PNFI) = 0.80, and parsimony goodness of fit index (PGFI) = 0.76. In another study, Önder and Besoluk [18] reported a Turkish validation of the instrument when administered to 528 undergraduate students. Their findings also identified a first-order two-factor model as the best explanation for the construct validity of the R-SPQ-2F. Their

results involved a significant χ^2 -value (166) = 487.95, $p < 0.05$, GFI = 0.89, SRMR = 0.07, RMSEA = 0.06, NNFI = 0.90, CFI = 0.93, PGFI = 0.92, incremental fit index (IFI) = 0.93, and relative fit index (RFI) = 0.88. A major difference between these results and that of Justicia, Pichardo, Cano, Berbén and De la Fuente [15] was the inclusion of error covariance between items 8 and 10 as well as between items 11 and 20.

Non-admissible solutions and poor fits for the hypothesized models in [9] were also reported in a study involving 269 university and non-university students [17]. Following confirmatory factor analysis results, a first-order two-factor model was identified as the best explanation for the construct validity of the R-SPQ-2F. A significant χ^2 -value (168) = 259.32, $p < 0.05$, was also reported and coupled with SRMR = 0.07, RMSEA = 0.05, Tucker-Lewis index (TLI) = 0.95, and CFI = 0.96. Similar to the findings of Önder and Besoluk [18] an error covariance was also defined between items 4 and 14 in order to achieve a good model fit. Corroborative results can also be found in a Chinese validation of the R-SPQ-2F involving 439 university students, in which a first-order two-factor model was also reported [16]. Table 1 presents a juxtaposition of the findings of these studies for easy comparison.

Table 1. Summary of findings on the first-order two-factor model of the revised two-factor study process questionnaire (R-SPQ-2F).

	[M1]-Justicia et al. [15]	[M2]-Önder and Besoluk [18]	[M3]-Merino and Kumar [17]	[M4]-Xie [16]
Error Covar.		8 and 10, 11 and 20	4 and 14	
Cor. DA/SA	−0.39	−0.51	−0.33	−0.35
df	169	166	168	169
χ^2 -value	645.77	487.95	259.32	489.40
GFI	0.95	0.89		
SRMR	0.09	0.07	0.07	
RMSEA	0.07	0.06	0.05	0.07
NNFI/TLI	0.91	0.90	0.95	0.91
CFI	0.92	0.93	0.96	0.92
PNFI	0.80			0.79
PGFI	0.76	0.92		0.72
IFI		0.93		
RFI		0.88		

There seems to be a consistency in the results of previous studies presented in Table 1. They corroborate the theoretical explanation of indicators measuring the DA and the SA as proposed in [9] with an exclusion of additional subdivisions of each factor into motive and strategy. The negative standard correlation coefficients found in all the studies between the deep approach and the surface approach subscales are indicative of discriminant validity. A close look at the results of Merino and Kumar [17] as well as Önder and Besoluk [18] suggests a better fit of their models as compared to others. This can be deduced from their reduced χ^2 -values and RMSEA within the range suggested in [19,20]. However, the inclusion of error covariance between some indicators in their models could pose some complications in the application and interpretation of the scale item scores by classroom teachers.

The second broad category of studies on the cultural specificity of the R-SPQ-2F are the reports of two first-order and four first-order factor structures with some items deleted to achieve good fits (e.g., [10,21]). The number of items deleted ranged from 2–5. Immekus and Imbrie [22] after establishing non-admissible solutions of the hypothesized Biggs' et al. [9] model, subjected the data from their first cohort of 1490 university students to an exploratory factor analysis (EFA). The results of their EFA gave four extracted latent factors which they identified as 'deep motive,' 'deep strategy,' 'surface motive,' and 'surface strategy' after rotating using Promax. Five items that exhibit substantial cross-loading were removed from the model. The first-order four-factor model was then subjected to a confirmatory factor analysis in an independent cohort of 1533 university students' sample. The

results of a confirmatory analysis suggest a modest fit with a significant χ^2 -value (114) = 568.54, $p < 0.05$, RMSEA = 0.05, and CFI = 0.96. Surprisingly, relatively high positive correlations of 0.76 and 0.59 were found between 'deep motive' and 'deep strategy' as well as 'surface motive' and 'surface strategy,' respectively.

No empirical evidence was found to support the first Biggs' et al. model in the Japanese validation of the R-SPQ-2F reported in [21]. However, a modest fit for the second Biggs' et al. model involving a first-order two-factor model with item parceling on each deep and surface approach latent factors was confirmed. The study involved 269 university students distributed across different programs in a Japanese tertiary institution. The results of their confirmatory analysis did not include the χ^2 -value, instead an RMSEA = 0, CFI = 1, and TLI = 1 coupled with a positive correlation coefficient of 0.30 between the deep and surface approach latent factors were reported [21]. There are some reservations with respect to these results. First, the goodness of fits (GOF) indices indicate a perfect fit of the model, which appears to be unrealistic. However, an observed methodological issue could stem from the degrees of freedom (though not reported), which is 1. This could make it difficult for the variance/covariance matrix to be positively definite. Unfortunately, nothing was mentioned in the article with respect to this matrix. Another methodological difficulty that could even lead to the rejection of this model is the positive correlation of 0.30 reported between deep and surface approaches. This shows a non-discriminating capacity of this model between the deep and surface approaches which is contrary to both the theoretical and the conceptual interpretations of the instrument.

More so, Stes, De Maeyer and Van Petegem [12] could not also find any supportive empirical evidence for both models hypothesized by Biggs et al. [9] in the validation of their Dutch version of the instrument involving 1974 effective sample of students distributed across diverse university programs. For this reason, an exploratory factor analysis was conducted on a randomly selected 963 cases from the total sample, using maximum likelihood for factor extraction and an oblique rotation. Five factors were initially identified, and these were later collapsed into four factors—study is interesting (SI), spending extra time (ST), minimal effort (ME), and learning by heart (LH)—after a series of confirmatory factor analyses and item deletions. The final fitted solution was a first-order four-factor model with three items measuring SI, four items measuring ST, five items measuring ME and three items measuring LH. The final chi-squared statistic as well as the degree of freedom were not reported. However, some GOF indices such as GFI = 0.95, absolute goodness of fit index (AGFI) = 0.93, RMSEA = 0.06, CFI = 0.94, and PGFI = 0.66 were reported as evidence of a good fit for their model. Further, relatively high correlation coefficients of 0.76 and 0.62 were found between SI and ST as well as between ME and LH, respectively.

In an attempt to reconcile between variant inconclusive models results on the R-SPS-2F, Socha and Sigler [13] conducted a validation study on the instrument involving 868 university students. In their study, eight models were compared using a confirmatory factor analysis, and a first-order two-factor solution was found as the best explanation for the construct validity of the instrument involving the deletion of two items from the original version. Their final results included a significant χ^2 -value (134) = 504.83, $p < 0.05$, SRMR = 0.05, RMSEA = 0.06, and CFI = 0.95 and a negative correlation of -0.38 between the deep and surface approach latent factors. Similar results can also be found in another Spanish validation of the R-SPQ-2F involving 279 university students, in which a first-order two-factor model coupled with two item deletion was also reported [10]. Table 2 presents a juxtaposition of the findings of these studies for easy comparison.

Table 2. Summary of findings on the two-factor and four-factor models of the R-SPQ-2F.

	[M5]-Immekus and Imbrie [22]	[M6]-Fryer et al. [21]	[M7]-Stes et al. [12]	[M8]-Socha and Sigler [13]	[M9]-López-Aguado and Gutiérrez-Provecho [10]
Item deleted	1, 3, 7, 13, 15		2, 3, 7, 10, 17	7, 8	7, 8
Cor. DA/SA	DM/DS = 0.76 SS/SM = 0.59	0.30	SI/ST = 0.76 ME/LH = 0.62	−0.38	−0.41
df	114			134	134
χ^2 -value	568.54			504.83	226.53
GFI			0.95		0.91
AGFI			0.93		
SRMR				0.05	0.05
RMSEA	0.05	0	0.06	0.06	0.05
CFI	0.96	1	0.94	0.95	0.91
TLI		1			0.90
PGFI			0.66		0.72

Note: DM = Deep motive; DS = Deep strategy; SM = Surface motive; SS = Surface strategy; SI = Study is interesting; ST = Spending extra time; ME = Minimal effort; and LH = Learning by heart.

The results presented in Table 2 reveal variant and inconclusive solutions of the models. These can be ascribed to some methodological issues inherent in the factor analysis procedure as well as the cultural sensitivity of the instrument. For example, Immekus and Imbrie [22], after establishing non-admissible solutions of the hypothesized models in [9], subjected their data to an exploratory factor analysis (FA). Difficulties arose when some indicators loaded (loadings greater than |0.3|) on more than one extracted factor. Rather than seeking theoretical explanations for this observation, they opted to delete these indicators from the scales. For instance, item 1 loaded on deep motive (DM) and deep strategy (DS) with 0.31 and 0.42 oblique rotated loadings, respectively. This could be suggestive of over-factoring in the extraction, especially when this item has been theorized to measure both DM and DS. To support this claim, the high positive correlations of 0.76 and 0.59 reported between DM and DS as well as surface strategy (SS) and surface motives (SM), respectively, are indications of multicollinearity, which could be addressed by collapsing the subcategories.

A similar methodological issue is also perceived in the analysis of Stes et al. [12] with high positive correlations of 0.76 and 0.62 between SI and ST as well as ME and LH, respectively. Another methodological issue involved in the analysis of Stes et al. [12] and Fryer et al. [21] is the use of maximum likelihood estimator, which has been found to perform poorly in the analysis of ordinal data (e.g., [23,24]). It is also important to remark that SI combined with ST and LH combined with ME are other ways to refer to the DA and the SA, respectively. Later studies (e.g., [10]) seem to address some of these methodological issues, yet the cultural specificity of the R-SPQ-2F still remains an important consideration when adapted to a different language from English. Therefore, the current study sought to build on this literature in searching and evaluating hypothesized models to explain the construct validity of the R-SQP-2F in the Norwegian context.

3. Methods

3.1. Research Design and Materials

An appropriate design for this study is a survey type which can be used to justify the collection of data from a large number of students using questionnaires. A Norwegian version of the R-SPQ-2F was prepared using translation and back-translation approaches by two associate professors and a full professor of mathematics education in the research team. A comparison was made between translated versions, and some adjustments for language differences were done before the final version was prepared in an electronic form.

3.2. Sample and Data Collection Procedure

The sample comprised 253 undergraduate students on engineering programs in their first year across two Norwegian government universities. This was made up of males (168) and females (72), including 13 students who did not respond to the gender item on the instrument, and their sample had a mean age of 19–22 years. The sample size was considered appropriate with a justification from reported computer simulated studies on the adequacy of a confirmatory factor analysis (CFA) sample size (e.g., [25]). The data were collected using an online questionnaire development kit (SurveyXact) and distributed to students (via university mails) who gave their consent to partake in the project. Class visits were organized to encourage students' participation, and some students completed a paper version of the questionnaires during the visits. The data were collected within 3 weeks, screened for outliers, and prepared for CFA. Very few data were missing at random and posed no concern for the analysis.

3.3. Data Analysis

CFA was used to test ten models with a weighted least square mean and variance adjusted (WLSMV) estimator using Mplus version 8.2 [26] for the analysis. The use of WLSMV was not accidental. It was because of its robust ability to perform better than maximum likelihood (ML), unweighted least squares (ULS), etc., in the analysis of ordinal data that violate multiple assumptions (see [23]). Model fits were assessed based on χ^2 -values and combinations of the goodness of fit indices with some relaxations in the suggested cut-offs (CFI/TLI closed to or ≥ 0.9 , SRMR ≤ 0.8 , and RMSEA ≤ 0.60) proposed by Hu and Bentler [19]. This was necessary because of the type of data (ordinal) as well as a different estimator (WLSMV) as compared to the continuous data and ML estimator used in some simulation studies on cut-off criteria (e.g., [19,27]). The emerging results and discussion are presented in the next few paragraphs.

4. Results

The first set of results as presented in Table 3 represent the tested and hypothesized two first-order factor models of the R-SPQ-2F, as in the literature. Analyzed results from hypothesized model of Xie [16] are included in Table 3, and those of Justicia et al. [15] were excluded, because they both practically advocated the same model and the former is more recent. Notations and abbreviations used in Table 1 are repeated in Table 3, with M2 used for Önder and Besoluk [18], M3 used for Merino and Kumar [17], and M4 used for Xie [16].

Table 3. Selected CFA results of the two first-order factor hypothesized model of the R-SPQ-2F.

	M2	M3	M4
Error Covar.	8 and 10, 11 and 20	4 and 14	
Cor. DA/SA	−0.519	−0.507	−0.512
df	167	168	169
χ^2 -value	495.212	517.980	522.179
<i>p</i> -value	0.000	0.000	0.000
CFI	0.785	0.771	0.769
TLI	0.756	0.741	0.740
SRMR	0.078	0.081	0.081
RMSEA	0.088	0.091	0.091
90% C.I.	0.079 0.097	0.082 0.100	0.082 0.100

The results presented in Table 3 show admissible solutions of the two first-order factor model of the R-SPQ-2F. Negative standard correlations found between deep and surface components are suggestive of a discriminant validity between these subscales. This could be interpreted to mean a student with a high score on deep approach items had a low score on surface approach items and vice versa. This

makes sense and is conceptually sound. However, the high χ^2 -values (496.21–522.18) coupled with out of range fit indices are indicative of the poor fit of these models. The model proposed by Önder and Besoluk [18] seems to perform better than others, with the lowest χ^2 -value (167, $p < 0.05$) = 495.21 and RMSEA ≤ 0.08 . Meanwhile, the two error covariances involved between item 8 and item 10, as well as item 11 and item 20, could pose some complications in classroom conceptual understanding and interpretation of scores from this instrument. Therefore, all these models are not statistically and conceptually fit to justify the construct validity of the R-SPQ-2F in the Norwegian context.

The second set of results concern the CFA of the two-factor and four-factor models hypothesized to explain construct validity of the R-SPQ-2F. The model result of Fryer et al. [21] was not included in Table 4 because it has been reported in [14]. The analyzed result showed a non-admissible solution of the model with a negative error variance on the SM indicator, a result that is suggestive of over-factoring in the model [14]. Further, analyzed results from the model hypothesized by López-Aguado and Gutiérrez-Provecho [10] were included in Table 4, but those of Socha and Sigler [13] were omitted because they are both practically advocating same model and the former is more recent. Notations and abbreviations used in Table 2 are repeated in Table 4 with M5 used for Immekus and Imbrie [22]. Mod. M5 was used for modified M5, M7 was used for Stes et al. [12], Mod. M7 was used for modified M7, and M9 was used for López-Aguado and Gutiérrez-Provecho [10].

Table 4. Selected CFA results of the two-factor and four-factor hypothesized models on the R-SPQ-2F.

	M5	Mod. M5	M7	Mod. M7	M9
Cor. DA/SA	DM/DS = 0.737 SS/SM = 0.366	−0.290	SI/ST = 1.020 ME/LH = 0.401	−0.284	−0.40
df	84	89	85	89	134
χ^2 -value	152.278	289.254	138.318	257.024	301.440
p -value	0.000	0.000	0.0002	0.000	0.000
CFI	0.925	0.780	0.950	0.844	0.869
TLI	0.906	0.741	0.939	0.816	0.850
SRMR	0.054	0.077	0.052	0.072	0.068
RMSEA	0.057	0.094	0.050	0.086	0.070
90% C.I.	0.042 0.071	0.082 0.107	0.034 0.065	0.074 0.099	0.060 0.081

There seems to be indications of good fits in all the models analyzed and reported in Table 4. The reduced χ^2 -values between 152.28 and 301.44 coupled with within suggested range indices may prompt one to conclude that M5 and M7 have been demonstrated as the best models. However, there was evidence of a gross misspecification and a high multicollinearity between DM and DS, which are suggestive of over-factoring in M5. This is evident with a high standardized correlation coefficient ($r = 0.74$, $p < 0.05$) between the DM and DS latent factors. This posed some methodological difficulties involved in trying to balance both the theoretical and conceptual understanding that could yield a substantive interpretation of scores from the instrument. Therefore, an attempt was made to revive this model as reported under the heading modified model 5 (Mod. M5). Here, items measuring DM and DS were merged to form a factor (DA), and those measuring SS and SM were merged to form another factor (SA). The resulting two-factor model was subjected to CFA, and selected GOF indices are presented in Table 4 with the heading Mod. M5. The χ^2 -value (89, $N = 253$) = 289.25 became bigger, and all the fit indices were out of range.

The analyzed results of the proposed model by Socha and Sigler [13] were even worse. The latent variance-covariance matrix was not positively definite, which is a necessary condition for an acceptable model (see [28]). This was observed with the presence of the Heywood case in terms of a standardized correlation coefficient great than 1 between latent factors SI and ST. In a similar manner to M5, this model was modified, and the CFA results were reported with the heading Mod. M7 in Table 4. The resulting χ^2 -value (89, $N = 253$) = 257.02, $p < 0.05$ is significant, but, when combined with GOF indices, qualifies the model to an appropriate fit of the data [29]. However, a comparison of this model

results with the ones reported in Table 4 with the heading M9 [10] favored the latter. This is evident with the higher CFI/TLI and lower SRMR and RMSEA values observed in M9. Therefore, from the foregoing discussion, what appears to be the best explanation of the factor structure of the R-SPQ-2F in Norwegian context is the hypothesized model of López-Aguado and Gutiérrez-Provecho [10].

5. Conclusions

Teaching is considered successful when accompanied by meaningful learning. A good way to ensure that successful learning takes place is to investigate the approaches adopted by students when learning. Several efforts have been expended to promote deep learning in higher education such that the emerging leaders will better be prepared for an ever-increasingly diverse society [30,31]. Both qualitative and quantitative studies have, for several years, been directed towards proper conceptualization and operationalization of students' approaches to learning. Prominent among these studies are the works of Marton and Säljö [5,6], Entwistle and Waterston [32], and Biggs [33]. These have led to the development of measuring instruments in which the study process questionnaire (SPQ) seems to have gained global attention. However, studies on the cultural specificity of the latest SPQ called the R-SPQ-2F have generated diverse and inconclusive results.

In this article, investigations were geared towards addressing the issue of R-SPQ-2F cultural specificity when applied in the Norwegian context. Several models were compared, and what seems to be the best explanation of the R-SPQ-2F construct validity is a two first-order factor model involving deep and surface approaches to learning subscales. Meanwhile, by comparing the identified tested hypothesized model proposed in [10,13], as reported in Table 4, with the results of the appropriate fit (χ^2 -value (151, N = 253) = 377.68, $p < 0.05$, SRMR = 0.072, CFI = 0.844, TLI = 0.824, and RMSEA = 0.077) found in [14], a conclusion could be drawn. There appears to be no obvious statistical difference with consideration for respective χ^2 -values and GOF indices of these two models. Therefore, a two first-order factor model with 10 items measuring the deep approach and nine items (contrary to eight items in [10]) measuring the surface approach is still considered the best explanation for the R-SPQ-2F construct validity.

The justification for removing one item from the instrument was previously explained in detail in the first article. This study is to be followed up with an independent sample that will be collected in the near future for a confirmation of the proposed model and the predictive validity of the R-SPQ-2F. An interpretation of item scores can be achieved simply by adding corresponding items on the deep approach (scaled by dividing the sum by 10) and those on the surface approach (scaled by dividing the sum by 9) for classroom decisions. It is hoped that future replications of this study across other universities and groups of students will be carried out. This instrument is therefore recommended for measuring year-one undergraduate students' approaches in Norwegian universities. For instructional purposes, both data and Mplus syntax codes used for this study are available upon request from the author.

Funding: The article processing charge (APC) was funded by the University of Agder library.

Acknowledgments: The author acknowledges the support of the faculty of engineering and science, University of Agder, Kristiansand, Norway as well as MatRIC, centre for research, innovation, and coordination of mathematics teaching. Thanks to K. Bjørkestøl, H. K. Nilsen, and S. Goodchild for providing guidance and help with the Norwegian language as well as M. Lorås for providing some data.

Conflicts of Interest: The author declares no conflict of interest.

References

1. Maciejewski, W.; Merchant, S. Mathematical tasks, study approaches, and course grades in undergraduate mathematics: A year-by-year analysis. *Int. J. Math. Educ. Sci. Technol.* **2016**, *47*, 373–387. [[CrossRef](#)]
2. Fryer, L.K.; Vermunt, J.D. Regulating approaches to learning: Testing learning strategy convergences across a year at university. *Br. J. Educ. Psychol.* **2018**, *88*, 21–41. [[CrossRef](#)] [[PubMed](#)]

3. Baeten, M.; Kyndt, E.; Struyven, K.; Dochy, F. Using student-centred learning environments to stimulate deep approaches to learning: Factors encouraging or discouraging their effectiveness. *Educ. Res. Rev.* **2010**, *5*, 243–260. [[CrossRef](#)]
4. Biggs, J.B.; Tang, C. *Teaching for Quality Learning at University*, 3rd ed.; The Society for Research into Higher Education & Open University Press: Berkshire, UK, 2007.
5. Marton, F.; Säljö, R. On qualitative differences in learning II: Outcome as a function of the learner's conception of the task. *Br. J. Educ. Psychol.* **1976**, *46*, 115–127. [[CrossRef](#)]
6. Marton, F.; Säljö, R. On qualitative differences in learning I: Outcome and process. *Br. J. Educ. Psychol.* **1976**, *46*, 4–11. [[CrossRef](#)]
7. Biggs, J.B. *The Study Process Questionnaire (SPQ): Manual*; Australian Council for Educational Research: Hawthorn, Australia, 1987.
8. Biggs, J.B. What do inventories of students' learning processes really measure? A theoretical review and clarification. *Br. Educ. Res. J.* **1993**, *63*, 3–19. [[CrossRef](#)]
9. Biggs, J.B.; Kember, D.; Leung, D.Y.P. The revised two factor study process questionnaire: R-SPQ-2F. *Br. Educ. Res. J.* **2001**, *71*, 133–149. [[CrossRef](#)]
10. López-Aguado, M.; Gutiérrez-Provecho, L. 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. *Cult. Educ.* **2018**, *30*, 105–141. [[CrossRef](#)]
11. Chan, Y.M.; Sheung Chan, C.M. Approach to learning of sub-degree students in Hong Kong. *Res. Educ.* **2010**, *84*, 65–78. [[CrossRef](#)]
12. Stes, A.; De Maeyer, S.; Van Petegem, P. Examining the cross-cultural sensitivity of the revised two-factor study process questionnaire (R-SPQ-2F) and validation of a Dutch version. *PLoS ONE* **2013**, *8*, e54099. [[CrossRef](#)]
13. Socha, A.; Sigler, E.A. Exploring and “reconciling” the factor structure for the revised two-factor study process questionnaire. *Learn. Individ. Differ.* **2014**, *31*, 43–50. [[CrossRef](#)]
14. Zakariya, Y.F.; Bjørkestøl, K.; Nilsen, H.K.; Goodchild, S.; Lorås, M. University students' learning approaches: An adaptation of the revised two-factor study process questionnaire to Norwegian. *Stud. Educ. Eval.* **2019**. under revision.
15. Justicia, F.; Pichardo, M.C.; Cano, F.; Berbén, A.B.; De La Fuente, J. The revised two-factor study process questionnaire (R-SPQ-2F): Exploratory and confirmatory factor analyses at item level. *Eur. J. Psychol. Educ.* **2008**, *23*, 355–372. [[CrossRef](#)]
16. Xie, Q. Validating the revised two-factor study process questionnaire among Chinese university students. *Int. J. Educ. Psychol. Assess.* **2014**, *16*, 4–21.
17. Merino, S.C.; Kumar, P.R. Validación estructural del R-SPQ-2F: Un análisis factorial confirmatorio. *Rev. Digit. Investig. Docencia Univ.* **2013**, *7*, 113–127. [[CrossRef](#)]
18. Önder, I.; Besuluk, S. Adaptation of revised two factor study process questionnaire (R-SPQ-2F) to Turkish. *Educ. Sci.* **2010**, *35*, 55–67.
19. Hu, L.T.; Bentler, P.M. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Struct. Equ. Modeling Multidiscip. J.* **1999**, *6*, 1–55. [[CrossRef](#)]
20. Hu, L.T.; Bentler, P.M. Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychol. Methods* **1998**, *3*, 424–453. [[CrossRef](#)]
21. Fryer, L.K.; Ginns, P.; Walker, R.A.; Nakao, K. The adaptation and validation of the CEQ and the R-SPQ-2F to the Japanese tertiary environment. *Br. J. Edu. Psychol.* **2012**, *82*, 549–563. [[CrossRef](#)] [[PubMed](#)]
22. Immekus, J.C.; Imbrie, P.K. A test and cross-validation of the revised two-factor study process questionnaire factor structure among western university students. *Educ. Psychol. Meas.* **2010**, *70*, 495–510. [[CrossRef](#)]
23. Suh, Y. The performance of maximum likelihood and weighted least square mean and variance adjusted estimators in testing differential item functioning with nonnormal trait distributions. *Struct. Equ. Modeling Multidiscip. J.* **2015**, *22*, 568–580. [[CrossRef](#)]
24. Brown, T.A. *Confirmatory Factor Analysis for Applied Research*; The Guilford Press: New York, NY, USA, 2015.
25. Wolf, E.J.; Harrington, K.M.; Clark, S.L.; Miller, M.W. Sample size requirements for structural equation models: An evaluation of power, bias, and solution propriety. *Educ. Psychol. Meas.* **2013**, *76*, 913–934. [[CrossRef](#)] [[PubMed](#)]
26. Muthén, L.K.; Muthén, B.O. *Mplus User's Guide*, 8th ed.; Muthén & Muthén: Los Angeles, CA, USA, 2017.

27. Bentler, P.M. Comparative fit indexes in structural models. *Psychol. Bull.* **1990**, *107*, 238–246. [[CrossRef](#)] [[PubMed](#)]
28. Byrne, B.M. *Structural Equation Modeling with Mplus: Basic Concepts, Applications, and Programming*; Routledge, Taylor & Francis Group: New York, NY, USA, 2012.
29. Marsh, H.W.; Hau, K.T.; Wen, Z. 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. *Struct. Equ. Modeling Multidiscip. J.* **2004**, *11*, 320–341. [[CrossRef](#)]
30. Lake, W.W.; Boyd, W.E. Is the university system in Australia producing deep thinkers? *Aust. Univ. Rev.* **2015**, *57*, 54–59.
31. Barattucci, M. Approach to study as an indicator of the quality of teaching and of learning environment: The contribution of John Biggs. *J. E Learn. Knowl. Soc.* **2017**, *13*, 77–88.
32. Entwistle, N.; Waterston, S. Approaches to studying and levels of processing in university students. *Br. Educ. Res. J.* **1988**, *58*, 258–265. [[CrossRef](#)]
33. Biggs, J.B. Individual and group differences in study processes. *Br. J. Educ. Psychol.* **1978**, *48*, 266–279. [[CrossRef](#)]



© 2019 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).