Title	Self-regulated learning in Singaporean context: A congeneric approach of confirmatory factor analysis
Author(s)	Betsy Ng, C. K. John Wang and W. C. Liu
Source	*International Journal of Research and Method in Education, 40*(1), 91-107
Published by	Taylor & Francis (Routledge)

Citation: Ng, B., Wang, J. C. K., & Liu, W. C. (2017). Self-regulated learning in Singaporean context: A congeneric approach of confirmatory factor analysis. *International Journal of Research and Method in Education, 40*(1), 91-107. https://doi.org/10.1080/1743727X.2015.1075125

Self-regulated learning in Singaporean context: A congeneric approach of confirmatory factor analysis

Betsy Ng, C.K. John Wang, W.C. Liu

Abstract

The motivated strategies for learning questionnaire (MSLQ) is widely used as a self-report instrument to assess students' motivation and self-regulation. This study utilized the MSLQ Junior High to examine the motivational beliefs and self-regulation of secondary school students (Grades 8 and 9) from Singapore. The instrument was slightly modified and administered to students (N = 610) in mathematics and science classes. In the first sample, 314 students completed the MSLQ Junior High while a second sample of 296 students completed the revised MSLQ Junior High. Using a congeneric approach of confirmatory factor analysis (CFA), the five-factor measurement model was determined with the first sample. This measurement model was further supported using a second sample and its goodness-of-fit indices were compared with other CFA models. Statistical findings showed that the five-factor structure of the revised MSLQ Junior High had a good model fit. The present study contributes a novel methodological approach by investigating the parsimony confirmatory factor structure of the revised MSLQ Junior High in local academic contexts.

Keywords: confirmatory factor analysis, motivation, MSLQ, self-regulated learning

Introduction

Despite the extensive empirical studies on the measures of motivation and cognition, there is still room for improvement on the psychometric properties of the developed instruments. The motivated strategies for learning questionnaire (MSLQ) had been used extensively in the States and western settings (e.g., Cheang, 2009; Lynch, 2006; Zusho, Pintrich, & Coppola, 2003). Yet, previous psychometric studies of the MSLQ (Davenport, 2003; Pintrich, Smith, Garcia, & McKeachie, 1993) did not obtain good findings of confirmatory factor analysis (CFA) analyses. The MSLQ is likely a valuable tool in research and practical settings but it still has varied factorial validity. A recent meta-analysis reported that problematic psychometric properties could likely explain the low validity of the MSLQ (Credé & Phillips, 2011). According to Credé & Phillips (2011), there is no explicit examination of the validity of the MSLQ's assumption – "students are differently motivated and use different learning strategies for different courses" (p. 344), indicating that the psychometric properties of the MSLQ still needs attention. Furthermore, it is essential to assess the construct validity of an instrument in a new or different cultural setting (Maneesriwongul & Dixon, 2004). Based on local educational settings, Rotgans and Schmidt (2009, 2010) investigated the psychometric properties of the MSLQ College (Pintrich, Smith, Garcia, & McKeachie, 1991) on first-year polytechnic students. They obtained satisfactory model fit by performing two separate CFA analyses on the MSLQ's motivational and self-regulated learning components, respectively. The MSLQ College assesses students' motivation and self-regulatory skills in colleges, universities or tertiary institutions, while the shortened version (known as MSLQ Junior High) is meant for students from junior high or secondary schools. Both MSLQs differ from each other: MSLQ College has 81 items with 15 subscales, whereas MSLQ Junior High has 44 items with 5 subscales. These subscales are described subsequently. As there are very few psychometric studies on the utility of the

MSLQ Junior High, the present study examines its reliability and validity across specific subjects, i.e., mathematics and science.

Motivational beliefs such as self-efficacy plays an important role in improving mathematical and scientific problem-solving skills of middle and high school students (e.g., Pajares & Graham, 1999; Zusho et al., 2003). Students with highly perceived self-efficacy in mathematics are more likely to use higher-order cognitive and metacognitive strategies (Berger & Karabenick, 2011). Likewise, there is an emphasis on the importance of inquiry skills where students discover, produce and evaluate scientific knowledge (Kim, Tan, & Talaue, 2013; Yoon, 2009). Coupled with scientific thinking and reasoning skills, students are encouraged to experience the knowledge construction process (Chinn & Malhotra, 2002). Taken together, students are likely more strategic and engaged in learning contexts such as mathematics and science classes.

With respect to the utility of the instrument in an Asian educational setting, we investigated the reliability and validity of the MSLQ Junior High (Liu et al., 2012) in mathematics and science classes. The MSLQ Junior High evaluates the effectiveness of learning amongst middle or junior high school students (Pintrich and De Groot, 1990). It includes three motivational subscales namely intrinsic value, self-efficacy and test anxiety, as well as two metacognitive subscales namely cognitive strategy use and self-regulation. Rao and Sachs (1999) performed CFA on the five-factor model of the 44-item MSLQ Junior High in Hong Kong and compared to that of Pintrich and De Groot (1990). Their findings of three motivational factors – intrinsic value, self-efficacy and test anxiety were consistent with Pintrich and De Groot (1990). However, there were differences in two metacognitive and cognitive factors, which could be due to the characteristics of the Chinese learner. Chinese students tend to learn by rote learning (Law, Chan, & Sachs, 2008), whereas learners in western culture are encouraged in metacognitive strategies (Rao & Sachs, 1999). Rote learning is considered as “mechanical” and a lower cognitive strategy (Marton, Watkins, & Tang, 1997; Purdie & Hattie, 1996). Hence, Rao and Sachs (1999) proposed to combine two self-regulated learning factors (i.e., cognitive and metacognitive strategies) into one factor. They also recommended three reverse-coded items of self-regulated learning strategies to form a separate factor.

Lee et al. (2010) examined the psychometric properties of the Chinese version of MSLQ Junior High (developed by Rao and Sachs, 1999) on Hong Kong junior secondary school students. They revised the reverse-coded items of self-regulated strategies into positive statements, and combined strategy use and self-regulation into a single factor. The revised instrument contained 44 items and 4 factors. Although the four-factor CFA model demonstrated good fit, test anxiety had non-significant correlation with self-efficacy as well as weak correlations with intrinsic value and strategy use.

Liu and her colleagues (Liu et al., 2012) investigated the psychometric properties of the MSLQ Junior High (Pintrich & De Groot, 1990) on eight secondary schools in Singapore. Combined exploratory factor analysis (EFA) and CFA with two samples, they conducted EFA on the first sample, followed by CFA on the second sample in search for a well-fit measurement model of the MSLQ Junior High. Consistent with Rao and Sachs’ (1999) suggestions, Liu and colleagues renamed the items on cognitive and metacognitive strategies as learning strategies since secondary school students might not differentiate them. In addition, three negatively worded items for self-regulation were renamed as “lack of self-regulation”. With the deletion of items by EFA, the 44-item MSLQ was then reduced to 28 items. Their CFA analysis showed a better model fit than the factor structure proposed by Pintrich and De Groot (1990). Nonetheless, it was still not a conventionally good model fit as comparative fit index (CFI) and non-normed fit index (NNFI) were still below .95 (Byrne, 2001). This suggests that there is still room for improvement – a need to re-examine the psychometric properties of the MSLQ Junior High. Liu et al. (2012) revealed a negative relationship between test anxiety and lack of self-regulation (r = -.68; 95% confidence interval of -.92, -.44). This finding is contradictory as lack of self-regulation can cause anxiety (Baumiester & Heatherton, 1996). Pintrich et al. (1993) found that test anxiety is negatively correlated to self-regulation whereas Bembenutty et al. (1998) did not obtain a significant relationship between test anxiety and self-regulation. Due to the inconsistent findings, this study aimed to uncover this relationship between test anxiety and lack of self-regulation.

Based on existing knowledge, there are only four studies that investigated the reliability and validity of the MSLQ Junior High (see Table 1). Within these studies, three of them did not confirm factorial validity of the identified scale with a second sample. Moreover, Pintrich and De Groot (1990) did not provide the details and procedures of factor analysis such as the type of factor extraction methods. According to Hu and Bentler's (1999) recommendation, it is essential to evaluate the hypothesized model with the goodness-of-fit indices.

#Insert Table 1 about here#

Despite the recent psychometric study by Liu et al. (2012), the construct of cognition in the 28-item MSLQ Junior High still consists of a large number of learning strategies items. This one-dimensional factor (i.e., learning strategies) still has room for refinement and item-reduction. In most research studies, the items of an original one-dimensional factor were reduced without much consideration of the integrity of the larger construct (MacCallum & Austin, 2000). To support the validity of a construct, other types of scale refinement were included in this study. This explains why we used the congeneric approach of CFA for first sample as EFA does not consider the structure of original factors that constitutes to the model structure (Larwin & Harvey, 2012). Congeneric CFA refers to the approach of testing multiple factors in the context of multifactor CFA models (Sinclair, Dowson, & McInerney, 2006), that is, from a single-factor CFA model to multifactor CFA models. As congeneric CFA is less documented in research, this study aimed to highlight this novel methodological contribution – congeneric models are considered parsimonious by determining "the qualities of these items and scales free from disturbance errors associated with other factors (Sinclair et al., 2006, p. 1144). In short, the congeneric approach provides clear guidelines on item reduction. On the other hand, EFA and Rasch analysis are more relevant to exploring of new inventories. Rasch analysis is the item response theory (IRT) model used to test the "theoretical construct validity and adequacy of a scale" (Bohn et al., 2008, p. 1107). Rasch model strongly depends on unidimensionality of the questionnaire whereby each scale is tested individually (Kaipper, Chachamovich, Hidalgo, da Silva Torres, & Caumo, 2010). Since the MSLQ is considered a well-established inventory to measure students' motivational beliefs and cognition (Larwin & Harvey, 2012), the congeneric CFA approach is considered more appropriate for generalization of existing scales.

To sum, the objective of the present study was to investigate the psychometric properties of the 28-item MSLQ Junior High on secondary school students' motivation and self-regulated learning towards academic subjects, namely mathematics and science. This study also aimed to refine the testing of the MSLQ's psychometric properties using congeneric CFA.

Method

Participants

Data were collected from 610 students (14 to 16 years old; Grades 8 and 9) from six secondary schools in Singapore. The sampling included two academic streams namely, express and normal. Express stream students complete their secondary schooling in 4 years and they sit for Education-Ordinary Level (GCE-O) examinations in their last year. However, students with weaker ability are channelled to normal stream whereby they sit for GCE-Normal Level (GCE-N) exam in the end of fourth year or GCE-O exam in the end of fifth year. All subjects (except second language) were taught in British English.

The first sample (Sample 1) contained 314 students (165 males, 146 females and 3 did not state gender), while the second sample (Sample 2) had 296 students (139 males, 149 females and 8 did not state gender). Within Sample 1, there were 213 express students and 101 normal students, with a mean age of 14.65 (*SD* = .86). Within Sample 2, there were 181 express students and 115 normal students, with a mean age of 14.70 (*SD* = .68).

Procedure

Prior to the data collection, ethic clearance from the university review board and permission from Ministry of Education were attained. For a representative sampling, schools from four cluster zones (north, east, south and west) were invited to participate in this study. Participants were briefed on the purpose of the study and were assured of the confidentiality of their responses. Specific instructions were given to the participants and their participation was voluntary. The MSLQ with reference to specific subjects (either mathematics or science) was administered in quiet classroom settings and the participants took about 20 minutes to complete the questionnaire (in English).

Measure

Motivated Strategies for Learning Questionnaire (MSLQ)

The 28-item MSLQ Junior High (Liu et al., 2012) was adapted in this study, with reference to specific subjects and slight modifications (described subsequently). Cronbach's alpha (α) coefficients for each subscale were computed (Cronbach, 1951). Three subscales of the motivation scales and two subscales of the learning strategies of the MSLQ were adapted as shown in Table 2.

#Insert Table 2 about here#

Students rated all the items on a 7-point Likert scale, from 1 (*not at all true of me*) to 7 (*very true of me*) in the specific academic context (either mathematics or science). For instance, an item concerning intrinsic value in mathematics or science is “I prefer Mathematics work that is challenging …” or “I prefer Science work that is challenging…”.

Slight modifications of MSLQ items

As the item wordings of the MSLQ originated in the American context, several items were punctuated to make them easier for the local students to comprehend. For example, an item for the self-efficacy scale (in mathematics context) was “Compared with other students in this class I think I know a great deal about mathematics.” There was no comma in the statement to allow pause for the ease of comprehension, in particularly to students with weaker reading abilities. Commas are known as “weaker cues that signal within-sentence clause boundaries” (Just & Carpenter, 1980, p. 346). In this case, this item was clarified by an addition of comma: “Compared with other students in this class, I think I know a great deal …”. Other items with long statements were also refined with the addition of commas (e.g., “When reading, I try to connect the things I am reading …”).

Another slight modification made was the use of verb. In the American context, the verb and noun for “practice” are of the same spelling. However, in the British English context (according to Singapore educational system), there is a difference between “practice” and “practise”. In British English, the word “practice” is the noun whereas the word “practise” is the verb (Oxford Dictionaries, 2013). With reference to the British reading system, the item for learning strategies scale was amended. The item that was originally “When I study for a test I practice saying …" was then refined to "When I study for a test, I practise saying…". Although these modifications made to some items might seem minimal, the key principle was to capture the true meaning of students' responses.

Data Analysis

To test the measurement model, an estimation of internal consistency and convergent validity of the MSLQ items were conducted. The measures of internal consistency included Cronbach's α and composite reliability (CR). The reliability measures should be at least .70 to indicate adequate internal consistency (Nunnally & Bernstein, 1994). Subsequently, convergent validity is adequate when constructs have an average variance extracted (AVE) of at least .5 (Fornell & Larcker, 1981). CFA was conducted with EQS 6.2 for Windows.

Confirmatory factorial analysis

For CFA, maximum likelihood estimation (MLE) was used to assess the parameter estimation and evaluate the goodness of fit (Myung, 2003). MLE is the "standard approach to parameter estimation and inference in statistics" (p. 90). MLE is not only an efficient method for parameter estimation; it also provides consistency in true parameter values that generate data of sufficiently large samples. Furthermore, it is a pre-requisite for the chi-square test and inference with missing data which seeks for the desired probability description.

To evaluate a good model fit, Hu and Bentler (1999) suggested the following criteria: >.95 for both NNFI and CFI, as well as <.06 for root mean square error of approximation (RMSEA). In addition, the chi-square and χ^2/df ratio were used to examine the adequacy of models. For an excellent model fit, RMSEA value should be below .05 (Kline, 2005). Alternatively, to conclude a good model fit, the robust NNFI and CFI should be at least .95 while the robust RMSEA is recommended .05 or less (Yu & Muthen, 2002). In addition, the Satorra-Bentler scaled chi-square ($SB\chi^2$) was used since its asymptotic statistic test is robust for a large sample size (Satorra & Bentler, 2001). SBχ^2/df ratio should be below the recommended value of 3.0 (Kline, 2005) to indicate a good fit of a specified model to the data.

As described subsequently, CFA analysis using MLE in EQS 6.2 for Windows was conducted to test the model fit of data for both samples (1 and 2). First, congeneric CFA was conducted to test the fit of each model structure before deriving the final five-factor model. The details of congeneric approach are described in the subsequent section. Using the derived five-factor structure, CFA was performed on Sample 2 to confirm the adequate fit of the final model. Its goodness-of-fit indices were then compared with other proposed models. Finally, the fit between latent factors of the final model was tested for construct reliability and validity.

Results

Sample 1

Single-scale analyses and internal consistency

A confirmatory measurement model was constructed using the one-factor structure with Sample 1. To construct a single-factor model, all items for each scale were included as one factor and tested for fit statistics. Through examination of the residuals and indices, one item was deleted from the self-efficacy scale and two items were deleted from the learning strategies scale. Table 3 shows the item loadings, internal consistency and fit statistical results for the single-scale analyses after item deletion. The Cronbach's α ranged from .70 to .90, indicating good internal consistency for each factor.

For all five single-factor models, the factor loadings were statistically significant, ranging from .60 to .89. In relation to the fit statistics, all NNFI and CFI were significantly acceptable, with exception to RMSEA. Several models (e.g., self-efficacy and test anxiety) did not meet the recommended value of RMSEA. As RMSEA is greatly influenced by its degrees of freedom, these single-factor models were accepted on the basis of NNFI and CFI. In addition, the patterns of residuals and factor loadings did not reveal any ambiguity which might cause model misspecification. Hence, the one-factor models (in Table 3) proceeded to the next phase of factorial analyses: combining each single-scale sequentially to achieve a complete five-factor model.

#Insert Table 3 about here#

Congeneric model of fit statistics

The second phase of the present study involved combining factor by factor and testing for statistical fit of items in each model. For each congeneric model, the parameters were estimated and scrutinized, followed by excluding the item one at a time. The aim of this congeneric method was to identify and eliminate any ambiguous loading of items. An example of such ambiguous item loading was cross-loading of items on a non-intended factor occurred. Some items could load on more than one factor, resulting in a poor model fit. Hence, these items were considered as factorially ambiguous (Markland & Ingledew, 1997).

Subsequently, the reduced model was estimated again with item(s) excluded. With each item deletion, the model's structure was evaluated after successive iterations. The resultant reduced model was intensively computed and assessed after successive iterations. Within each confirmatory measurement model, the factorial structure was generated using robust fit statistical analyses. First, the one-factor model included all items from intrinsic value. Second model comprised of all items from two factors, namely intrinsic value and self-efficacy. The third model combined intrinsic value with the previously one-factorial structure of self-efficacy (see Table 3), representing the reduced two-factor model. Subsequent factorial model was built on the previous measurement model. Ranging from a three-factor structure to the five-factor model, the final measurement model was achieved. In total, seven measurement models were formulated and their fit statistics were compared. Table 4 shows the model fit indices for each model using congeneric CFA.

#Insert Table 4 about here#

Evaluation of congeneric models using CFA

With reference to Table 4, the impact of item deletion could infer from the fit indices of two reduced models (10-item and 22-item, respectively). Using congeneric CFA, the complete measurement model was finalized to a five-factor structure with a total of 25 items.

Among five factors, there was no item deletion for the scales of intrinsic value, anxiety and lack of self-regulation. The scale of learning strategies that comprised of 10 items was the largest, followed by the scale of self-efficacy with 6 items. Firstly, the one-factor model was established on the 5-item intrinsic value and its CFA results demonstrated excellent model fit. Next, CFA results of the second model (11 items) did not reveal a good model fit, resulting in the scrutiny of items on the self-efficacy scale (described in next section).

Overall, three items were removed because of ambiguous loadings (e.g., low regression weights or poor loadings) on the self-efficacy scale and the learning strategies scale. These items did not load significantly and the model fit indices improved with deletion of items. For example, the reduced model of two-factor structure had significantly improved fit indices (NNFI and CFI) and RMSEA value, as compared to the original two-factor model. When two models were compared, the chi-square difference was 31.38, which was significant at $p < .001$ with a change of degree of freedom of 9. This indicated that the reduced model is significantly a better model fit. Likewise, the reduced model of the four-factor structure had an improved chi-square difference of 97.64, which was significant at

$p < .001$ with a change of degree of freedom of 43. Its fit indices (NNFI and CFI) increased by .14, with an acceptable value of RMSEA (close to .05). The stepwise deletion of items resulted in a five-factor model with a total of 25 items. CFA results supported the final measurement model established with Sample 1, indicating an acceptable fit between model and data ($SB\chi^2 = 447.06$, df = 265, NNFI = .93, CFI = .94, RMSEA = .048, 90% confidence interval (CI) = .040, .056). Additionally, the $SB\chi^2$/df ratio of the finalized five-factor model was 1.69, which was well below 3.0.

Deleted MSLQ items

To understand why some measurement models did not have good fit indices, there is a need to examine beyond those figures. One approach is to scrutinize the item wordings as students may perceive or comprehend the statement differently. Table 5 shows all the items on the self-efficacy scale of the MSLQ.

#Insert Table 5 about here#

An item (SE2) was removed from the original scale of self-efficacy (Liu et al., 2012; Pintrich and De Groot, 1990) due to low regression weights. A closer examination of this item's statement revealed student's expectations to out-perform his or her classmates, indicating performance goal orientation. For students who do not have such performance goal orientation, they might score a low rating for this item. Yet, this might not imply that these students have low self-efficacy. Due to this possibility, item SE2 had poor loading on the self-efficacy scale.

Table 6 presents all the items on learning strategies of the MSLQ. Two items (LS3 and LS4) were removed from the scale of learning strategies (Liu et al., 2012) due to poor loading. Items LS3 and LS4 were originally from the scale of cognitive strategy use (Pintrich & De Groot, 1990). The wordings of the item LS3 might appear ambiguous to local secondary school students as they might not understand “make everything fit together”. In local educational contexts, instead of making “everything fit together”, teachers tend to say “making connections to concepts” or “seeing the whole picture” when teaching mathematics or science. Likewise, item LS4 might not apply in the context of mathematics. Mathematics requires practice and there is no necessity to regurgitate words over and over again. Although Pintrich et al. (1991, 1993) asserted that the MSLQ scales and subscales could be used collectively or modularly according to the researcher’s needs, item LS4 might not seem to support this notion. Additionally, students might not apply such rehearsal strategy in their mathematics class.

#Insert Table 6 about here#

Sample 2

Comparison across alternative models

A confirmatory measurement model was constructed using the five-factor structure with Sample 2 (see Table 7). The results of the current five-factor model (Model A) fit the data significantly well, SBχ^2 = 370.89, df = 265, NNFI = .95, CFI = .96, RMSEA = .038, 90% confidence interval (CI) = .028, .046. The validity of Model A was further tested and compared with three alternative models, namely Models B, C and D, respectively.

#Insert Table 7 about here#

To provide the evidence that the current model was better, Model A was compared to the five-factor model (i.e., Model B) proposed by Liu et al. (2012). Examining the fit indices of Model B and subsequently compared those of Model A, their chi-square difference was 193.52, which was significant at $p < .001$ with a change of degree of freedom of 75. Conversely, a large change in SBχ^2 as compared to the difference in degrees of freedom indicates that the eliminated items constitute a real improvement in fit (Bagozzi & Yi, 1988).

Thus, Model A had significantly better fit than Model B. Moreover, the $SB\chi^2$/df ratio of Model A which was well below 3.0 was significantly lower than that of Model B. Likewise, the CFI for Model A is greater than .95 and its RMSEA is less than .05, indicating a very good model fit.

To explore the possibilities of other factorial structures of the 25-item MSLQ and their fit statistics, two alternative models were generated and tested. Model C was a four-factor model with combined grouping of intrinsic value and self-efficacy. Due to some cross-loading of items between intrinsic value and self-efficacy, these two positive constructs seemed strongly correlated. So, it was of interest to test the factorial structure of this model. Interestingly, Model C had satisfactory fit statistics, with NNFI and CFI close to .90 and RMSEA below .06.

In Model D, the negative constructs, namely test anxiety and lack of self-regulation were combined and tested with the individual three constructs (intrinsic value, self-efficacy and learning strategies). According to the aforementioned literature, test anxiety is likely associated positively with lack of self-regulation. This is why the four-factor measurement model was tested for its fit statistics. Not surprisingly, Model D had better fit than Model C but its NNFI and CFI were still comparatively lower than the recommended value of .95.

By comparing across the four models, the results validated the credibility of Model A in the present study, thus supporting the model parsimony for all factors in the 25-item MSLQ.

Specific CFA evidence of convergent validity

Factorial validity was tested using the analyses of covariance structures. Table 8 presents the CR, AVE and latent factor correlation matrix with confidence intervals. All five factors had CR and AVE values of at least .70 and close to .50, respectively. The CR values of the four factors (except for lack of self-regulation) were above .80, indicating good construct reliability. The fifth factor, lack of self-regulation revealed acceptable reliability. Convergent validity for the revised version of MSLQ (25-item) was also supported by AVE values. Discriminant validity of the latent factors provided evidence to question for latent factors, namely intrinsic value, learning strategies and lack of self-regulation. For example, correlation coefficient between intrinsic value and learning strategies was .78, with 95% CI of .56, 1.00. However, the correlation between test anxiety and lack of self-regulation exceeded an estimate of 1.00 (95% CI of .54, 1.08). The explanation for these findings are described in the next section.

#Insert Table 8 about here#

Discussion

This study evaluated the validity of the 28-item MSLQ Junior High (Liu et al., 2012) on secondary school students' motivation and cognition in academic subjects. The five-factor model comprising intrinsic value, self-efficacy, test anxiety, learning strategies and lack of self-regulation was supported by the results of Sample 1. Consequently, goodness-of-fit statistics and correlational analysis on Sample 2 supported the reliability and validity of the revised MSLQ scales. To complete the construct validity of the 25-item MSLQ, goodness-of-fit indices and convergent validity of the measurement models generated by CFA proved psychometrically sound. Besides obtaining good fit indices, the observed sample data could also sufficiently estimate the latent theoretical constructs. As having a good model fit may not be a sufficient indicator of construct validity (Wong & Lo, 2012), AVE of the model which denotes convergent validity was included. Going beyond fit indices can explicitly examine the construct validity of measurement model and ascertain that the items in the revised MSLQ actually reflect the theoretical constructs. Furthermore, comparison of the fit indices and alternative models would better represent the construct validity of the questionnaire. Overall, the psychometric study established a congeneric approach of item reduction and revised MSLQ Junior High, resulting in a parsimonious model with integrity of the original model.

In the preceding sections, statistical tests were conducted on Samples 1 and 2 to identify the following: first, item-reduction of the 28-item MSLQ; second, internal consistency and fit indices of the reduced five-factor model; third, confirmation of the model with Sample 2; fourth, comparison of fit indices with alternative models; and finally, convergent validity of the 25-item MSLQ. Collectively, the statistical findings on Samples 1 and 2 provided significant support for the psychometric properties of the 25-item MSLQ. The MSLQ scores showed excellent internal consistency of the five scales, thus providing internal structure validity evidence.

Previous empirical studies (Dunn et al., 2012; Pintrich & De Groot, 1990; Rao & Sachs, 1999) failed to offer clear guidelines in terms of condensation of construct and how an acceptable level of model reduction was achieved. A novel, congeneric approach (Markland & Ingledew, 1997) provides a clear method of item reduction through CFA. This investigation incorporates four different fit indices ($SB\chi^2$, NNFI, CFI and RMSEA) to demonstrate the influence of each item's elimination on the fit statistics. As $SB\chi^2$ difference is sensitive to sample size (Dimitrov, 2010), improved changes in the NNFI and CFI goodness-of-fit indices were reported. According to Cheung and Rensvold (2002), a decrease of .01 or larger in CFI indicates a lack of invariance. But this issue did not surface when CFI difference between the four-factor structure (reduced model) and five-factor model was .002 (see Table 4). Furthermore, the impact of item deletion was determined by the degree of fit statistics improved when a specific item was eliminated from the data set. This may provide an unambiguous psychometric method which demonstrated good stability as well as improved fit between model and data.

CFA results revealed good fit indices which supported the five-factor model with intrinsic value, self-efficacy, anxiety, learning strategies and lack of self-regulation. This differs slightly from the original model of Pintrich and De Groot (1990), yet remains consistent with the MSLQ's theoretical underpinnings. The five-factor model still retains the integrity of all constructs, with the exception for the reverse-coded learning strategies items that were renamed as "lack of self-regulation". Fit indices alone may not deliver conclusive evidence of construct validity. Hence, there is a need to examine convergent validity to buttress the result. With CR and AVE values, the construct validity of the five-factor model is considered significantly acceptable. According to Fornell and Larker (1981), AVE index which is close to or above .50, accounts for the 50% or more variance of the indicators. AVE indexes for the first three factors (intrinsic value, self-efficacy and anxiety) are close to .50. Convergent validity is considered adequate despite approximately 50% of variance is due to error. On average, each construct is significantly related to its items, thus delivering conclusive evidence of construct validity. Instead of equating those low values of AVE to a lack of convergent validity, AVE is a more conservative measure than construct reliability (Fornell & Larcker, 1981). Given the acceptable construct reliability of the five-factor model A, convergent validity can be considered as satisfactory with AVE close to .50. Taken together, these statistical findings sufficiently document an overall acceptable convergent validity of the current revised MSLQ Junior High.

Discriminant validity is determined by comparing the AVE for each factor with the squared correlation coefficents (Wong & Lo, 2012). It is robust when the "confidence intervals (±2 standard errors) around estimated correlations between two latent indicators never include 1.00" (Cegarra-Navarro, 2011, p. 34). In this case, discriminant validity of latent factors (e.g., intrinsic value and learning strategies) seemed not robust, thus indicating weak measurement invariance. Consistent to this finding, Dunn et al. (2012) also identified that the original latent structure of the MSLQ was problematic.

With reference to latent factor correlations, two factors namely self-efficacy and learning strategies were strongly related to intrinsic value. Likewise, self-efficacy was strongly related to learning strategies while test anxiety was strongly correlated to lack of self-regulation. Although the five-factor model denoted acceptable construct reliability and convergent validity, its discriminant validity may seem doubtful. However, these highly correlated factors could be explained in terms of the overlapping positive latent constructs relating to motivational beliefs or cognition. Supported by previous findings (Liu et al., 2012; Pintrich & De Groot, 1990), intrinsic value was strongly related to the use of learning strategies. In addition, latent factor correlation matrix revealed that there was a significant positive relationship between latent factors (test anxiety and lack of self-regulation), which opposed to earlier findings (Liu et al., 2012). On the other hand, Pintrich and De Groot (1990) did not find any association between test anxiety and self-regulation. Yet, there is evidence that anxiety can impede self-regulation (Zimmerman, 1989). Therefore, this study demonstrated the positive correlation between anxiety and lack of self-regulation which proved similar for university and secondary school students in previous studies described by Pekrun et al., 2002. Anxiety is also identified as negative affect that can influence components of self-regulation and may reduce one's effort put toward self-regulation (Schutz & Davis, 2000). As there has been limited empirical research on the association between test anxiety and lack of self-regulation, such conceptual relationship suggests the importance of examining students' test anxiety and how it may influence their motivation and self-regulation towards learning.

Congeneric CFA highlights a methodological contribution in refining the MSLQ Junior High. Although EFA or Rasch analysis is an approach to reduce the items of an instrument, this study applied congeneric CFA to delete the MSLQ items with Sample 1, followed by confirming the factorial structure of model with Sample 2. As mentioned previously, one-dimensional factor comprising a large number of items should be refined by item-reduction with consideration of the integrity of the construct. In this study, learning strategies factor was reduced from 10 to 8 items. The reduced five-factor structure demonstrated better fit by maintaining items that are truly working well in the model. The present approach may consider a superior model to earlier research (e.g., EFA, Rasch model) as it addresses primary criterion in structural equation modelling (SEM), resulting in a parsimonious model.

Nevertheless, there are several possible limitations that need to be acknowledged. First, this method may seem questionable to some researchers or statisticians as using CFA to delete items is not the norm. But to support the rationale of this method, Larwin and Harvey (2012) asserted that EFA is more relevant for exploring new inventories and they proposed a congeneric approach of item deletion using CFA. In view of this, the MSLQ Junior High is considered as a well-established inventory to measure students' motivational beliefs and cognition, thus congeneric CFA is considered more appropriate than EFA. Furthermore, this study posits the notion of scale refinement in the MSLQ Junior High and the current congeneric approach supports the validity of constructs. Second, CFA models assume that the items used in the study are ordered categorical. To support this, all the MSLQ's items were based on a 7-point Likert scale, which should not treated as continuous variables (Pintrich et al., 1991; Rowe, 2006). Hence, categorical items are suitable for CFA. Third, when the number of constructs increases, parameter estimation and model fit statistics may become unstable. Future research should consider improving the constructs by adding specific and relevant items, restructuring the constructs, or further refining the item wording to avoid ambiguity. Finally, this study did not test the invariance of the revised MSLQ Junior High across gender. Further study should examine the invariance of the measurement tool across gender. To sum, this study provides evidence that the congeneric CFA approach has valuable potential in refining and validating the revised MSLQ Junior High. It also extended previous research by examining the relationship between test anxiety and lack of self-regulation. Finally, applicability of the revised MSLQ's constructs on junior high students in a different cultural context (i.e., Singaporean sample) and its findings add to the existing studies.

References

Bagozzi, R. P., & Yi, Y. (1988). On the evaluation of structural equation models. *Journal of the academy of marketing science, 16*(1), 74-94.

Baumiester, R. F., & Heatherton, T. F. (1996). Self-regulation failure: An overview. *Psychological Inquiry, 7*(1), 1-15.

Bembenutty, H., McKeachie, W. J., Karabenick, S. A., & Lin, Y. G. (1998). The relationship between test anxiety and self-regulation on students' motivation and learning. Retrieved from http://files.eric.ed.gov/fulltext/ED424244.pdf

Berger, J. L., & Karabenick, S. A. (2011). Motivation and students' use of learning strategies: Evidence of unidirectional effects in mathematics classrooms. *Learning & Instruction, 21*(3), 416-428.

Bohn, K., Doll, H. A., Cooper, Z., O'Connor, M., Palmer, R. L., & Fairburn, C. G. (2008). The measurement of impairment due to eating disorder psychopathology. *Behaviour Research & Therapy*, *46*(10), 1105-1110.

Byrne, B. (2001). Structural equation modeling with AMOS, EQS, and LISREL: Comparative approaches to testing for the factorial validity of a measuring instrument. *International Journal of Testing, 1*(1), 55-86.

Cegarra-Navarro, J. G., Wensley, A. K. P., & Sánchez-Polo, M. T. (2011). Improving quality of service of home healthcare units with health information technologies. *Health Information Management Journal, 40*(2), 30-38.

Cheang, K. I. (2009). Effect of learner-centered teaching on motivation and learning strategies in a third-year pharmacotherapy course. *American Journal of Pharmaceutical Education*, *73*(3), 1-8.

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modelling, 9*(2), 233-255.

Chinn, C. A., & Malhotra, B. A. (2002). Epistemologically authentic inquiry in schools: A theoretical framework for evaluating inquiry tasks. *Science Education, 86*(2), 175-218.

Credé, M., & Phillips, L. A. (2011). A meta-analytic review of the motivated strategies for learning questionnaire. *Learning & Individual Differences*, *21*(4), 337-346.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16*(3), 297-334.

Davenport, M. (2003). Modeling motivation and learning strategy use in the classroom: An assessment of the factorial, structural, and predictive validity of the motivated strategies for learning questionnaire. *Dissertation Abstracts International Section A, 64*, 394.

Dimitrov, D. M. (2010). Testing for factorial invariance in the context of construct validation. *Measurement & Evaluation in Counseling & Development, 43*(2), 121-149.

Dunn, K. E., Lo, W-J., Mulvenon, S. W., & Sutcliffe, R. (2012). Revisiting the motivated strategies for learning questionnaire: A theoretical and statistical reevaluation of the metacognitive self-regulation and effort regulation subscales. *Educational & Psychological Measurement, 72*(2), 312-331.

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable and measurement error. *Journal of Marketing Research, 18*(1), 39-50.

Hu, L-T, & Bentler, P. M. (1999). Cut-off criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modelling, 6*(1), 1-55.

Just, M. A., & Carpenter, P. A. (1980). Theory of reading: From eye fixations to comprehension. *Psychological Review, 87*(4), 329-354.

Kaipper, M. B., Chachamovich, E., Hidalgo, M. P. L., da Silva Torres, I. L., & Caumo, W. (2010). Evaluation of the structure of Brazilian State-Trait Anxiety Inventory using a Rasch psychometric approach. *Journal of Psychosomatic Research, 68*(3), 223-233.

Kim, M., Tan, A. L., & Talaue, F. T. (2013). New vision and challenges in inquiry-based curriculum change in Singapore. *International Journal of Science Education, 35*(2), 289-311.

Kline, R. B. (2005). *Principles and practice of structural equation modelling* (2nd ed.). New York, NY: Guilford.

Larwin, K., & Harvey, M. (2012). A demonstration of a systematic item-reduction approach using structural equation modeling. *Practical Assessment, Research & Evaluation, 17*(8), 1-19.

Law, Y-K., Chan, C. K. K., & Sachs, J. (2008). Beliefs about learning, self-regulated strategies and text comprehension among Chinese children. *British Journal of Educational Psychology, 78*(1), 51-73.

Lee, J. C. K., Yin, H., & Zhang, Z. (2010). Adaptation and analysis of Motivated Strategies for Learning Questionnaire in the Chinese setting. *International Journal of Testing*, *10*(2), 149-165.

Liu, W. C., Wang, C. K. J., Koh, C., Chye, S., Chua, B. L., & Lim, B. S. C. (2012). Revised motivated strategies for learning questionnaire for secondary school student. *The International Journal of Research & Review, 8,* 19-31.

Lynch, D. J. (2006). Motivational factors, learning strategies and resource management as predictors of course grades. *College Student Journal, 40*(2), 423-426.

MacCallum, R. C., & Austin, J. T. (2000). Applications of structural equation modelling in psychological research. *Annual Review of Psychology*, *51*(1), 201-226.

Maneesriwongul, W., & Dixon, J. K. (2004). Instrument translation process: A methods review. *Journal of Advanced Nursing, 48*(2), 175-186.

Markland, D. & Ingledew, D. K. (1997). The measurement of exercise motives: Factorial validity and invariance across gender of a revised exercise motivations inventory. *British Journal of Health Psychology, 2*(4), 361-376.

Marton, F., Watkins, D., & Tang, C. (1997). Discontinuities and continuities in the experience of learning: An interview study of high-school students in Hong Kong. *Learning & Instruction*, *7*(1), 21-48.

Myung, I. J. (2003). Tutorial on maximum likelihood estimation. *Journal of Mathematical Psychology*, *47*(1), 90-100.

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York: McGraw-Hill.

Oxford Dictionaries. (2013). http://oxforddictionaries.com/definition/english/practice

Pajares, F., & Graham, L. (1999). Self-efficacy, motivation constructs, and mathematics performance of entering middle school students. *Contemporary Educational Psychology, 24*(2), 124-139.

Pekrun , R., Goetz, T., Titz, W., & Perry, R. P. (2002). Academic emotions in students' self-regulated learning and achievement: A program of qualitative and quantitative research. *Educational Psychologist, 37*(2), 91-105.

Pintrich, P. R., & De Groot, E. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology, 82*(1), 33-40.

Pintrich, P. R., Smith, D. A. F., Garcia, T., & McKeachie, W. J. (1991). *A manual for the use of the Motivated Strategies for Learning Questionnaire (MSLQ).* Ann Arbor, MI: National Center for Research to Improve Postsecondary Teaching and Learning.

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

Purdie, N., & Hattie, J. (1996). Cultural differences in the use of strategies for self-regulated learning. *American Educational Research Journal*, *33*(4), 845-871.

Rao, N., & Sachs, J. (1999). Confirmatory Factor analysis of the Chinese version of the Motivated Strategies for Learning Questionnaire. *Educational & Psychological Measurement, 59*(6), 1016-1029.

Rotgans, J., & Schmidt, H. (2009). Examination of the context-specific nature of self-regulated learning. *Educational Studies, 35*(3), 239-253.

Rotgans, J., & Schmidt, H. (2010). The motivated strategies for learning questionnaire: A measure for students' general motivational beliefs and learning strategies? *The Asia-Pacific Education Research, 19*(2), 357-369.

Rowe, K. (2006). The measurement of composite variables from multiple indicators: Applications in Quality Assurance and Accreditation Systems–Childcare. *Camberwell, Victoria: Australian Council for Educational Research*, 1-16.

Satorra, A., & Bentler, P. M. (2001). A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika*, *66*(4), 507-514.

Schutz, P. A., & Davis, H. A. (2000). Emotions and self-regulation during test taking. *Educational Psychologist, 35*(4), 243-256.

Sinclair, C., Dowson, M., & McInerney, D. (2006). Motivations to teach: Psychometric perspectives across the first semester of teacher education. *The Teachers College Record*, *108*(6), 1132-1154.

Wong, A. K-Y., & Lo E. S-C. (2012). Assessing the construct validity of the conception of teaching and learning questionnaire (CTLQ) for Chinese university students in Hong Kong: Going beyond the use of goodness of fit indices. *The Asia-Pacific Education Researcher, 21*(2), 402-413.

Yoon, C. H. (2009). Self-regulated learning and instructional factors in the scientific inquiry of scientifically gifted Korean middle school students. *Gifted Child Quarterly, 53*(3), 203-216.

Yu, C. Y., & Muthen, B. O. (2002). Evaluation of model fit indices for latent variable models with categorical and continuous outcomes (Technical Report) Los Angeles: University of California at Los Angeles. *Graduate School of Education & Information Studies.*

Zimmerman, B. J., & Schunk, D. H. (1989). *Self-regulation learning and academic achievement: Theory, research and practice*. New York: Springer-Verlag.

Zusho, A., Pintrich, P. R., & Coppola, B. (2003). Skill and will: The role of motivation and cognition in the learning of college chemistry. *International Journal of Science Education, 25*(9), 1081-1094.

Table 1

Factor analysis of past studies on MSLQ Junior High

Study	Sample (*N*, mean age)	No. of factors	Factor analysis	Factor extraction (rotation method)	Factors supported by CFA using second sample
Lee et al. (2010)	Hong Kong Chinese (*N* = 1447, 13.88)	4	CFA	Did not perform	Did not validate
Liu et al. (2012)	Singaporean (*N* = 780, 13.29)	5	EFA and CFA	Principal component analysis (varimax and oblique)	Validated
Pintrich & De Groot (1990)	American (*N* = 173, 12.5)	5	Did not state	Did not state	Did not validate
Rao & Sachs (1999)	Hong Kong Chinese (*N* = 477, 15.4)	5	CFA and unweighted least squares (ULS)	Did not perform	Did not validate

Table 2

The 28-item MSLQ

Part 1: Motivation scales		Part 2: Learning strategies scale	
Scale	No. of items	Scale	No. of items
1. Intrinsic value	5	1. Learning strategies (cognitive & meta-cognitive items)	10
2. Self-efficacy for learning & performance	6	2. Lack of self-regulation	3
3. Test anxiety	4		
Total number of items	15	Total number of items	13

Table 3

Fit indices and item loadings of each single factor following item deletion (*n* = 314)

Model (Scale/Item coding)	Loading	α	χ^2	df	χ^2/df	NNFI	CFI	RMSEA (90% CI)
Intrinsic value		.85	12.20	5	2.44	.979	.989	.068 (.018, .117)
IV1	.60							
IV2	.87							
IV3	.81							
IV4	.63							
IV5	.78							
Self-efficacy		.90	32.89	9	3.65	.945	.967	.092 (.060, .127)
SE1	.79							
SE3	.79							
SE4	.83							
SE5	.80							
SE6	.77							
Test anxiety		.87	13.02	2	6.51	.936	.979	.133 (.070, .205)
AX1	.83							
AX2	.89							
AX3	.72							
AX4	.70							

Table 3.

Continued

Model (Scale/Item coding)	Loading	α	χ^2	df	χ^2/df	NNFI	CFI	RMSEA (90% CI)
Learning strategies		.86	54.19	20	2.71	.944	.960	.075 (.051, .099)
LS1	.67							
LS2	.70							
LS5	.63							
LS6	.65							
LS7	.69							
LS8	.67							
LS9	.75							
LS10	.60							
Lack of self-regulation		.70	11.41	5	2.28	.958	.979	.065 (.010, .115)
LSR1	.61							
LSR2	.70							
LSR3	.69							

Note. IV: intrinsic value; SE: self-efficacy; AX: test anxiety; LS: learning strategies; & LSR: lack of self-regulation. All estimates are significant at $p < .001$. NNFI = Non-normed Fit Index; CFI = Comparative Fit Index; RMSEA = Root Mean Square Error of Approximation (90% confidence interval).

Table 4

Fit indices for congeneric CFA models (n = 314)

Model (Factor)	$SB\chi^2$	df	$SB\chi^2/df$	NNFI	CFI	RMSEA
One-factor (IV, 5 items)	9.53	5	1.91	.98	.99	.054 (.00, .105)
Two-factor (IV and SE, 11 items)	118.25	43	2.75	.93	.95	.075 (.059, .091)
Two-factor (Reduced model, IV and SE, 10 items)	86.87	34	2.55	.95	.96	.071 (.052, .089)
Three-factor (IV, SE and AX, 14 items)	148.13	74	2.00	.95	.96	.057 (.043, .070)
Four-factor (IV, SE, AX and LS, 24 items)	460.70	246	1.87	.92	.92	.054 (.046, .061)
Four-factor (Reduced model, IV, SE, AX and LS, 22 items)	363.06	203	1.79	.93	.94	.051 (.042, .059)
Five-factor (IV, SE, AX, LS and LSR, 25 items)	447.06	265	1.69	.93	.94	.048 (.040, .056)

Note. IV: intrinsic value; SE: self-efficacy; AX: test anxiety; LS: learning strategies; & LSR: lack of self-regulation. All estimates are significant at $p < .001$. NNFI = robust Non-normed Fit Index; CFI = robust Comparative Fit Index; RMSEA = robust Root Mean Square Error of Approximation (90% confidence interval).

Table 5

Self-efficacy items with item-deletion (e.g., for mathematics)

Item coding	Item wording
SE1	Compared with other students in this class, I think I know a great deal about mathematics.
SE2 (deleted item)	Compared with other students in this class, I expect to do well.
SE3	I am sure I can do an excellent job on the problems and tasks assigned for mathematics.
SE4	I think I will receive a good grade in mathematics.
SE5	My study skills are excellent compared with others in this class.
SE6	Compared with others in this class, I think I’m a good student.

Table 6

Learning strategies items with item-deletion (e.g., for mathematics)

Item coding	Item wording
LS1	When I study for mathematics test, I practise saying the important facts over and over to myself.
LS2	I use what I have learned from old homework assignments and the textbook to do new assignments.
LS3 (deleted)	When I am studying a topic, I try to make everything fit together.
LS4 (deleted)	When I read materials for mathematics class, I say the words over and over to myself to help me remember.
LS5	I outline the chapters in my book to help me study.
LS6	When reading, I try to connect the things I am reading about with what I already know.
LS7	I ask myself questions to make sure I know the material I have been studying.
LS8	Even when study materials are dull and uninteresting, I keep working until I finish.
LS9	Before I begin studying, I think about the things I will need to do to learn.
LS10	When reading, I stop once in a while and go over what I have read.

Table 7

Fit indices for CFA models (n = 296)

Model	$SB\chi^2$	df	$SB\chi^2/df$	NNFI	CFI	RMSEA (90% CI)
Model A Five-factor	370.89	265	1.40	.949	.955	.038 (.028, .046)
Model B Five-factor (Liu et al., 2012)	564.41	340	1.66	.912	.921	.048 (.042, .055)
Model C Four-factor	516.05	269	1.92	.883	.895	.057 (.050, .064)
Model D Four-factor	461.71	269	1.72	.909	.918	.050 (.042, .058)

Note. All estimates are significant at $p < .001$. NNFI = robust Non-normed Fit Index; CFI = robust Comparative Fit Index; RMSEA = robust Root Mean Square Error of Approximation (90% confidence interval).

Table 8

Reliability, validity and latent factor correlations for Model A (*n* = 296)

Scale	CR	AVE	IV	SE	AX	LS
Intrinsic value, IV	.87	.47				
Self-efficacy, SE	.88	.48	.75*			
			(.10)			
			.55, .95			
Test anxiety, AX	.84	.47	-.01*	-.21*		
			(.08)	(.09)		
			-.17, .15	-.39, .03		
Learning strategies, LS	.86	.50	.78*	.75*	.14*	
			(.11)	(.11)	(.09)	
			.56, 1.00	.53, .97	-.04, .32	
Lack of self-regulation	.70	.50	-.36*	-.26*	.81*	- .21*
			(.09)	(.09)	(.14)	(.09)
			-.54, -.18	-.44, .08	.54, 1.08	-.39, .03

Note. * $p < .05$. In each cell, first row = latent factor correlation, second row = SE of latent correlation coefficient, last row = correlation confidence intervals within plus/minus 2 SE.