British Journal of Educational Psychology (2000), **70**, 405–418 *Printed in Great Britain*

Item-level and subscale-level factoring of Biggs' Learning Process Questionnaire (LPQ) in a mainland Chinese sample

John Sachs
University of Hong Kong

Lingbiao Gao*
Institute of Curriculum Studies and Teaching Material Development, South China Normal University

Background. The learning process questionnaire (LPQ) has been the source of intensive cross-cultural study. However, an item-level factor analysis of all the LPQ items simultaneously has never been reported. Rather, items within each subscale have been factor analysed to establish subscale unidimensionality and justify the use of composite subscale scores.

Aims. It was of major interest to see if the six logically constructed items groups of the LPQ would be supported by empirical evidence. Additionally, it was of interest to compare the consistency of the reliability and correlational structure of the LPQ subscales in our study with those of previous cross-cultural studies.

Methods. Confirmatory factor analysis was used to fit the six-factor item level model and to fit five representative subscale level factor models.

Sample. A total of 1070 students between the ages of 15 to 18 years was drawn from a representative selection of 29 classes from within 15 secondary schools in Guangzhou, China. Males and females were almost equally represented.

Results. The six-factor item level model of the LPQ seemed to fit reasonably well, thus supporting the six dimensional structure of the LPQ and justifying the use of composite subscale scores for each LPQ dimension. However, the reliability of many of these subscales was low. Furthermore, only two subscale-level factor models showed marginally acceptable fit. Substantive considerations supported an oblique three-factor model.

Conclusions. Because the LPQ subscales often show low internal consistency reliability, experimental and correlational studies that have used these subscales as dependent measures have been disappointing. It is suggested that some LPQ items should be revised and other items added to improve the inventory's overall psychometric properties.

*Requests for reprints should be addressed to Dr. Lingbiao Gao, Institute of Curriculum Studies and Teaching Material Development, South China Normal University, China.

The learning process questionnaire or LPQ (Biggs, 1987) is a self-report inventory designed to measure three approaches to learning – surface, deep, and achieving – in secondary students. The three approaches reflect the extrinsic, intrinsic, and achieving motivation, respectively, of the learner (Biggs, 1992, 1995). Each approach to learning is defined by a question related to the student's motivation for learning and the strategy that the student employs to achieve it (Biggs, 1992, 1995). Consequently, the LPQ measures six dimensions of student approaches to learning (Biggs, 1992): surface motive (SM), surface strategy (SS); deep motive (DM), deep strategy (DS); achieving motive (AM), achieving strategy (AS).

Since identifying the approaches students use in learning is of considerable practical educational importance in both improving student learning and improving the curriculum, the LPQ has been the source of intense cross-cultural research from its inception (Andrews, Violato, Rabb, & Hollingsworth, 1994; Biggs, 1987, 1991, 1992; Kember & Leung, 1998; Watkins & Akande, 1994; Watkins, Akande, & Mpofu, 1994; Watkins, Dong, & Xia, 1995; Watkins & Mpofu, 1997; Wong, Lin & Watkins, 1996). All these studies reported on the reliability of the six LPQ subscales and on their factor structure. None of these studies, however, reported an item-level factoring of the LPQ except for Biggs (1992, 1993). In those cases, the item-level factoring was done separately for the items comprising each subscale and not simultaneously for all the LPQ items. It seems reasonable, therefore, to enquire if the logically constructed item groups of the LPQ (Biggs, 1992) can be confirmed with empirical data before subscale-level factoring is attempted.

Inventories of student learning such as the LPQ, the slightly longer version, the study process questionnaire or SPQ (Biggs, 1992) and Entwistle, Hanley, and Hounsell's (1979) approaches to study inventory (ASI; Entwistle & Ramsden, 1983; Ramsden & Entwistle, 1981), including the revised version (RSAI; Duff, 1997; Entwistle & Tait, 1995), are commonly referred to as student approaches to learning (SAL) inventories because they emphasise the context within which student learning takes place (Biggs, 1993; Entwistle & Waterston, 1988). This is in contrast to inventories of student learning such as Schmeck, Ribich, and Ramanaiah's inventory of learning processes (1977; Schmeck & Grove, 1979; Schmeck, Geisler-Brenstein, & Cercy, 1991) that place greater emphasis on the cognitive aspects of learning such as information processing.

Both the SPQ and ASI are appropriate for tertiary level students and measure similar constructs (Biggs, 1993; Entwistle & Ramsden, 1983; Entwistle & Waterston, 1988; Harper & Kember, 1989; Kember & Leung, 1998; Watkins, 1996). Watkins (1996) notes that the adoption of the deep/surface terminology for constructs measured by these inventories owes much to the early qualitative work of Marton and Säljö (1976).

Since the underlying theory of SAL as assessed by the LPQ has been described at length in Biggs (1987, 1992, 1993, 1995, 1996), Kember and Leung (1998), and Watkins (1996), it will not be repeated here. The interested reader is referred to the cited references. The emphasis in this paper will be on the psychometric properties of the LPQ within a cross-cultural context. In that regard, the purpose of this paper is threefold. First, the reliability of the six LPQ subscales will be assessed and compared with those reported in previous studies such as Wong *et al.* (1996). Second, item-level factoring will be conducted to determine if the LPQ does in fact measure six dimensions of student learning as proposed by Biggs (1987, 1992, 1993, 1995). Third, various theoretically derived subscale-level factor models will be fit and compared with relevant ones reported in previous studies such as Andrews *et al.* (1994), Wong *et al.* (1996), and Kember and Leung (1998). Important cross-cultural differences will be discussed.

Method

Sample

A sample of 1070 secondary students was drawn from 29 classes within 15 secondary schools in Guangzhou, China. The sample was comprised of 557 males and 459 females, with 54 students failing to indicate their gender. Student ages ranged from 15 to 18 years.

Instrument

The LPQ is a self-report inventory consisting of 36 statements about student approaches to learning. The scale is usually administered to secondary school students. Students are asked to respond to each statement by indicating how true that statement is of them using on a 5-point Likert scale: 1-*never* or *only rarely* true of me, 2-*sometimes* true of me, 3-true of me about *half the time*, 4-*frequently* true of me, 5-*always* or *almost always* true of me. These 36 statements are grouped into six subscales with six statements (items) per subscale. The six subscales are measures of the six dimensions of student approaches to learning proposed by Biggs (1987, 1992, 1993, 1995): surface motive (SM), surface strategy (SS); deep motive (DM), deep strategy (DS); and achieving motive (AM), achieving strategy (AS).

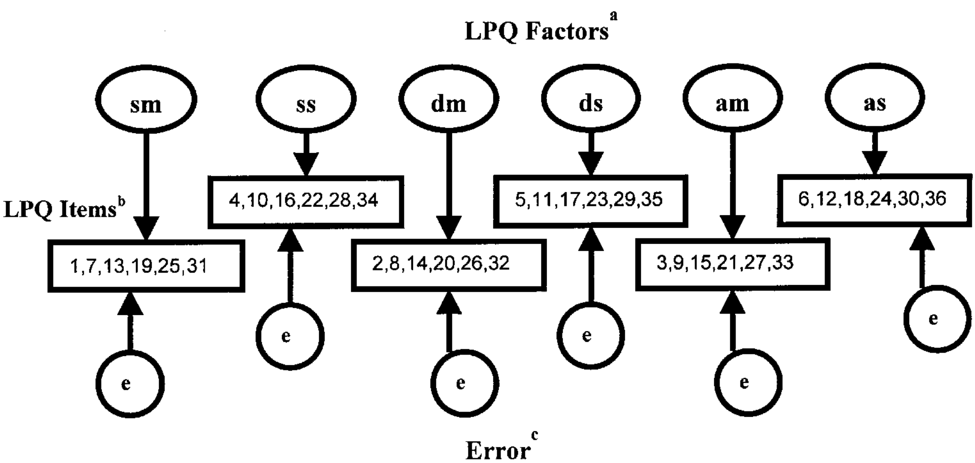
The LPQ data used in this study were part of a larger study conducted by Gao (1999) on 'conceptions of teaching held by school physics teachers in Guandong, China and their relations to student learning'. Therefore, where necessary, some LPQ statements (items) were slightly reworded so that their reference to learning a subject in general was focused on physics instead. For example, LPQ item 14, 'I find that many subjects can become very interesting once you get into them', would be reworded to read 'I find that physics can become very interesting once you get into it'. But LPQ item 18, 'I try to do all of my assignments as soon as they are given to me' would remain unchanged. This modification to some LPQ statements is only minor and is unlikely to invalidate comparisons of our results with results from studies that used an unmodified version of the LPQ.

Once the necessary subject specific changes were made to some of the LPQ statements, experts in English to Chinese translation were used to translate all 36 LPQ statements into Chinese. Care was taken to ensure that each translated statement retained as similar a meaning as possible to the original.

Design

There were three parts to this study. First, internal-consistency reliability estimates along with other descriptive measures were obtained for each LPQ subscale. Second, Biggs' six-factor model was fit to all 36 LPQ items. We decided to use the correlation matrix for this item-level confirmatory factor analysis because the categorical 5-point Likert scale used to measure the item-level responses is clearly arbitrary and because the interest in this analysis centred on the relationships between the 36 LPQ items. The use of the correlation matrix instead of the covariance matrix for fitting the confirmatory factor model was also acceptable since no item loaded on more than one factor in this six-factor model (see Figure 1), and each factor was standardised (Cudeck, 1989).

Figure 1. Simplified schematic of the six-factor item-level model of the LPQ



[a] All LPQ factors were free to correlate with each other.

[b] List of items defining each LPQ factor.

[c] There is a unique latent error term associated with each LPQ item.

We also decided to use unweighed least squares (ULS) estimation as opposed to the more common maximum likelihood (ML) estimation because item-level data of this type tends to be skewed and heterogeneous and, therefore, is not multivariate normal. The use of ML estimation with such non-normal categorical data has been shown to substantially inflate the chi-square statistic (Babakus, Ferguson, & Jöreskog, 1987). Even with ULS estimation, the chi-square will be inflated. Therefore, overall (or absolute) model fit was evaluated with the root mean square residual (RMR) and the goodness of fit index (GFI). Values of 0.05 or less for the RMR and values of 0.90 or more for the GFI indicate acceptable model fit. Although the RMR tends to be somewhat inflated with non-normal categorical data (Babakus *et al.*, 1987), the GFI is robust against non-normality and is independent of sample size (Jöreskog & Sörbom, 1981). But neither the RMR nor the GFI is independent of the number of model parameters estimated, and the RMR is biased in favour of ULS estimation by definition. To obviate these problems, we also reported the adjusted goodness of fit index (AGFI) which does take into account the number of estimated model parameters and Steiger's (1990) root mean square error approximation (RMSEA), which is based on the noncentrality parameter (McDonald & Marsh, 1990). The RMSEA adjusts for both the sample size and the model degrees of freedom. Values in the range of 0.90 to 1 for the AGFI and values in the range of 0 to 0.08 for the RMSEA indicate acceptable model fit.

Third, confirmatory factor analysis of the LPQ subscales was used to fit the five models shown in Figure 2. These five models were chosen because they were the most interesting substantive models tested in previous research (Andrews *et al.*, 1994; Kember & Leung, 1998; Wong *et al.*, 1996). The first four models were two-factor models while the last was a three-factor model.

Two-factor models

Model 1: A Motive factor (M) and a Strategy factor (S) (Wong *et al.*, 1996).

Model 2: A Surface plus Achieving (S+ A) factor and a Deep (D) factor (Wong *et al.*, 1996)

Model 3: A Surface (S) factor and an Achieving plus Deep (A+ D) factor (Wong *et al.*, 1996). Kember and Leung (1998) also tested this model. However, they referred to the surface factor as a 'reproducing' (RP) factor and the achieving plus deep factor as a 'meaning' (ME) factor.

Model 4: This model is a combination of models 2 and 3. It has an (S+ A) factor and an (A+ D) factor, which means that the LPQ subscales AS and AM crossload on these two factors (Kember & Leung, 1998).

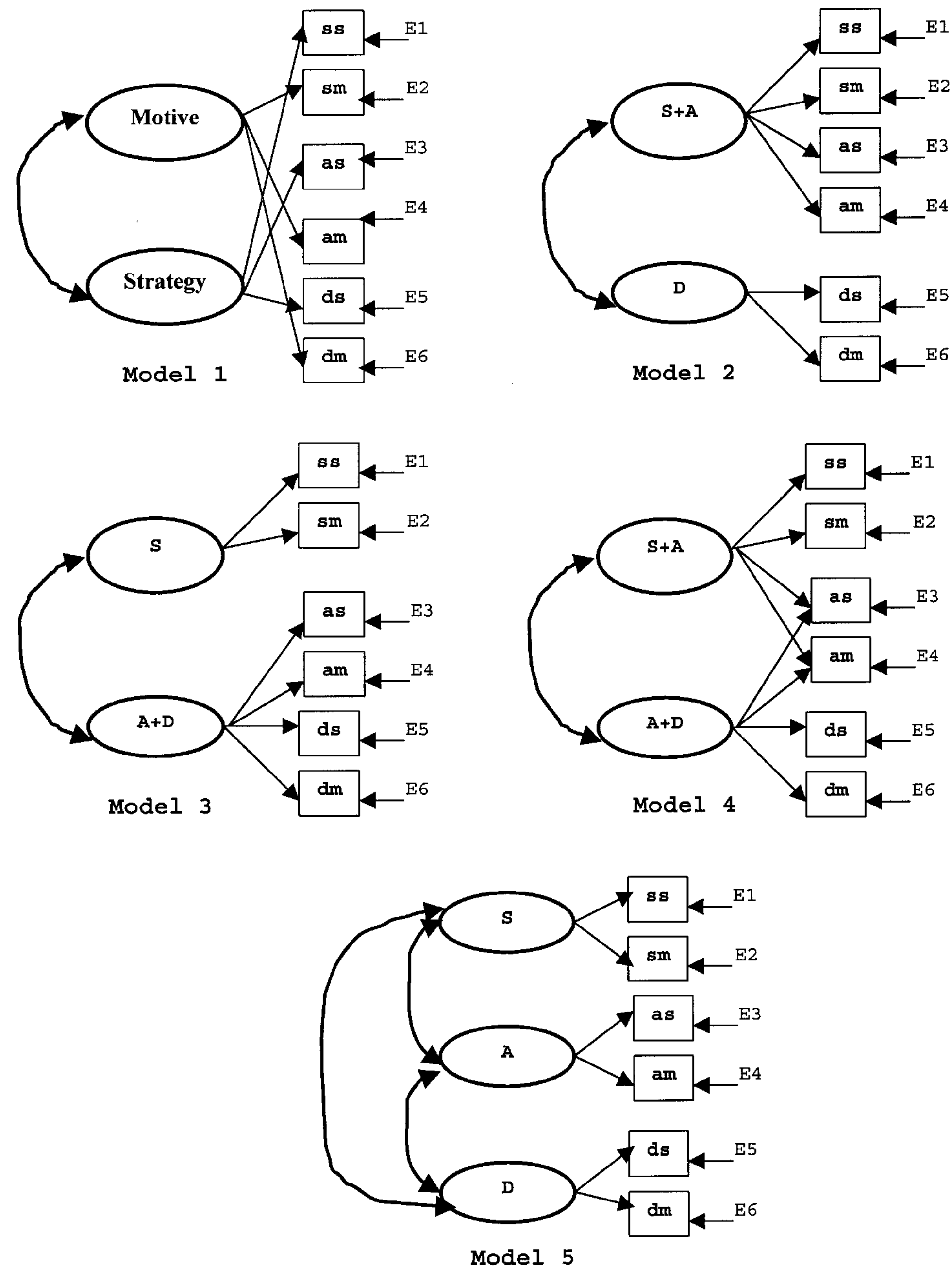
Three-factor model

Model 5: The three factors are Surface (S: subscales SS, SM), Achieving (A: subscales AS, AM), and Deep (D: subscales DS, DM) (Andrews *et al.*, 1994; Biggs, 1992; Kember & Leung, 1998; Wong *et al.*, 1996). These reflect the three approaches to student learning as measured by the surface, achieving, and deep scales of the LPQ (Biggs, 1992, p. 45).

Since subscale scores are composite measures, these five factor models were fitted to the covariance matrix using the generalised least squares (GLS) procedure in LISREL 8 (Jöreskog & Sörbom, 1993). We used the GLS procedure because the subscale score distributions were not normal, and GLS, unlike ML, yields an approximate chi square (χ^2) statistic even when the data are not multivariate normal (Browne, 1977). In addition to the conventional χ^2 statistic, we also assessed model fit with a number of other fit indices.

To determine overall or absolute model fit, we used Steiger's (1990) RMSEA, which was described earlier. For assessing incremental model fit, a comparison of the fitted model to a base model, we used the Tucker Lewis index (TLI) (Tucker & Lewis, 1973), where the base model is the null model (i.e., a model of no relationship). The TLI is also called the nonnormed fit index (NNFI). Conventionally, values of 0.90 to 1 for the TLI are indicative of a well-fitted model (Marsh, Balla, & McDonald, 1988; McDonald & Marsh, 1990), with values in the open interval 0.80 to 0.89 indicating marginal fit. To adjust for the number of parameters estimated by each of our models, we used the adjusted goodness-of-fit index (AGFI) described earlier. Finally, since Kember and Leung (1998) reported the comparative fit index (CFI), we included it here too to facilitate comparisons with that work. The CFI belongs to the same class of incremental fit indices as the TLI. Values of the CFI lie between 0 and 1, with values close to 1 indicating better fit (Jöreskog & Sörbom, 1993).

Figure 2. Five subscale level LPQ factor models



Results

Descriptive measures and reliability

The estimates of internal consistency reliability (coefficient alpha) reported in Table 1 tended to be in line with those reported in previous studies (Andrews *et al.*, 1994; Kember & Leung, 1998; Wong *et al.*, 1996). Nevertheless, there is considerable variability between the internal consistency reliability estimates of the LPQ subscales from one study to the next. Some of this variability no doubt reflects cross-cultural differences. However, direct comparison of reliability estimates from one study to the next is generally not advisable since the magnitude of the reliability coefficient is dependent on the sample variance (Stanley, 1971). A better choice of comparison would be the standard error of measure (SEM), which tends to be invariant of the sample's variance (Nunnally, 1978). For example, when this is done for the LPQ subscale means of the Andrews *et al.* (1994) data and the Kember and Leung (1998) data, we see that the SEMs for the subscale means in both these studies are very similar to those reported in the current study (see Table 1). Therefore, any observed differences in subscale reliability reflects greater true score variability (i.e., greater group heterogeneity) in the study that reports the larger subscale reliability.

Table 1. Reliability and descriptive measures for the LPQ subscales

					SEM[a]		
Subscale	*M*	SD	Skew	Reliability	Current[b]	Andrews[c]	Kember[d]
Surface Motive (SM)	3.20	.69	−.104	.53	.47	.49	.46
Deep Motive (DM)	3.54	.64	−.136	.50	.45	.40	.43
Achieving Motive (AM)	3.52	.64	−.379	.51	.45	.43	.44
Surface Strategy (SS)	2.49	.61	.291	.41	.47	.39	.46
Deep Strategy (DS)	3.18	.67	.073	.62	.41	.44	.39
Achieving Strategy (AS)	3.39	.67	−.215	.61	.43	.41	.40

[a] Standard error of the measure ($SEM_x = SD_x\sqrt{1-REL_x}$). All SEMs were computed using the mean subscale score.
[b] SEMs for current study.
[c] SEMs computed from Andrews *et al.* (1994) data.
[d] SEMs computed from Kember and Leung (1998) data.

The LPQ subscale means and subscale skew are also reported in Table 1. Inspection of the means shows that respondents tended to use the centre of the item scale when responding with a slight upward bias between '3-true of me *half the* time' and '4-*frequently* true of me'. This was true for all the subscales except SS. Most of the LPQ subscale distributions were negatively skewed except for subscales SS and DS. Also subscale SS had the lowest reliability (coefficient alpha = 0.41). Internal consistency reliability estimates as low as 0.32 have been reported for this LPQ subscale (see Wong, *et al.*, 1996).

Item-level factoring
Biggs' six-factor item-level model fits the data reasonably well. The RMSEA and the RMR were 0.048 and 0.053, respectively. Even though the RMR is a little above 0.05, this is acceptable because of the tendency of the RMR to become slightly inflated with non-normal data. Both the GFI and AGFI were 0.93 and 0.92, respectively. These values are well above the 0.90 value indicative of acceptable model fit. The ULS coefficients (loadings) of the items on their respective factors are shown in Table 2. Most items had factor loadings of 0.30 or above on their respective factors. The remaining items had loadings above 0.20 on their factors, while three items (16, 21, and 25) had factor loadings below 0.20 but above 0.10. The only really problematic item was item 10, which had a zero loading on its factor. Item 10: 'I find that the best way to learn physics is to memorise by heart', belongs to the SS subscale, and its corrected item-total correlation with that subscale was only .08. As was noted earlier, the SS subscale also had the lowest reliability estimate of all six subscales. On these grounds, we would be justified in deleting this item. This item, however, has not been shown to be problematic in previous research reported by Biggs (1992, 1993), although in a factor analysis of this subscale, Biggs (1993) did report obtaining a second eigenvalue of almost 1 (i.e., 0.99), suggesting that a second factor might be present. But item level data of this type are inherently noisy (Nunnally, 1978), often resulting in spurious factors especially when item distributions are heterogeneous (Bernstein & Teng, 1989). Furthermore, Biggs (1993) did not identify any SS items associated with this putative second factor. Therefore, the poor psychometric properties of item 10 could reflect either a characteristic peculiar to our sample or it could reflect some real cross-cultural difference. The latter is a distinct possibility since work by Marton, Dall'Alba, and Tse (1996) suggests that memorising for many Asians is perceived as being similar to understanding. But because the reliability coefficient increased by only a small amount (0.02) when this item was deleted, and because deleting this item would make it difficult to compare our results with previous work, we decided to retain it when computing the SS subscale scores.

Subscale-level factoring
While fitting our five subscale-level factor models, it was found that the error variance for the SS subscale was always negative. Rather than setting its error variance to zero, we fixed its value to a computed estimate of its error variance by multiplying its observed score variance by one minus its reliability. Therefore, the reported results for all models have one degree of freedom less than would be expected.

Inspection of the results in Table 3 shows that none of the models tested fit particularly well. If a choice had to be made between these models, then models 4 and 5 are to be preferred since they show the best fit. But since both these models fit equally well, a choice between them has to be made on purely substantive grounds. Therefore, model 5 might be preferred since it corresponds to the relationship between the three approaches to learning (surface, deep, and achieving) and the subscale type (motive or strategy) outlined by Biggs (1992, p. 45).

Although the model fit results are not particularly good, they do agree with results summarised in Wong *et al.* (1996) for a Hong Kong (HK1) and a Beijing (BEIJ) sample. Also, the fit of model 4 was similar to that reported by Kember and Leung (1998) for

Table 2. Standardised ULS coefficients (loadings) for Biggs' six-factor LPQ model on 36 LPQ items

	LPQ FACTORS					
Item No.	SM	DM	AM	SS	DS	AS
1	0.20					
2		0.26				
3			0.48			
4				0.45		
5					0.21	
6						0.33
7	0.45					
8		0.42				
9			0.58			
10				0.00		
11					0.57	
12						0.58
13	0.56					
14		0.42				
15			0.25			
16				0.11		
17					0.60	
18						0.45
19	0.43					
20		0.46				
21			0.12			
22				0.52		
23					0.59	
24						0.51
25	0.19					
26		0.32				
27			0.44			
28				0.32		
29					0.44	
30						0.48
31	0.48					
32		0.45				
33			0.59			
34				0.35		
35					0.39	
36						0.41

their model 7. Using the Tucker-Lewis index (TLI) and the adjusted goodness-of-fit index (AGFI), the fit of models 4 and 5 can best be described as marginal. The fit of models 4 and 5 could have been improved somewhat by allowing various subscale error terms to correlate. This practice, however, seems unjustified without strong substantive reasons to support it.

Table 3 Summary of fit for LPQ subscale-level factor models

Model	Degrees of Freedom and Fit Indices					
	d.f.	χ^2	RMSEA[a]	TLI[b]	AGFI[c]	CFI[d]
1. M – S	9	210.99	0.14	0.77	0.85	0.86
2. S+ A – D	9	202.61	0.14	0.78	0.85	0.87
3. S – A+ D	9	185.64	0.14	0.80	0.86	0.88
4. S+ A – A+ D	7	115.83	0.12	0.84	0.89	0.92
5. S – A – D	7	115.83	0.12	0.84	0.89	0.92

[a] Root mean square error approximation.
[b] Tucker Lewis index or nonnormed fit index.
[c] Adjusted goodness-of-fit index.
[d] Comparative fit index.

Discussion

An obvious problem with the LPQ is that its subscales are often reported as having low reliability, usually less than 0.70 and often even less than 0.50. There are of course exceptions. Wong *et al.* (1996) reported LPQ subscale reliability estimates for a Hong Kong international school (HK-I) sample that were mostly greater than 0.80, and Andrews *et al.* (1994) reported four reliability estimates of greater than or equal to 0.70 for their Canadian data. The very low reliability of the SS subscale in our study (coefficient alpha = 0.41) is particularly noteworthy since both the Andrews *et al.* (1994) and the Kember and Leung (1998) studies reported reliability estimates of less than 0.50 for this subscale, too. Referring to Biggs (1993), Kember and Leung attribute this low reliability to the fact that this subscale contains two components. As discussed earlier, however, Biggs' (1993) evidence for a second factor is not compelling.

The poor reliability of the LPQ subscales no doubt explains the disappointing results associated with their use as dependent measures in a number of studies. For example, Lai (1991) introduced cooperative learning in one class of secondary 6 (S6) geography, while another class of S6 geography was taught in the conventional manner. After eight weeks she found that the cooperative learning group performed better in geography, but no differences in LPQ subscale scores were found even though differences in strategies were reported in interviews with students from both groups. Leung (1992) applied the small group 'learning how to learn' approach with an S4 economics and public affairs (EPA) class. A significant increase in EPA test performance was observed while no significant differences in the LPQ subscale scores were observed. Again, students did report using different learning strategies during the time the new teaching approach was used. Gao (1999) categorised physics teachers into three groups in terms of their conceptions of teaching. It was expected that students being taught by teachers using a 'developmental' conception of teaching over the year would adopt a more deep learning approach, and those students who were taught by teachers using a 'knowledge delivery' conception of teaching would adopt a more surface learning approach. The third group of teachers adopted a 'conflict' conception of teaching, which was not expected to impact on their students' learning approach. However, when LPQ subscale scores were used as predictors of students' learning approach, no significant relationship was found between teachers' conceptions of teaching and changes in students' approaches to learning even though significant relations between teachers' conceptions of teaching and students' learning outcomes were identified.

Watkins (1996), in summarising the results from a number of studies on LPQ correlates with achievement, reports mean correlations (over 16 studies) of only −0.11, 0.20, and 0.19 between measures of academic achievement and the surface, deep, and achieving approachs, respectively. Citing Fraser, Walberger, Welch, and Hattie's (1987) extensive synthesis of educational research on academic correlates, Watkins (1996) notes that correlations of 0.20 are worth pursuing. Fraser *et al.* cite Rosenthal and Rubin's (1982) use of the binomial effect size to show that a correlation as small as 0.20 is associated with a 20% improvement in the rate of learning. Looked at in this way, correlations of 0.20 with any of the LPQ subscales are, therefore, of considerable substantive importance. But many other correlates of academic achievement reported in Fraser *et al.* are much higher than 0.20. For example, the average correlation of student characteristics was found to be 0.40. This correlation translates into a binomial effect size of 40%, double that associated with a correlation of 0.20. Surely, one could consider a student's approach to learning to be a student characteristic; therefore, expectations of larger correlations with academic achievement seem reasonable. If the reliability of the LPQ subscales were improved, higher correlations with standardised measures of academic achievement would be expected, which in turn would translate into greater power and larger effect sizes in studies using the LPQ subscales as dependent measures.

Improved reliability of the LPQ subscales would also improve item level factoring results. As it stands now, the unidimensionality of some LPQ subscales – SM, DM, and SS – is questionable. Using a composite score for such subscales makes little sense. For example, Biggs' (1996) principal components analysis of the SM subscale would lead one to construct two subscales: a two-item subscale (items 7 and 19) and a four-item subscale (items 1, 13, 21, 31). Yet a two- or even a four-item subscale hardly seems adequate for any serious enquiry into student approaches to learning. A major problem with item level factoring, as mentioned earlier, is that items are inherently noisy because responses are categorised and because item response distributions are often very heterogeneous, even for items measuring a similar concept. Heterogeneity of response distributions for items measuring a similar concept will attenuate their correlation. This was certainly true of our observations of the item distributions within each LPQ subscale and could account for the low measures of internal consistency reliability (alpha) reported for many of the subscales.

Subscale item heterogeneity probably also accounts for the fact that not only the magnitude but also the sign of the covariances (correlations) between the LPQ subscales can vary greatly from one study to the next. For example, the DM and SM subscales show a significant negative relationship in Andrews *et al.* (1994) and a significant positive relationship in the current study (see Appendix). One could argue that this reflects some cultural difference. If this were true, then one would expect to see greater agreement between the covariance matrix of this study and one based on Hong Kong data such as Kember and Leung's (1998) study, but the opposite is the case.

From a psychometric viewpoint, the problems outlined above represent some shortcomings of the LPQ and would suggest the need for improvement. An obvious first step would involve the rewriting and or removal of items that have been shown to be problematic in the past. Another possibility would be the addition of one or two items to each subscale to improve the overall reliability of the subscales. A slightly larger item pool would have the added advantage of making it possible to drop a subscale item if it were found to function poorly within a particular cultural context without adversely affecting the subscale's validity.

One could argue, however, the shortcomings of the LPQ discussed here, especially those not reported in other studies, are more indicative of sampling differences than of any inherent psychometric deficits. For example, the current sample consists of mainland secondary school science students. Science students are usually more quantitatively oriented than arts students. Furthermore, the mainland school system is different from that in Western countries and even from that in Hong Kong. These sample specific characteristics need to be kept in mind when making any generalisations about the psychometric properties of the LPQ, especially when these generalisations would suggest changes to the inventory's item content.

Although student approaches to learning (SAL) inventories such as the LPQ emphasise the context within which learning takes place, this does not mean that a student's generally preferred approach to learning is dependent on a particular subject that students might happen to be studying. If this were not the case, then inventories of SAL would be of little utility in identifying preferred approaches to learning. So differences associated with the subject learned are usually not an issue. The second sampling issue relates to the characteristics of the mainland school system and what impact this might have had on the results.

The teaching environment in both the mainland and Hong Kong is characterised by a formal teaching style that is largely curriculum bound and examination oriented, especially when compared to the West. The situation in the mainland is perhaps even more examination oriented and curriculum bound than that in Hong Kong. However, Biggs and Watkins (1996) argue that even with basic differences in Hong Kong and mainland China, the 'Chinese Learner' is essentially a product of a 'Confucian-heritage' (CH) culture and, therefore, shares important similarities. At least as far as the LPQ is concerned, this argument is supported by the Wong *et al.* (1996) study which found basic similarities in LPQ confirmatory factor analysis results for Hong Kong and Beijing samples. More importantly, they found that LPQ confirmatory factor analysis results were similar for Western, Asian, and African samples.

In conclusion, although evidence for the LPQ's cross-cultural validity is generally positive, the low reliability associated with of some of its subscales and their corresponding low correlations with measures of academic achievement needs to be addressed. This is especially true if the inventory's continued use is to prove useful in experimental and correlational studies on student approaches to learning.

References

Andrews, J., Violato, C., Rabb, K., & Hollingsworth, M. (1994). A validity study of Biggs' three-factor model of learning approaches: A confirmatory factor analysis employing a Canadian sample. *British Journal of Educational Psychology*, *64*, 179–185.

Babakus, C.E., Ferguson, J.R., & Jöreskog, K.G. (1987). The sensitivity of confirmatory

maximum likelihood factor analysis to violations of measurement scale and distributional assumptions. *Journal of Marketing Research*, *24*, 222–228.
Bernstein, I.H., & Teng, G. (1989). Factoring items and factoring scales are different: Spurious evidence for multidimensionality due to item categorization. *Psychological Bulletin*, *105* (3), 467–477.
Biggs, J.B. (1987). *Student approaches to learning and studying.* Hawthorn, Vic.: Australian Council for Educational Research.
Biggs, J.B. (1991). Approaches to learning in secondary and tertiary students in Hong Kong: Some comparative studies. *Educational Research Journal*, *6*, 27–39.
Biggs, J.B. (1992). *Why and how do Hong Kong students learn? Using the learning and study process questionnaires.* Education Paper 14, Faculty of Education, University of Hong Kong.
Biggs, J.B. (1993). What do inventories of students' learning processes really measure? A theoretical review and clarification. *British Journal of Educational Psychology*, *63*, 3–19.
Biggs, J.B. (1995). Learning in the classroom. In J.B. Biggs & D.A. Watkins (Eds.), *Classroom learning* (pp. 147–166). New York: Prentice Hall.
Biggs, J.B. (1996). Western misperceptions of the Confucian-heritage learning culture. In D.A. Watkins & J.B. Biggs (Eds.), *The Chinese learner: Cultural psychology and contextual influences* (pp. 45–68). Hong Kong: CERC and ACER.
Biggs, J.B., & Watkins, D.A. (1996). The Chinese learner in retrospect. In D.A. Watkins & J.B. Biggs (Eds.), *The Chinese learner: Cultural psychology and contextual influences* (pp. 269–285). Hong Kong: CERC and ACER.
Browne, M.W. (1977). Generalized least-squares estimators in the analysis of covariance structures. In D.J. Aigner & A.S. Goldberger (Eds.), *Latent variables in socio-economic models* (pp. 205–266). Amsterdam: North-Holland.
Cudeck, R. (1989). The analysis of correlation matrices using covariance structure models. *Psychological Bulletin*, *105*, 317–327.
Entwistle, N.J., & Ramsden, P. (1983). *Understanding student learning.* London: Croom Helm.
Entwistle, N.J., & Tait, H. (1995). *The Revised Approaches to Study Inventory.* Edinburgh: Centre for Research on Learning and Instruction, University of Edinburgh.
Entwistle, N.J., & Waterston, S. (1988). Approaches to studying and levels of processing in university students. *British Journal of Educational Psychology*, *58*, 258–265.
Entwistle, N.J., Hanley, M., & Hounsell, D. (1979). Identifying distinctive approaches to studying. *Higher Education*, *8*, 365–380.
Duff, A. (1997). A note on the reliability and validity of a 30-item version of Entwistle & Tait's Revised Approaches to Study Inventory. *British Journal of Educational Psychology*, *67*, 529–539.
Fraser, B.J., Walberg, H.J., Welch, W.W., & Hattie, J.A. (1987). Synthesis of educational productivity research. *International Journal of Educational Research*, *11*, 145–252.
Gao, L. (1999). *Conceptions of teaching held by school physics teachers in Guandong China and their relations to student learning.* Unpublished doctoral dissertation, University of Hong Kong, SAR, China.
Harper, G., & Kember, D. (1989). Interpretation of factor analyses from the approaches to study inventory. *British Journal of Psychology*, *59*, 66–74.
Jöreskog, K.G., & Sörbom, D. (1981). *LISREL versions V and VII user's guide* (1st ed.). Chicago, IL: International Educational Services.
Jöreskog, K.G., & Sörbom, D. (1993). *LISREL 8: Structural equation modeling with SIMPLIS command language.* Chicago, IL: Scientific Software, Inc.
Kember, D., & Leung, D.Y.P. (1998). The dimensionality of approaches to learning: An investigation with confirmatory factor analysis on the structure of the SPQ and LPQ. *British Journal of Educational Psychology*, *68*, 395–407.
Lai, L.Y.E. (1991). *Effects of cooperative learning on students' learning outcomes and approaches to learning in S3 geography.* Unpublished master's thesis, University of Hong Kong.
Leung, W.Y.W. (1992). *Effects of student-student interaction on approaches to learning and on academic performance.* Unpublished master's thesis, University of Hong Kong.
Marsh, H.W., Balla, J.R., & McDonald, R.P. (1988). Goodness-of-fit indexes in confirmatory factor analysis: The effect of sample size. *Psychological Bulletin*, *103* (3), 391–410.

Marton, F., & Säljö (1976). On quantitative differences in learning – I: Outcome and process. *British Journal of Educational Psychology*, *46*, 4–11.

Marton, F., Dall' Alba, G., & Tse, L.K. (1996). Memorizing and understanding: The keys to the paradox? In D.A. Watkins & J.B. Biggs (Eds.), *The Chinese learner: Cultural psychology and contextual influences* (pp. 69–84). Hong Kong: CERC and ACER.

McDonald, R.P., & Marsh, H.W. (1990). Choosing a multivariate model: Noncentrality and goodness of fit. *Psychological Bulletin*, *107* (2), 247–255.

Nunnally, J.C. (1978). *Pychometric theory* (2nd ed.). New York: McGraw-Hill.

Ramsden, P., & Entwistle, N.J. (1981). Effects of academic departments on students' approaches to studying. *British Journal of Educational Psychology*, *51*, 368–383.

Rosenthal, R., & Rubin, D.B. (1982). A simple general purpose display of magnitude of experimental effect. *Journal of Educational Psychology*, *74*, 166–169.

Schmeck, R.R., & Grove, E. (1979). Academic achievement and individual differences in learning processes. *Applied Psychological Measurement*, *3*, 43–49.

Schmeck, R.R., Geisler-Brenstein, E., & Cercy, S. (1991). The revised inventory of learning processes. *Educational Psychology*, *11*, 343–362.

Schmeck, R.R., Ribich, F.D., & Ramanaiah, N. (1977). Development of a self-report inventory for assessing individual differences in learning processes. *Applied Psychological Measurement*, *1*, 413–431.

Stanley, J.C. (1971). Reliability. In R.L. Thorndike (Ed.), *Educational measurement* (2nd ed.) (pp. 356–442). Washington, DC: American Council on Education.

Steiger, J.H. (1990). Structural model evaluation and modification: An interval estimation approach. *Multivariate Behavioral Research*, *25*, 173–180.

Tucker, L.R., & Lewis, C. (1973). The reliability coefficient for maximum likelihood factor analysis. *Psychometrika*, *38*, 1–10.

Watkins, D.A. (1996). Learning theories and approaches to research: A cross-cultural perspective. In D.A. Watkins & J.B. Biggs (Eds.), *The Chinese learner: Cultural psychology and contextual influences* (pp. 3–24). Hong Kong: CERC and ACER.

Watkins, D.A., & Akande, A. (1994). Approaches to learning of Nigerian secondary school children: Emic and etic perspectives, *International Journal of Psychology*, *29* (2), 165–182.

Watkins, D.A., & Mboya, M. (1997). Assessing the learning processes in Black South African students. *Journal of Psychology*, *131* (6), 632–640.

Watkins, D.A., Akande, A., & Mpofu, E. (1994). The assessment of learning processes: Some African data. *IFE Psychologia*, *2* (2), 1–18.

Watkins, D.A., Dong, Q., & Xia, Y. (1995). Towards the validation of a Chinese version of the self-description questionnaire. *Psychologia*, *38*, 22–30.

Wong, N.Y., Lin, W.Y., & Watkins, D.A. (1996). Cross-cultural validation of models of approaches to learning: An application of confirmatory factor analysis. *Educational Psychology*, *16*, 317–327.

Received 26 April 1999; revised version received 12 January 2000

Appendix

Covariance matrix for LPQ subscale total scores ($N = 1070$)

	SS	SM	DS	DM	AS	AM
SS	13.245					
SM	4.252	17.359				
DS	−1.971	−0.298	16.112			
DM	−2.111	1.545	7.072	14.570		
AS	−1.352	3.204	5.473	4.557	16.194	
AM	0.025	5.277	4.709	6.352	5.416	14.975