Open Research Online

Validating the Course Experience Questionnaire in West Bengal higher secondary education

Journal Item

For guidance on citations see FAQs.

Version: Accepted Manuscript

oro.open.ac.uk

Validating the Course Experience Questionnaire in West Bengal higher secondary education

Anup Kumar Chakrabarty[a,*], John T. E. Richardson[b,**], Malay Kumar Sen[c]

[a]*Department of Education, University of Calcutta, Kolkata 700 027, India.* E-mail: anup2169@gmail.com

[b]*Institute of Educational Technology, The Open University, Milton Keynes MK7 6AA, United Kingdom.* E-mail: John.T.E.Richardson@open.ac.uk

[c]*Department of Education, University of Calcutta, Kolkata 700 027, India.* E-mail: malaykscaluniv@gmail.com

* Corresponding author [post-publication].

    E-mail address: anup2169@gmail.com (A. K. Chakrabarty)

** Corresponding author [prior to publication].

    E-mail address: John.T.E.Richardson@open.ac.uk

Acknowledgements

ABSTRACT

The Course Experience Questionnaire (CEQ) is a 36-item instrument that is intended to measure six different aspects of students' perceptions of the academic quality of their programmes. It has been widely used in Western countries, and it has also been used in non-Western countries, including China, Hong Kong, Japan and Pakistan. Nevertheless, in the latter countries, it has sometimes not been possible to identify the full range of constructs that were supposed to be measured by the original CEQ. We translated the CEQ into Bengali and administered this to 552 science students at 15 higher secondary schools in West Bengal, India. A confirmatory factor analysis found that their responses provided a poor fit to the original six-factor model of the CEQ. An exploratory factor analysis identified just four constructs, which reflected good teaching, generic skills, student support and appropriate workload. The items with salient loadings on the four factors were used to construct four scales. The students' scores on three of the four scales showed satisfactory levels of internal consistency. A factor analysis of their scores on all four scales yielded one overarching factor that could be interpreted as a measure of perceived academic quality. A reduced version of the CEQ consisting of the 30 items that constitute these four scales can be recommended as a measure of students' perceptions of the academic quality of programmes in West Bengal.

Keywords: Confirmatory factor analysis, Course Experience Questionnaire, Exploratory factor analysis, Perceptions of academic quality, West Bengal

Introduction

An important source of evidence about the quality of any educational provision is feedback from the students themselves. This can be obtained through meetings with students, but in recent years there has been an increased interest in the use of formal questionnaires to obtain student feedback (for reviews, see Richardson, 2005; Spooren, Brockx, & Mortelmans, 2013). The constructs that are reflected in students' responses to such questionnaires might be expected to vary from one educational system to another, and so it is important to consider whether or not they generalise across different cultures. In the present study, we evaluated a questionnaire that is often used in English-speaking countries in the different and distinct context of higher secondary education in the Indian state of West Bengal.

The Course Experience Questionnaire

The Course Experience Questionnaire (CEQ) was originally developed by Ramsden (1991a) as a performance indicator for monitoring the quality of teaching on individual programmes of study at Australian universities. In the light of preliminary evidence, a national trial of the CEQ was commissioned, and this yielded usable responses from 3,372 final-year students at 13 Australian universities and colleges of advanced education (see also Ramsden, 1991b). The instrument used in this trial consisted of 30 items in five scales which had been identified in previous research as reflecting different dimensions of effective instruction: good teaching (8 items); clear goals and standards (5 items); appropriate workload (5 items); appropriate assessment (6 items); and emphasis on independence (6 items). The defining items of the five scales (according to the results of the national trial) are shown in Table 1. The respondents were instructed to indicate their level of agreement or

disagreement (along a scale from "definitely agree", scoring 5, to "definitely disagree",

scoring 1) with each statement as a description of their programme of study. Half of the items

referred to positive aspects, whereas the other half referred to negative aspects and were to be

scored in reverse.

Table 1

Defining items of the scales in the original Course Experience Questionnaire.

Scale	Defining item
Good Teaching	Teaching staff here normally give helpful feedback on how you are going
Clear Goals and Standards	You usually have a clear idea of where you're going and what's expected of you in this course.
Appropriate Workload	The sheer volume of work to be got through in this course means you can't comprehend it all thoroughly.[a]
Appropriate Assessment	Staff here seem more interested in testing what we have memorised than what we have understood.[a]
Emphasis on Independence	Students here are given a lot of choice in the work they have to do.

[a]Items to be coded in reverse.

As a result of this national trial, it was decided that the Graduate Careers Council of

Australia should administer the CEQ on an annual basis to all new graduates through the

Graduate Destination Survey, which is conducted a few months after the completion of their

degree programmes. In these surveys, shorter versions of the CEQ have been employed. The

Emphasis on Independence scale was replaced by a new scale concerned with Generic Skills

to "investigate the extent to which higher education contributes to the enhancement of skills

relevant to employment" (Ainley & Long, 1994, p. xii). Additional items were included but

not assigned to any of the scales; one of these measures the respondents' general level of

satisfaction with their programmes. Wilson, Lizzio, and Ramsden (1997) argued that for research purposes the original version of the CEQ should be augmented with the Generic Skills scale to yield a 36-item instrument measuring six scales.

Research studies have typically focused on the validity of the CEQ: that is, whether it measures the personal traits that it purports to measure (for a review, see Richardson, 2009). Its *construct validity* is demonstrated by confirmatory or exploratory factor analysis showing that it measures several distinct scales that reflect different aspects of effective instruction. For instance, Wilson et al. (1997) used both exploratory and confirmatory factor analysis to show that a six-factor model provided a good fit to Australian students' responses to the 36-factor CEQ. They and other researchers have found that students' scores on the six scales are positively correlated with each other, which is consistent with the assumption that they reflect different aspects of perceived academic quality.

The CEQ's *criterion validity* is demonstrated by correlations between students' scale scores and their scores on some independent criterion, such as their responses to the additional item concerned with general satisfaction. For instance, Wilson et al. (1997) found that students' scores on all of the CEQ's scales were positively correlated with their ratings of general satisfaction, and this too has been confirmed by other researchers.

The CEQ's *discriminative validity* is demonstrated by the extent to which it yields different scores on groups of participants who would be expected to differ on the underlying traits. For instance, Wilson et al. (1997) found that students who had followed programmes in the same disciplines but with different teaching philosophies and methods produced different patterns of responses on the CEQ. In addition, the Australian graduate surveys have consistently shown that students' scores on the constituent scales of the CEQ vary across different academic disciplines and across different institutions of higher education offering programmes in the same discipline (see, e.g., Ainley & Long, 1994).

Finally, the CEQ's *convergent validity* is demonstrated by correlations with students' scores on other instruments with which the CEQ should theoretically be related. For instance, Wilson et al. (1997) found that students' who had more positive perceptions of their courses and programmes were more likely to adopt a deep approach to studying and were less likely to adopt a surface approach to studying. This pattern has been confirmed by a number of researchers (e.g., Parpala, Lindblom-Ylänne, Komulainen, & Entwistle, 2013; Richardson, 2007; Rytkönen, Parpala, Lindblom-Ylänne, Virtanen, & Postareff, 2012). As another example, Dorman (2014) found that students' scores on the CEQ were strongly related to their general perceptions of their classroom environment.

The CEQ in non-Western contexts

The CEQ has been mainly used in Australia and the United Kingdom, although items from the CEQ were incorporated in the Experiences of Teaching and Learning Questionnaire, which is used to obtain student feedback annually at the University of Helsinki (Parpala et al., 2013). A Dutch version of the CEQ has also been validated with postgraduate students in The Netherlands (Jansen, van der Meer, & Fokkens-Bruinsma, 2013). However, the CEQ has also been used in some non-Western contexts.

Zhang, Lu, and Cheng (2006) administered a Chinese version of the CEQ to 3,125 first-year students at Xi'an Jiaotong University in China. They carried out a factor analysis of their responses and found just four factors. Three were concerned with appropriate workload, generic skills and good teaching, and the fourth combined items measuring clear goals and standards and appropriate assessment. Ten items failed to show salient loadings on any of the four factors. However, Price, Richardson, Robinson, Ding, Sun, and Han (2011) suggested that Zhang et al. (2006) had extracted too many factors. Price et al. (2011) themselves

administered a Chinese version of the CEQ to 356 students at Gansu Radio and Television University, a distance-learning institution in China. A factor analysis of their responses led to just two factors: one was concerned with the quality of the academic support; the other was concerned with the academic demands of the courses.

Ning and Dowling (2010) administered a short English version of the CEQ on two occasions a year apart to 396 students at a university in Hong Kong. This version contained 23 items intended to measure five of the six scales. It should be noted that the educational system in Hong Kong was imposed under colonial rule by the British, many teachers are not themselves ethnically Chinese, and Western cultural influences have been prominent for 170 years. Confirmatory factor analysis showed a good fit to the intended scale structure on both occasions. However, Law and Meyer (2011) found different results when they administered a Chinese version of the CEQ to 1,572 students at six post-secondary colleges in Hong Kong. They carried out exploratory factor analysis of the students' responses and found four factors: three were concerned with good teaching, appropriate workload and generic skills, and the fourth combined items concerned with appropriate assessment, good teaching and clear goals.

Yin, Lu, and Wang (2014) administered the CEQ to a large sample of students at 15 universities in different regions of China. In contrast to the studies by Zhang et al. (2006) and Price et al. (2011), they found a marginally acceptable fit to the intended six scales in a confirmatory factor analysis after one item had been dropped. Yin and Wang (2015) obtained similar results in a separate study of students at one teaching-oriented university and one research-oriented university in China (see also Yin, Wang, & Han, 2016). These findings may indicate that the more nuanced concepts that are reflected in the original CEQ scales are gradually emerging in the general discourse about higher education in China.

Ullah, Richardson, and Hafeez (2011) included the CEQ in a survey of 912 students at two universities in Pakistan. The CEQ was administered in English (the language of

instruction), but some of the items were reworded for the Pakistani context. A factor analysis of the students' responses yielded just four factors: three were concerned with generic skills, appropriate assessment and appropriate workload; the fourth factor was concerned with good teaching, specifically construed in terms of the students' perceptions of the prevalent instructional practices. Ullah et al. (2011) concluded that the clarity of goals and standards and the role of independence were not salient dimensions along which Pakistani students evaluated different learning environments.

Fryer, Ginns, Walker, and Nakao (2012) employed a Japanese version of the CEQ with students at a post-secondary institution in Japan. Exploratory factor analysis of the responses from 370 students yielded four factors that reflected good teaching, generic skills, appropriate workload and appropriate assessment. Confirmatory factor analysis of the responses produced by a separate sample of 269 students confirmed these four factors. As in the study that had been carried out in Pakistan by Ullah et al. (2011), neither clear goals and standards nor an emphasis on independence emerged as a salient construct for Japanese post-secondary students.

Context for this study

The region of Bengal is in the north-eastern part of the Indian subcontinent. In 1947, it was divided in two, with the western part becoming a state of modern India and the eastern part becoming East Pakistan (later Bangladesh). The education system in West Bengal originated under the period of colonial occupation by the United Kingdom. Compulsory education lasts for 10 years from the age of 6 to the age of 15 and is undertaken in primary and secondary schools. This may be followed by 2 years of higher secondary education in liberal arts, commerce or science before admission to a 3-year university programme.

Throughout the educational system, teaching methods are old-fashioned and didactic. Classes are teacher dominated; interaction in classes is hardly seen; argumentation is viewed as disobedience; raising questions on the part of the students is seen as a challenge to the teacher; and silence on the part of the students is seen as the best form of class management. From the students' point of view, there is strong emphasis on rote learning at the expense of more elaborative learning strategies (Kurtz, 1990), and this extends even to the education of the teachers themselves (Joyce & Showers, 1985).

Successive national enquiries have tried to promote a shift away from assessment that required only the bare memorisation and reproduction of information (e.g., National Advisory Committee, 1993, p. 22; National Council of Educational Research and Training, 2005, pp. 3, 74, 82). However, it remains the case that examinations are seen as the best way to measure students' knowledge and their ability to reproduce it from memory under standard conditions. The aim for both teachers and students is to obtain good grades, not to achieve understanding, and to obtain good grades students are entirely dependent on their teachers' hand-outs.

In West Bengal, there are both public and private educational institutions. However, one consequence of the highly competitive focus upon certification is that increasing numbers of students at public institutions also receive private tuition. Indeed, the Ministry of Statistics and Programme Implementation (2015, p. 19) found that 78% of students in secondary and higher education in West Bengal were receiving private tuition. This was initially provided by individual tutors, but it is being increasingly supplied by organised institutions (Snehi, 2010). These coaching centres provide notes for the students to memorise. This in turn is just one way in which parents continue to have a major influence over the educational choices and decisions of their children, even those in higher secondary education (Rammohan, 2010).

Finally, the national government has been taking affirmative action to increase the educational opportunities available for people from Scheduled Castes and Tribes (formerly

known as "untouchables"). In West Bengal, the Left Front government (which held power from 1977 to 2011) was also concerned to widen educational participation. As a result, large numbers of first-generation learners are now accessing higher secondary education, often with distinctive beliefs, orientations, attitudes, motivations and aspirations. This is therefore a timely moment at which to assess the experiences and perceptions of students in higher secondary education in West Bengal.

Aims of this study

The aim of the present study was to identify the main constructs underlying students' perceptions of science programmes at higher secondary schools in West Bengal. We employed the conventional procedures that were described earlier to assess the psychometric properties of the CEQ in this new context. In particular, we focused on its construct validity, its criterion validity, and its discriminative validity.

Method

Participants and procedure

Fifteen Bengali-medium higher secondary schools were randomly selected from the publicly recognised and funded institutions in West Bengal, and the Class XI science students in each school were surveyed with regard to their previous 6 months' experience of science courses. These are typically high-performing students who are aiming to study engineering, medicine or science in higher education. Usable responses were obtained from 305 boys and 247 girls aged 16–17 years, of whom 254 lived in urban areas and 298 lived in rural areas.

The response rate was 82.7%, mainly because students were absent from relevant classes.

Instrument

The 36-item version of the CEQ recommended by Wilson et al. (1997) was employed along with the 37th item concerned with the students' general satisfaction. The English version of the questionnaire was translated into Bengali and back-translated into English for checking. The results were discussed with experienced colleagues, and minor changes were then made to ensure that the wording was appropriate for the local context. This draft Bengali version was used in a pilot study with 30 urban and rural Class XI science students, and this led to additional amendments for the Bengali context. For instance, *staff*, *lecturer* and *tutor* were all rendered as *teacher*, and statements in the second person were rendered in the first person. The detailed amendments are shown in the appendix. English translations of the final wording of the 36 items are shown in Table 2 below.

Data analysis

The responses to the items in the CEQ were coded from 1 for "definitely disagree" to 5 for "definitely agree". The coding of the 15 negatively worded items was reversed so that higher scores corresponded to more positive ratings.

IBM SPSS AMOS 22 was employed to carry out confirmatory factor analysis on the covariance matrix among the students' responses to the 36 items. This evaluated the fit of the theoretical model proposed by Wilson et al. (1997), which postulates six correlated factors. The assignment of the 36 items to the six scales is shown in Table 2 below. The solution was generated on the basis of maximum-likelihood estimation, and the model's fit was assessed

using the chi-squared degrees of freedom ratio (χ^2/df), the comparative fit index (CFI), the

Tucker-Lewis index (TLI), and the root-mean-square error of approximation (RMSEA).

Values of χ^2/df less than 2 are taken to indicate an acceptable fit. The CFI and the TLI vary

between 0 and 1, and values greater than 0.95 are usually taken to indicate an acceptable fit.

In the case of the RMSEA, values less than 0.06 are usually taken to indicate an acceptable fit

(Jackson, Gillaspy, & Purc-Stephenson, 2009; Ullman, 2014).

IBM SPSS Statistics 21 was then employed to carry out exploratory factor analysis on

the students' responses to the 36 items. First, the Kaiser-Meyer-Olkin (KMO) measure of

sampling adequacy was calculated; this varies between 0 and 1, and values greater than 0.5

indicate that exploratory factor analysis is appropriate (Kaiser, 1974). The number of factors

to be extracted was determined by comparing the eigenvalues of the correlation matrix

among the responses to the individual items with those produced by the parallel analysis of

1,000 random correlation matrices using the program written by O'Connor (2000). Principal

axis factoring was used to extract the relevant number of factors, and these were subjected to

oblique rotation using the promax procedure. Loadings greater than 0.30 in absolute

magnitude were regarded as being salient for the purposes of interpretation, and the factors

were labelled on the basis of the items that showed the highest loadings.

Following Pedhazur and Schmelkin (1991, pp. 625–626), the students were assigned

scores on factor-based scales according to the means of their scores on the salient items in

each of the factors. Descriptive statistics were calculated from the students' scores on these

scales, and their reliability was estimated from values of Cronbach's (1951) coefficient alpha.

We examined the correlation coefficients between the students' scale scores and their ratings

of general satisfaction with their courses, and a second-order exploratory factor analysis was

carried out to investigate the relationships among the first-order factor-based scales using the

same procedures as the first-order exploratory factor analysis.

A multivariate analysis of variance was carried out to compare the scale scores that had been obtained by male and female students attending schools in urban and rural areas. This technique is a generalisation of the simple (univariate) analysis of variance to situations in which there are several dependent variables (in this case, the students' scores on the first-order factor-based scales). It makes the assumption that the within-cell variance–covariance matrices are homogeneous across the individual groups, an assumption that can be evaluated using Box's test. When significant multivariate effects are found, univariate analyses can then be carried out to identify the dependent variables that are responsible for those effects.

Both the multivariate tests and the univariate tests yield the statistic partial η^2 (eta squared). This quantifies the proportion of variance in the dependent variables that is explained by each independent variable or interaction when the effects of other variables and interactions have been statistically controlled (see Tabachnick & Fidell, 2014, pp. 285–354). We expected urban students to obtain higher scores, both because of the influence of socioeconomic factors on community attitudes to education and because of the different level of facilities available in urban and rural schools.

Results

The results of the confirmatory factor analysis did not support the six-factor model proposed by Wilson et al. (1997). A chi-square test showed significant disparity between the model and the data, $\chi^2(579, N = 552) = 2602.03$, $p < 0.001$, which is unsurprising given the large sample size. The RMSEA was marginal at 0.08, although its 95% confidence interval, [0.077, 0.083], did not include the cut-off of 0.06. The other fit indexes all indicated a poor fit to the model: χ^2/df was 4.49, the CFI was 0.55, and the TLI was 0.51.

The KMO measure of sampling adequacy was 0.79, which implied that exploratory factor analysis was appropriate. The parallel analysis of 1,000 random correlation matrices indicated that four factors should be extracted by means of exploratory factor analysis. Their eigenvalues were 5.65, 2.92, 2.77 and 1.62. The proportion of variance explained by each factor was 15.69%, 8.12%, 7.68% and 4.51%, respectively. The pattern factor matrix for the rotated solution is shown in Table 2, where the items have been sorted in descending order of their salient factor loadings.

Table 2

Factor loadings for the 36-item Course Experience Questionnaire.

	Factor 1	Factor 2	Factor 3	Factor 4
20. Teachers here make real effort to understand the difficulties we are facing during learning. (GT)	**0.72**	-0.04	0.08	-0.03
9. Teachers of this course put a lot of time into commenting on our work. (GT)	**0.67**	-0.09	-0.07	-0.11
4. The teachers of this course motivate us to do the best work. (GT)	**0.63**	0.14	0.04	-0.21
22. Teachers here normally give helpful feedback on our performance. (GT)	**0.61**	-0.09	0.01	0.07
25. Teachers of this course work hard to make subjects interesting. (GT)	**0.61**	-0.02	-0.03	0.12
23. Teachers here are extremely good in explaining content to us. (GT)	**0.57**	-0.06	0.14	-0.01
35. Teachers here make it clear right from the start what they expect from us. (CG)	**0.52**	0.15	-0.20	-0.13
30. We often discuss with our teachers here on our ways of learning. (IN)	**0.41**	0.16	-0.06	0.09
7. Teachers of this course frequently give the impression that they have nothing to learn from us.* (AA)	**0.33**	-0.07	**0.32**	0.25
13. This course has improved my written communication skills. (GS)	-0.03	**0.60**	0.04	-0.05
2. This course has helped me to develop my problem-solving skills. (GS)	-0.03	**0.57**	0.15	0.04
15. This course has developed my academic interest as far as possible. (IN)	0.09	**0.57**	0.13	0.07
28. This course has helped me in developing the ability to plan my own work. (GS)	0.07	**0.53**	-0.13	0.02
12. As a result of pursuing this course, I feel more confident about tackling unfamiliar problems. (GS)	0.03	**0.53**	0.06	-0.11
6. This course has sharpened my analytic skills. (GS)	0.01	**0.48**	0.15	-0.11
33. This course really tries to get the best out of all the students. (GT)	-0.01	**0.45**	-0.07	-0.07
16. Students of this course are given a great deal of opportunity in selecting for themselves their ways of learning. (IN)	0.10	**0.40**	-0.05	0.24
1. It is always easy here to know the standard of work expected. (CG)	-0.18	**0.39**	0.04	0.07
11. This course has helped me in developing my ability to work as a team member. (GS)	0.06	**0.37**	0.07	-0.01
24. The aims and objectives of this course are not made very clear.* (CG)	-0.21	0.19	**0.56**	-0.02
31. Teachers here show no real interest on students' opinions.* (GT)	0.27	0.02	**0.56**	-0.12

34. There is a very little choice in this course in the ways I am assessed.* (IN)	0.17	-0.04	**0.56**	0.05
32. It is possible to get through this course just by working hard around examination time.* (AA)	-0.07	0.06	**0.52**	0.17
18. It is often hard to know what is expected of me in this course.* (CG)	-0.17	0.26	**0.50**	0.05
27. There is a lot of pressure on me as a student here.* (AA)	-0.04	0.06	**0.40**	0.19
29. Usually feedback on our work is provided only in marks.* (AA)	0.12	-0.28	**0.33**	0.24
5. The workload of this course is too heavy.* (AW)	-0.06	-0.07	0.05	**0.45**
36. The volume of work to be got through in this course means I cannot comprehend it all thoroughly.* (AW)	-0.08	-0.03	0.25	**0.44**
19. In this course we are given enough time to understand the things that we have to learn. (AW)	0.14	0.13	-0.24	**0.43**
21. Students of this course are given a lot of choice in the work that they have to do. (IN)	0.23	0.13	-0.07	**0.40**
14. It seems to me that the syllabus of this course is vast.* (AW)	-0.05	-0.14	0.11	**0.30**
3. There are few opportunities to choose the particular areas I want to study.* (IN)	0.14	-0.01	0.21	0.06
8. I have a clear idea of where I am going and what is expected of me. (CG)	0.11	0.24	-0.04	-0.18
10. To do well in this course we all really need a good memory.* (AA)	-0.14	-0.20	-0.04	0.23
17. Teachers of this course are more interested in testing what I have memorised than what I have understood.* (AA)	-0.03	-0.06	0.22	0.03
26. Most of the teachers here ask questions on facts.* (AA)	-0.06	-0.05	0.19	0.15
Factor intercorrelations				
Factor 1	1.00	0.43	0.15	0.30
Factor 2	0.43	1.00	0.21	-0.04
Factor 3	0.15	0.21	1.00	0.05
Factor 4	0.30	-0.04	0.05	1.00

[Table note] This table shows English translations of the items in the final questionnaire. Each item is preceded by a number showing the order in which it was presented in the questionnaire and followed by an abbreviation indicating the scale of the original CEQ from which it is taken: GT, Good Teaching; CG, Clear Goals and Standards; GS, Generic Skills; AA, Appropriate Assessment; AW, Appropriate Workload; IN, Emphasis on Independence. Items that were scored in reverse are indicated by asterisks. Factor loadings greater than 0.30 in absolute magnitude are shown in bold.

Factor 1 showed salient loadings on six of the eight items from the Good Teaching scale, one item from the Appropriate Assessment scale, one item from the Clear Goals and Standards scale, and one item from the Emphasis on Independence scale. It was interpreted as measuring Good Teaching. Factor 2 showed salient loadings on all six items from the Generic Skills scale, two items from the Emphasis on Independence scale, one item from the Good Teaching scale, and one item from the Clear Goals and Standards scale. It was interpreted as measuring Generic Skills. Factor 3 showed salient loadings on three items from the Appropriate Assessment scale, two items from the Clear Goals and Standards scale, one item from the Appropriate Workload scale, one item from the Good Teaching scale, and one item from the Emphasis on Independence scale. All of these items refer to a lack of student support, but they are all coded in reverse. This factor was therefore interpreted as measuring Student Support. Whereas Factor 1 is concerned with teachers' instructional practices, this factor is concerned with institutional support for students in a more generic sense, including their teachers' attitudes. It might also be noted from Table 2 that there was only a weak correlation between Factors 1 and 3. Factor 4 showed salient loadings on four of the five items in the Appropriate Workload scale and one item from the Emphasis on Independence scale. This factor was interpreted as measuring Appropriate Workload.

Out of the 36 items in the CEQ, three items from the Appropriate Assessment scale, one item from the Clear Goals and Standards scale, and one item from the Emphasis on Independence scale failed to show any salient loadings. Item 7 showed low cross-loadings on two different factors, and so this item too was omitted in the construction of the four factor-based scales (cf. Worthington & Whittaker, 2006). Descriptive statistics of the participants' scores on these scales, values of coefficient alpha, and correlation coefficients with their ratings of general satisfaction are shown in Table 3. The values of coefficient alpha for the first three scales would be regarded as satisfactory on conventional research-based criteria,

but the value for the fourth would not (Robinson, Shaver, & Wrightsman, 1991). The

correlation coefficients between the scores on the four factor-based scales and the students'

ratings of their general satisfaction were all positive and statistically significant, which

confirms the criterion validity of the scale scores as measures of perceived academic quality.

Table 3

Descriptive statistics for factor-based scales.

Scale	Number of items	Mean	Standard deviation	Coefficient alpha	Correlation with satisfaction
Good Teaching	8	3.67	0.24	0.81	0.22**
Generic Skills	10	3.89	0.66	0.76	0.31**
Student Support	7	3.08	0.37	0.69	0.11**
Appropriate Workload	5	2.61	0.61	0.53	0.15**

**$p < 0.01$ (two-tailed tests).

Table 4 shows the correlation coefficients among the scores on the four factor-based

scales; these too were all positive and statistically significant. The KMO measure of sampling

adequacy among the four factor-based scales was 0.56, which implied that a second-order

factor analysis was appropriate. The parallel analysis of 1,000 random correlation matrices

indicated that just one factor should be extracted. Its eigenvalue was 1.67, and it explained

41.68% of the variance. The final column in Table 4 shows the loadings of the four factor-

based scales on this second-order factor. It had salient loadings on all four scales, but it was

most closely related to scores on Good Teaching and Generic Skills.

Table 4

Correlation coefficients among factor-based scales and second-order factor loadings.

Scale	Correlation coefficients				
	Good Teaching	Generic Skills	Student Support	Appropriate Workload	Second-order factor loadings
Good Teaching	—	0.42**	0.15**	0.29**	0.72
Generic Skills		—	0.25**	0.10*	0.57
Student Support			—	0.11**	0.31
Appropriate Workload				—	0.32

*$p < 0.05$; **$p < 0.01$ (two-tailed tests).

Finally, a multivariate analysis of variance was carried out on the students' scale scores using the independent variables of gender (male versus female) and setting (urban versus rural). Box's test showed that the variance–covariance matrices were homogeneous across the four subgroups, $F(30, 735497) = 0.91$, $p = 0.60$, which implies that the factor structure that we had identified applies to all four subgroups. The multivariate effect of gender was not statistically significant, $F(4, 545) = 0.57$, $p = 0.69$, partial $\eta^2 = 0.00$, and none of the univariate effects of gender was significant.

In contrast, the multivariate effect of setting was highly significant, $F(4, 545) = 6.07$, $p < 0.001$, partial $\eta^2 = 0.04$. Univariate tests showed that this was due to a highly significant difference in scores on Student Support, $F(1, 548) = 15.19$, $p < 0.001$, partial $\eta^2 = 0.03$: students at urban schools ($M = 21.56$; $SD = 5.37$) tended to rate their courses more positively than did students at rural schools ($M = 20.25$; $SD = 5.50$). Finally, the multivariate interaction between the effects of gender and setting was not statistically significant, $F(4, 545) = 0.19$, $p = 0.95$, partial $\eta^2 = 0.00$, and none of the univariate interactions was significant.

Discussion

The aim of this study was to identify the main constructs underlying the perceptions of science students at higher secondary schools in West Bengal. The CEQ was developed in Western countries but has subsequently been used in China, Hong Kong, Japan and Pakistan. So far as we are aware, this is the first time that the CEQ has been used in the distinct cultural and contextual situation of India. The results are interesting, both in theory and in practice, and they contribute to the growing literature on the perceptions of students in non-Western educational contexts. Specifically, they indicate the extent and the limitations of the construct validity of the CEQ when used in higher secondary education in West Bengal.

Based on the findings of other research in non-Western contexts, we thought it was unlikely that the perceptions of science students taking higher secondary education in West Bengal would show the full range of constructs tapped by the CEQ. As we expected, in the confirmatory factor analysis, the students' responses showed a poor fit to the theoretical model that was originally proposed by Wilson et al. (1997). Instead, the exploratory factor analysis identified just four constructs: good teaching, generic skills, student support and appropriate workload. Three of the four constructs (good teaching, generic skills and appropriate workload) broadly match scales in the extended CEQ advocated by Wilson et al. (1997). In addition, they represent the CEQ scales that have been most consistently identified in previous investigations in China (Yin et al., 2014; Yin & Wang, 2015; Zhang et al., 2006), Hong Kong (Law & Meyer, 2011; Ning & Downing, 2010), Japan (Fryer et al., 2012) and Pakistan (Ullah et al., 2011). They would appear to reflect fairly universal aspects of the perceptions of students in most, if not all, educational systems around the world, although, clearly, more research is needed to confirm this notion.

The emergence of generic skills as a construct in this study is of particular interest,

because it was added to the original CEQ in response to concerns about the employability of

Australian graduates (see Ainley & Long, 1994, p. xii). In his own research with students at

secondary schools in West Bengal, Sen (1999) had discounted a scale concerned with the

vocational relevance of their courses, but our own students who were following a science

curriculum in higher secondary education seem to recognise a cluster of key skills or core

skills relevant to their subsequent employment. (In fact, most of the students were probably

anticipating careers in engineering, medicine or science.)

Even so, neither clear goals and standards nor an emphasis on independence emerged

as a salient construct in the present study, even though many of the constituent items loaded

on other factors. This outcome is consistent with the findings of many of the previous

investigations in non-Western countries just cited. These definitely do not appear to reflect

universal aspects of the perceptions of students around the world. This may reflect variations

in whether assessment tasks are linked to specific learning outcomes and whether the students

are expected to engage in independent critical thinking in particular educational systems.

The present study also failed to identify a construct that could be characterised as

appropriate assessment. Such a construct has been identified in some studies carried out in

non-Western countries (Fryer et al., 2012; Ning & Downing, 2010; Ullah et al., 2011; Yin et

al., 2014) but not in others (Law & Meyer, 2011; Price et al., 2011; Zhang et al., 2006). In the

context of West Bengal, this could be because all students are exposed almost exclusively to

assessments that test their rote memory. Instead, the present results identified a relatively

undifferentiated construct reflecting student support that was measured by items from five of

the original six scales in the CEQ. A similar construct has been identified in some previous

studies in non-Western countries (Law & Meyer, 2011; Price et al., 2011; Zhang et al., 2006),

but not all. In the present case, it might be noted that all of the items with salient loadings on

this factor were coded in reverse (see Table 2), suggesting that the students may simply have

been responding to a common negative semantic or syntactic theme.

Scales based on three of the four constructs identified in this study achieved adequate levels of reliability, but the fourth (based on appropriate workload) did not (see Table 3). This is not surprising, since it was also the scale containing the fewest items, and it is well known that the value of coefficient alpha tends to vary directly with the number of items (Cronbach, 1951). Researchers and practitioners who wish to use this scale in the context of West Bengal may choose to incorporate additional items to try to increase its reliability. Poor reliability of factor-based scales measuring appropriate workload was noted previously by Law and Meyer (2011) in Hong Kong and by Fryer et al. (2012) in Japan.

This would be less important if the scores on the four factor-based scales were simply combined to yield a single overall measure of perceived academic quality that could then be compared across different programmes. This interpretation of the CEQ would be warranted by the fact that a second-order factor analysis yielded just one factor that had salient loadings on all four of the first-order scales. Following the logic of previous research using the CEQ, this factor can be interpreted as a measure of perceived academic quality (Richardson, 2009). In other words, the *construct validity* of the CEQ in the present context is shown by the fact that it measures four distinct scales that reflect different aspects of effective instruction.

The *criterion validity* of the CEQ is shown by the fact that the students' scores on all four of the scales were positively and significantly correlated with their ratings of general satisfaction with their programmes. Moreover, its *discriminative validity* is shown by the fact that, as expected, students in urban areas tended to rate their courses more positively in terms of Student Support than students in rural areas. There was no significant difference between male and female students in their CEQ scores, but this is consistent with the findings of previous research. In the Australian graduate surveys, any gender differences on the CEQ have been small in magnitude and have usually been attributed to male and female students

opting for different subjects of study (e.g., Ainley & Long, 1994, p. 10). In our study, the

male and the female students were matched on the basis of subject of study, since they had all

chosen to study science in higher secondary education.

On the basis of our results, we suggest that a reduced version of the CEQ containing

just the 30 items that constituted these four scales could be employed to measure students'

perceptions of the academic quality of their programmes and institutions in West Bengal. In

fact, our findings agree with those of Jansen et al. (2013) in suggesting that the CEQ can be

adapted to diverse educational contexts in many countries around the world. Of course, our

participants had chosen to study the sciences, and the science curriculum contains its own

challenges. It will therefore be important to establish whether the present findings can be

generalised to students taking liberal arts and commerce in higher secondary education and to

students taking other subjects in Indian higher education. In principle, however, such an

instrument would be of immense practical value in monitoring and enhancing the quality of

the educational provision in both secondary and higher education.

As an overall measure of perceived academic quality, CEQ scores could be used to

rank and benchmark individual schools in terms of the quality of their teaching. This kind of

information could be disseminated to ensure public accountability and to assist prospective

students and their families in their choice of schools at higher secondary level. More detailed

information could be used by high-scoring schools to identify aspects of good practice and by

low-scoring schools to identify areas needing improvement (as the CEQ was originally used

at Australian universities). Time-series data over successive years could be used to monitor

improvements in the quality of science teaching at higher secondary level in West Bengal.

The detailed composition of the CEQ scales serves to identify how the abstract notion

of academic quality is construed in this context. In particular, "good teaching" encompasses:

an emphatic approach whereby teachers try to understand the difficulties students are having;

giving helpful feedback that diagnoses strengths and weaknesses and improves performance; encouraging positive expectations that motivate students to achieve the best results; adopting stimulating learning designs; giving clear and useful expectations; and scaffolding students' ways of learning. These characteristics can all be promoted in programmes for training both new teachers and existing teachers to enhance the quality of higher secondary education.

Despite the relative weakness of the appropriate workload scale, it provides useful information about the student experience for educational planners and administrators. They might note that it yielded the lowest overall score of the four scales (see Table 3), and this might lead them to reconsider the volume of material in the curriculum, the amount of time available to study it, and the amount of choice that students have in covering that material. Finally, they might also consider the extent to which the curriculum is promoting generic or "soft" skills in higher secondary education. This yielded the highest overall score of the four scales, and so it is an aspect of academic quality to which students themselves seem to attach importance. The CEQ provides a systematic way of monitoring both of these aspects of the curriculum in the distinctive context of higher secondary education in West Bengal.

References

Ainley, J., & Long, M. (1994). *The Course Experience Survey 1992 Graduates*. Canberra: Australian Government Publishing Service.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3), 297–334. doi:10.1007/BF02310555

Dorman, J. P. (2014). Classroom psychosocial environment and course experiences in pre-service teacher education courses at an Australian university. *Studies in Higher Education*, *39*(1), 34–47. doi:10.1080/03075079.2012.674936

Fryer, L. K., Ginns, P., Walker, R. A., & Nakao, K. (2012). The adaptation and validation of the CEQ and the R-SPQ-2F to the Japanese tertiary environment. *British Journal of Educational Psychology*, 82(4) 549–563. doi:10.1111/j.2044-8279.2011.02045.x

Jackson, D. L., Gillaspy, J. A., Jr., & Purc-Stephenson, R. (2009). Reporting practices in confirmatory factor analysis: An overview and some recommendations. *Psychological Methods*, *14*, 6–23. doi:10.1037/a0014694

Jansen, E., van der Meer, J., & Fokkens-Bruinsma, M. (2013). Validation and use of the CEQ in The Netherlands. *Quality Assurance in Education*, *21*(4), 330–343. doi:10.1108/QAE-11-2012-0041

Joyce, B., & Showers, B. (1985). Teacher education in India: Observations on American innovations abroad. *Educational Researcher*, 14(8), 3–9. doi:10.3102/0013189X014008003

Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika*, *39*, 31–36. doi:10.1007/BF02291575

Kurtz, B. E. (1990). Cultural influences on children's cognitive and metacognitive development. In W. Schneider & F. E. Weinert (Eds.), *Interactions among aptitudes,*

strategies, and knowledge in cognitive performance (pp. 177–199). New York: Springer-Verlag.

Law, C. S., & Meyer, J. H. F. (2011). Adaptation and validation of the Course Experience Questionnaire in the context of post-secondary education in Hong Kong. *Quality Assurance in Education*, 19(1), 50–66. doi:10.1108/09684881111107753

Ministry of Statistics and Programme Implementation, National Sample Survey Office. (2015). *Key indicators of social consumption in India education: NSS 71st round.* Kolkata, India: Ministry of Statistics and Programme Implementation. Retrieved from http://mail.mospi.gov.in/index.php/catalog/160/download/1933

National Advisory Committee. (1993). *Learning without burden.* New Delhi: Ministry of Human Resource Development, Department of Education. Retrieved from http://www.teindia.nic.in/Files/Reports/CCR/Yash%20Pal_committe_report_lwb.pdf.

National Council of Educational Research and Training. (2005). *National Curriculum Framework.* (2005). New Delhi: Author. Retrieved from http://www.ncert.nic.in/rightside/links/pdf/framework/english/nf2005.pdf.

Ning, H. K., & Downing, K. (2010). Connections between learning experience, study behaviour and academic performance: A longitudinal study. *Educational Research*, 52(4), 457–468. doi:10.1080/00131881.2010.524754

O'Connor, B. P. (2000). SPSS and SAS programs for determining the number of components using parallel analysis and Velicer's MAP test. *Behavior Research Methods, Instruments, and Computers*, *32*, 396–402. doi:10.3758/BF03200807

Parpala, A., Lindblom-Ylänne, S., Komulainen, E., & Entwistle, N. (2013). Assessing students' experiences of teaching–learning environments and approaches to learning: Validation of a questionnaire in different countries and varying contexts. *Learning Environments Research*, 16(2), 201–215. doi:10.1007/s10984-013-9128-8

Pedhazur, E. J., & Schmelkin, L. P. (1991). *Measurement, design, and analysis: An integrated approach*. Hillsdale, NJ: Erlbaum.

Price, L., Richardson, J. T. E., Robinson, B., Ding, X., Sun, X., & Han, C. 2011. Approaches to studying and perceptions of the academic environment among university students in China. *Asia Pacific Journal of Education*, 31(2), 159–175. doi:10.1080/02188791.2011.566996

Rammohan, L. (2010). Learning orientation and perceived parental attitudes of students at the senior secondary level. *Journal of Indian Education*, 36(2), 128–138. Retrieved from http://www.ncert.nic.in/publication/journals/pdf_files/jie/JIE_August_2010.pdf.

Ramsden, P. (1979). Student learning and perceptions of the academic environment. *Higher Education*, 8(4), 411–427. doi:10.1007/BF01680529

Ramsden, P. (1991a). A performance indicator of teaching quality in higher education: The Course Experience Questionnaire. *Studies in Higher Education*, 16(2), 129–150. doi:10.1080/03075079112331382944

Ramsden, P. (1991b). Report on the course experience questionnaire trial. In R. D. Linke (Ed.), *Performance indicators in higher education: Report of a trial evaluation study commissioned by the Commonwealth Department of Employment, Education, and Training. Volume 2: Supplementary papers* (pp. 1–85). Canberra: Australian Government Publishing Service.

Richardson, J. T. E. (2005). Instruments for obtaining student feedback: A review of the literature. *Assessment and Evaluation in Higher Education*, 30(4), 387–415. doi:10.1080/02602930500099193

Richardson, J. T. E. (2007). Variations in student learning and perceptions of academic quality." In N. Entwistle, P. Tomlinson, & J. Dockrell (Eds.), *Student learning and university teaching* (pp. 61–71). Leicester: British Psychological Society.

doi:10.1348/000709906X162082

Richardson, J. T. E. (2009). What can students' perceptions of academic quality tell us? Research using the Course Experience Questionnaire. In M. Tight, K. H. Mok, J. Huisman, & C. C. Morphew (Eds.), *The Routledge international handbook of higher education* (pp. 199–210). New York: Routledge.

Robinson, J. P., Shaver, P. R., & Wrightsman, L. S. (1991). Criteria for scale selection and evaluation. In J. P. Robinson, P. R. Shaver & L. S. Wrightsman (Eds.), *Measures of personality and social psychological attitudes* (pp. 1–16). San Diego, CA: Harcourt Brace Jovanovich.

Rytkönen, H., Parpala, A., Lindblom-Ylänne, S., Virtanen, V., & Postareff, L. (2012). Factors affecting bioscience students' academic achievement. *Instructional Science*, 40(2), 241–256. doi:10.1007/s11251-011-9176-3

Sen, M. K. (1999). *Learning strategy: Its relationship with environmental factors, course perceptions and academic performance*. Unpublished doctoral dissertation, University of Calcutta.

Snehi, N. (2010). Private tuition and its implications on quality secondary education. *Journal of Indian Education*, 36(1), 61–75. Retrieved from http://www.ncert.nic.in/publication/journals/pdf_files/jie/jie_may_2010.pdf.

Spooren, P., Brockx, B., & Mortelmans, D. (2013). On the validity of student evaluation of teaching: The state of the art. *Review of Educational Research*, 83(4), 598–642. doi:10.3102/0034654313496870.

Tabachnick, B. G., & Fidell, L. S. (2014). *Using multivariate statistics* (6th ed.). Harlow, UK: Pearson Education.

Ullah, R., Richardson, J. T. E., & Hafeez, M. (2011). Approaches to studying and perceptions of the academic environment among university students in Pakistan. *Compare*, 41(1),

113–127. doi:10.1080/03057921003647065

Ullman, J. B. (2014). Structural equation modeling. In B. G. Tabachnick & L. S. Fidell,
Using multivariate statistics (6th ed., pp. 731–836). Harlow, UK: Pearson Education.

Wilson, K. L., Lizzio, A., & Ramsden, P. (1997). The development, validation and
application of the Course Experience Questionnaire. *Studies in Higher Education*,
22(1), 33–53. doi:10.1080/03075079712331381121

Worthington, R. L., & Whittaker, T. A. (2006). Scale development research: A content
analysis and recommendations for best practices. *The Counseling Psychologist*, 34(6),
806–838. doi:10.1177/0011000006288127

Yin, H., Lu, G., & Wang, W. (2014). Unmasking the teaching quality of higher education:
Students' course experience and approaches to learning in China. *Assessment and
Evaluation in Higher Education*, 39(8), 949–970.
doi:10.1080/02602938.2014.880107

Yin, H., & Wang, W. (2015). Assessing and improving the quality of undergraduate teaching
in China: The Course Experience Questionnaire." *Assessment and Evaluation in
Higher Education*, 40(8), 1032–1049. doi:10.1080/02602938.2014.963837

Yin, H., Wang, W., & Han, J. (2016). Chinese undergraduates' perceptions of teaching
quality and the effects on approaches to studying and course satisfaction. *Higher
Education*, *71*(1), 39–57. doi:10.1007/s10734-015-9887-5

Zhang, P., Lu, G., & Cheng, W. (2006). 教学情景对大学生学习方式的影响 [Impact of
teaching situations on undergraduates' approaches to learning]. *Research in Teaching*,
29(4), 301–305.

Appendix

Amendments to the original Course Experience Questionnaire

The following items were amended from the wording provided by Wilson et al. (1997).

Words in italics were deleted; those in square brackets were added.

3. There are few opportunities to choose the particular areas *you* [I] want to study

4. The *teaching staff* [teachers] of this course motivate *students* [us] to do *their* [the] best

   work

5. The workload [of this course] is too heavy

7. *Lecturers here* [Teachers of this course] frequently give the impression they have nothing

   to learn from *students* [us]

8. *You* [I] *usually* have a clear idea of where *you're* [I'm] going and what's expected of *you*

   [me]

9. *Staff* [Teachers] here put a lot of time into commenting on *students'* [our] work

10. To do well on this course *all you* [we all] really need *is* a good memory

12. As a result of *doing* [pursuing] this course, 1 feel more confident about tackling

   unfamiliar problems

14. It seems to me that the syllabus [of this course] *tries to cover too many topics* [is vast]

17. *Staff* [Teachers] seem more interested in testing what *you've* [I've] memorised than what

   you've [I've] understood

18. It's often hard to discover what's expected of *you* [me] in this course

22. *Teaching staff* [Teachers] here normally give helpful feedback on how *you are going* [our

   performance]

23. *Our lecturers* [Teachers here] are extremely good at explaining *things* [content] to us

25. *Teaching staff* [Teachers] *here* [of this course] work hard to make subjects interesting

26. *Too many staff* [Most of the teachers here] ask *us* questions *just about* [on] facts

27. There's a lot of pressure on *you* [me] as a student here

29. [Usually] Feedback on *student* [our] work is provided only in *the form of* marks *and grades*

30. We often discuss with our *lecturers or tutors* [teachers] *how we are going to learn in this course* [on our ways of learning]

31. *Staff* [Teachers] here show no real interest in *what students have to say* [students' opinions]

32. It *would be* [is] possible to get through this course just by working hard around exam times

33. This course really tries to get the best out of all *its* [the] students

34. There's very little choice in this course in the ways *you are* [I am] assessed

35. *The staff* [Teachers] here make it clear right from the start what they expect from *students* [us]

36. The *sheer* volume of work to be got through in this course means *you can't* [I cannot] comprehend it all thoroughly