Daniel T.L. Shek*, Lu Yu and Joe Ngai

Evaluation of a general education program in Hong Kong: secondary data analyses based on student feedback questionnaires

Abstract: The General University Requirements (GUR) at The Polytechnic University of Hong Kong (PolyU) has been developed and implemented since the 2012–13 academic year under the reform of education system in Hong Kong. To examine the effectiveness of GUR at PolyU, the present study investigated student's subjective perception in the academic years of 2012–13 and 2013–14 using the Student Feedback Questionnaires. Results showed that the GUR subjects were generally well-received by the students. Besides, students held different perceptions of different GUR components, and there was improvement in students' learning experience and staff teaching over time. While there was an increased satisfaction with Language and Communication Requirements subjects, there was a decline in satisfaction with Service Learning subjects.

Keywords: Chinese adolescents; general education; secondary data analyses; subjective outcome evaluation; university students.

Introduction

General Education (GE) is an important component of higher education and its importance is growing. GE has long been embedded in the curriculum of higher education in the Western world, particularly North America [1–3]. To balance the overemphasis on professional education and the need to promote holistic development of undergraduate students, many colleges and universities have designed courses on a broad range of skills and knowledge as requirements, which become known as GE or core curriculum, to all students before graduation [4, 5]. The key objective of GE is to enhance and promote students' generic competencies through introducing the broad-based subjects. Boyer and Arthur [6] found that roughly three-quarter of students from American colleges and universities responded that general education courses "added to the enrichment of other courses" and "helped prepare them for lifelong learning" (p. 85).

Although universities in Hong Kong have been generally doing well in terms of graduates' professional competences in the past decades, their ability to train graduates who are welcomed by the potential employers has been questioned recently. For example, according to the result of a survey reported in the *South China Morning Post* [7], <20% of the business leaders agreed that Hong Kong university graduates possessed sufficient and relevant skills to cope with the frequently changing market. In view of this, the Education Commission of Hong Kong recommended to lengthen the period of university education from 3 to 4 years to allow more time for students' all-round development and thus to "make Hong Kong a more vibrant, economically powerful, cultured, civilized and socially active and responsible society" (p. 4) [8]. It was also expected that the proposed change of education system by the Government of Hong Kong would "play a facilitating role in linking the mainland and the world at large" and become "the education hub of the region" (p. 5) [8]. In the higher education reform, every university is required to create its general education program.

At The Hong Kong Polytechnic University (PolyU), the General Education program is called "General University Requirements" which aims to develop six core attributes of PolyU graduates, including effective communication, innovative problem solving, critical thinking, ethical

*Corresponding author: Daniel T.L. Shek, PhD, FHKPS, SBS, JP, Associate Vice President and Chair Professor, Faculty of Health and Social Sciences, Department of Applied Social Sciences, The Hong Kong Polytechnic University, Hunghom, Hong Kong, P.R. China, E-mail: daniel.shek@polyu.edu.hk; Centre for Innovative Programs for Adolescents and Families, The Hong Kong Polytechnic University, Hong Kong, P.R. China; Department of Social Work, East China Normal University, Shanghai, P.R. China; Kiang Wu Nursing College of Macau, Macau, P.R. China; and Division of Adolescent Medicine, Department of Pediatrics, Kentucky Children's Hospital, University of Kentucky College of Medicine, Lexington, Kentucky, USA
Lu Yu and Joe Ngai: Department of Applied Social Sciences, The Hong Kong Polytechnic University, Hong Kong, P.R. China

leadership, and lifelong learning. There are several components in the GUR as follows:

1. Freshman Seminar (FS): Introduce students to the professional world of a broad discipline, cultivate higher order thinking skills, encourage entrepreneurship, and learn to engage in self-directed and autonomous study.
2. Language and Communication Requirements (LCR): Develop language competence for academic and professional needs.
3. Leadership and Intra-Personal Development (LIPD): Develop self-awareness, acquire interpersonal and self-reflection skills essential for an effective leader.
4. Cluster Area Requirements (CAR): Expand intellectual capacity beyond disciplinary domain.
5. Service Learning (SL): Apply the knowledge and skills from university learning to meet community needs.
6. Healthy Lifestyle (HL): Acquire the basic knowledge and skills to maintain a healthy lifestyle and quality of life by improving their physical well-being.

There are many ways to evaluate teaching and learning in higher education, and subjective outcome evaluation is one of the commonly adopted strategies [9]. In subjective outcome evaluation, the primary aim is to understand whether the program participants are satisfied with the subject, such as the subject design, instructor, and benefits of the subject. There are many subjective outcome evaluation tools in the context of higher education. For example, Kember and Leung [10] developed the Student Feedback Questionnaire (SFQ) at PolyU which comprised six dimensions, namely learning outcomes, interaction, individual help, organization and presentation, motivation, and feedback to gauge the effectiveness of teaching. In view of the changing curriculum and student needs, a modified SFQ was developed and validated at PolyU [11]. To address the varying and specific contexts and needs of different GUR components, a custom version of SFQ was further designed for GUR subjects [11].

In this study, student evaluation of the GUR subjects was examined using data based on the custom version SFQ. Three research questions were addressed in this study.

– Were students satisfied with the GUR subjects as shown by the general profiles of SFQ ratings in 2012–13 and 2013–14 academic years?
– Were there any differences in students' SFQ ratings on various GUR components?
– Were there any changes in students' perceived GUR learning experiences between 2012–13 academic year and 2013–14 academic year in terms of students' SFQ ratings?

Methods

All students taking GUR subjects in 2012–13 and 2013–14 academic years completed a SFQ for each GUR subject. In 2012–13 and 2013–14 years, a total of 21,080 and 32,339 full-time students were invited to participate in the survey. In total, 15,810 from 747 classes and 17,463 students from 1004 classes completed the questionnaire in 2012–13 and 2013–14, respectively. As the questionnaire survey was conducted in an anonymous manner and no identifiable individual information was available, the characteristics of individual respondent were unknown.

Procedures

Two different ways of SFQ were used for collecting student feedback on GUR subjects: traditional paper-and-pencil survey and online survey (eSFQ). In the first semester of 2012–13 academic year, paper-and-pencil survey was used for all but two GUR subjects (which adopted eSFQ as a trial). Starting from the second semester of 2012–13 academic year, all GUR subject were evaluated by eSFQ. For the paper-and-pencil survey, academic departments arranged the data collection and the person(s) responsible for distributing and collecting the forms strictly following the procedures specified by the University. For eSFQ, the Educational Developmental Center (EDC) at PolyU sent emails to students reminding them to respond to the survey online in the period of one week before and after the completion of the subject. The data were saved automatically to the database of the university Information Technology Services (ITS) office and were directly downloaded by EDC staff for data analysis and report writing. To avoid confounding of data collection methods, data collected in the first semester of 2013–14 academic year (based on both paper-and-pencil survey and online survey) were excluded from analyses in the present study. The number of respondents and response rates in each component in semester 2 of 2012–13 and semester 1 and semester 2 of 2013–14 are presented in Table 1.

Instruments

The SFQ for GUR subjects developed by the Educational Development Centre [11] at PolyU is a school or faculty-based system for collecting student feedback on different GUR subjects taken at the end of each semester. The questionnaire comprised two parts. The first part focuses on student's learning experience in terms of student's evaluation of the subject (e.g. "I have a clear understanding of what I am expected to learn from this subject"). The second part examines student's perceptions of the performance of the teaching staff (e.g. "the teaching of the staff member has provided me with valuable learning experience"). All items were scored on a 5-point Likert scale (5=strongly agree, 4=agree, 3=no strong view, 2=disagree, 1=strongly disagree), with higher scores indicating a higher teaching effectiveness and a better quality of the subject. For each part, there are both standardized items which are common to all GUR subjects and subject-specific items drafted by teaching staffs of respective GUR subjects. According to Kwan [12], the scales based on these two parts showed high internal consistency.

Table 1: Number of respondents and response rates of SFQ for each GUR component in 2012–13 and 2013–14 academic years.

Components	2012–13		2013–14	
	Number of respondents	Response rate	Number of respondents	Response rate
FS	954	68%	2448	44%
LCR	2871	64%	6894	58%
CAR	2008	65%	5603	47%
LIPD	722	64%	1671	56%
SL	52	59%	847	40%
Total	6607	64%	17,463	54%

The results for 2012–13 academic year were based on the second semester only in which all GUR subjects were evaluated via eSFQ.

In the present study, students' responses to four standardized items in the first part (learning experience) and two standardized items in the second part (teaching of the staff) were analyzed and reported. The six common items are listed below.

- IA1: I have a clear understanding of what I am expected to learn from this subject.
- IA2: The teaching and learning activities (e.g. lectures, discussions, case studies, projects, etc.) have helped me to achieve the subject learning outcomes.
- IA3: The assessments require me to demonstrate my knowledge, skills and understanding of the subject.
- IA4: I understand the criteria according to which I will be regarded.
- IIA6: The teaching of the staff member has provided me with valuable learning experience.
- IIA7: Overall, I think that staff member is an effective teacher.

Data analysis

With reference to the three research questions outlined earlier, there were three parts in the analyses. First, descriptive analyses were carried out to look at the general profiles of the students' responses in the two academic years. Second, multivariate analysis of variance (MANOVAs) were conducted to examine the differences of SFQ ratings in all six items across GUR components in 2013–14 academic year with each captioned GUR component (i.e. FS, LCR, CAR, LIPD, and SL) as independent variables and students' SFQ scores on the six items as dependent variables. Third, to explore the differences of SFQ ratings between two academic years overall and by components, independent-samples t-tests were conducted with the academic year of data collection (2012–13 vs. 2013–14) as independent variables and students' SFQ scores on GUR subjects as a whole and on each captioned GUR component as the dependent variables.

Results

Table 2 presents the means and standard deviations of the six common SFQ items of overall GUR and each component in 2012–13 and 2013–14 academic years, respectively, which reflects the profile of student's subjective outcome evaluation about GUR and each component. All ratings were above 3.89 out of a maximum of 5 in 2012–13 and 2013–14, indicating a general satisfaction with GUR subjects by students.

The results of MANOVA demonstrated that SFQ ratings differed significantly in all six items across

Table 2: Descriptive profile of SFQ ratings in overall GUR and different components in 2012–13 and 2013–14 academic years.

	FS		LCR		CAR		LIPD		SL		GUR	
	2012–13	2013–14	2012–13	2013–14	2012–13	2013–14	2012–13	2013–14	2012–13	2013–14	2012–13	2013–14
IA1												
Mean	3.49	3.52	4.01	4.09	3.86	3.91	3.80	3.72	4.19	3.87	3.96	4.01
SD	0.21	0.17	0.35	0.34	0.24	0.31	0.20	0.29	0.20	0.24	0.34	0.36
IA2												
Mean	3.48	3.57	3.98	4.07	3.90	3.90	3.93	3.80	4.21	3.88	3.95	4.00
SD	0.14	0.18	0.37	0.37	0.22	0.33	0.23	0.31	0.17	0.31	0.35	0.38
IA3												
Mean	3.55	3.63	4.02	4.11	3.91	3.93	3.87	3.79	4.25	3.93	3.98	4.03
SD	0.18	0.20	0.34	0.34	0.21	0.29	0.17	0.26	0.15	0.27	0.33	0.35
IA4												
Mean	3.49	3.54	3.94	4.02	3.82	3.82	3.75	3.72	4.11	3.73	3.89	3.94
SD	0.15	0.22	0.39	0.37	0.23	0.32	0.14	0.27	0.24	0.29	0.36	0.38
IIA6												
Mean	3.60	3.70	4.04	4.14	3.94	3.94	4.07	3.98	4.19	3.87	4.01	4.06
SD	0.22	0.34	0.45	0.43	0.27	0.38	0.21	0.33	0.33	0.33	0.41	0.42
IIA7												
Mean	3.73	3.74	4.13	4.22	3.99	3.98	4.14	4.08	3.99	3.87	4.09	4.14
SD	0.21	0.35	0.45	0.42	0.29	0.39	0.22	0.32	0.22	0.35	0.41	0.43

The results for 2012–13 academic year were based on the second semester only in which all GUR subjects were evaluated via eSFQ.

GUR components, Wilks' Λ=0.824, F (24, 3301.41)=7.84, p<0.001. Based on one-way ANOVA results, post-hoc comparisons of SFQ ratings using Bonferroni procedure showed several observations (see Table 3).

a. For IA1 and IA2, ratings in FS were significantly lower than the ratings in other GUR components ($ps \leq 0.022$), and the ratings in LCR were significantly higher than the ratings in other components ($ps \leq 0.019$). The ratings among CAR, SL, and LIPD were not significantly different.

b. For IA3, the rating in FS was significantly lower than the ratings in LCR, CAR, and SL ($ps \leq 0.010$), and the rating in LCR was significantly higher than the ratings in other GUR components ($ps \leq 0.018$). The ratings among CAR, SL, and LIPD were not significantly different.

c. For IA4, the rating in FS was significantly lower than the ratings in CAR and LCR ($ps \leq 0.010$) and the rating in LCR was higher than the rating in other GUR components ($ps < 0.001$). The ratings among CAR, LIPD,

and SL, as well as the ratings among LIPD, SL, and FS were not significantly different.

d. For IIA6 and IIA7, the rating in FS was significantly lower than the ratings of other GUR components except for SL ($ps \leq 0.033$), and the rating in LCR was significantly higher than the ratings of other GUR components except for LIPD ($ps \leq 0.025$). The ratings among LIPD, SL, and CAR; the ratings between LCR and LIPD; and the ratings between FS and SL were not significantly different.

Regarding the cross-year comparison of SFQ scores on GUR as a whole, the findings can be seen in Table 4. There was a statistically significant increase in three items about students' learning experience: IA1, t (1356)=2.02, p<0.05; IA2, t (1356)=2.18, p<0.05; IA3, t (1355) =2.43, p<0.05, and one item on students' perceived teaching of the staff, IIA6, t (1366)=2.01, p<0.05. The results indicated that there were good improvements in terms of students' subjective evaluation about GUR subjects in the second year of GUR implementation.

With regard to each GUR component, the respondents in 2013–14 reported a significantly higher mean rating for LCR on all six SFQ items than did the respondents in 2013–14 reported (all $ps \leq 0.01$). Regarding SL, there were statistically significant decreases in four SFQ items measuring students' learning experience (all $ps \leq 0.001$) and one item on teaching of the staff (IIA6), t (1366)=2.01, p<0.05. No significant difference in SFQ ratings was found in FS, CAR, and LIPD between the two academic years. The findings suggested that there was cross-time and cross-subject difference in students' satisfaction with GUR subjects.

Discussion

The present study attempted to investigate student's subjective outcome evaluation on the effectiveness of GUR

Table 3: Results of MANOVA on SFQ ratings across GUR components in 2013–14 academic year.

DV	df	df error	F	GUR components
IA1	4	951	33.68[a]	LCR>CAR/LIPD/SL/FS CAR/LCR/SL>FS CAR>LIPD
IA2	4	951	21.24[a]	LCR>CAR/LIPD/SL/FS CAR/LCR/SL>FS CAR>LIPD
IA3	4	951	25.53[a]	LCR>CAR/LIPD/SL/FS CAR/LCR/SL>FS
IA4	4	951	24.20[a]	LCR>CAR/LIPD/SL/FS CAR/LCR>FS
IIA6	4	951	14.80[a]	LCR>CAR/SL/FS CAR/LCR>FS
IIA7	4	951	20.93[a]	LCR>CAR/SL/FS CAR/LCR>FS

[a]p<0.001.

Table 4: Comparison on SFQ ratings in GUR subjects between two academic years overall and by components.

	FS		LCR		CAR		LIPD		SL		GUR	
	t	p	t	p	t	p	t	p	t	p	t	p
IA1	0.39	0.70	2.99	0.00	1.15	0.25	−1.21	0.23	−3.01	0.01	2.02	0.04
IA2	1.62	0.11	3.42	0.00	−0.03	0.98	−1.48	0.15	−2.49	0.02	2.18	0.03
IA3	1.42	0.16	3.34	0.00	0.70	0.48	−1.26	0.22	−2.74	0.01	2.43	0.02
IA4	0.76	0.45	3.01	0.00	0.03	0.98	−0.43	0.67	−3.08	0.00	1.89	0.06
IIA6	0.82	0.42	3.02	0.00	−0.10	0.92	−1.18	0.24	−2.26	0.03	2.01	0.04
IIA7	0.12	0.90	2.73	0.01	−0.37	0.71	−0.78	0.44	−0.81	0.42	1.61	0.11

The results for 2012–13 academic year were based on the second semester only in which all GUR subjects were evaluated via eSFQ.

in the academic years of 2012–13 and 2013–14. There are several strengths of this study. First, a large sample size was involved. Second, longitudinal data over 2 years were collected. Third, validated objective measures of course evaluation were used. Fourth, differences in students' evaluative ratings on GUR subjects were examined across both different GUR domains and years. This is the first known scientific study examining the views of the students to general education subjects under the 4-year undergraduate curriculum using a standardized evaluation instrument since the inception of the new 4-year undergraduate curriculum in Hong Kong.

Regarding the first research question, the results showed that the students generally held positive perceptions of GUR subjects in terms of their learning experiences as well as the teaching of the instructor. Basically, positive responses to the items in the first part and second part were identified. These findings are generally in line with the qualitative evaluation findings on the implementation of GUR at PolyU [13]. In a study based on students focus groups (n=62), results showed that the participants generally had positive perceptions of GUR subjects, the teachers, and the benefits of the subjects [13]. Another study based on teacher focus groups (n=20) revealed similar findings that the subject contents, teaching staff, and the effectiveness of GUR subjects were perceived in a positive manner by the teachers who were an important stakeholder in GUR implementation [14]. In an objective outcome evaluation study using the Collegiate Learning Assessment Plus (CLA+) at PolyU, results generally showed that sophomores performed better than did freshmen after 1-year GUR study [15]. The present findings are also consistent with the evaluation findings associated with a subject entitled "Tomorrow's Leaders" under the GUR domain of LIPD (Leadership and Intrapersonal Development) [16, 17].

Regarding differences of SFQ ratings among different GUR components in 2013–14 academic year, there are several observations deserving our attention. First, students had different views on different GUR components. There are several possible explanations for the observed variations. The first explanation is that students' ratings in evaluating university course teaching were greatly influenced by different subject matters [12, 18, 19]. The second explanation is that the effectiveness of teaching varies across different instructional methods. For example, as language teaching is usually more interactive, this may lead to better student evaluation as compared to subjects taught with less interactive methods. The third explanation is that students in different instructional groups may have distinct perceptions about teacher's teaching skills

and effectiveness [20]. It is also possible that they have different attitudes toward learning, academic goals and skills, as well as preferred learning styles [21]. All these factors may affect students' subjective evaluation about different GUR components.

Second, the ratings of FS appeared to be lower than other GUR components. There are three possible explanations for this observation. First, since FS is compulsory for all first-year students, it is possible that the freshmen could not understand the intended learning outcomes immediately. Students' focus group interviews [13] showed that FS subjects were viewed more theoretical and conceptual while less practical by students since these subjects focused more on the introduction of the history of one's majors or broad discipline. Students perceived the lecture contents of FS subjects as "boring" and "hard to understand" and the workload in these subjects (e.g. long essay writing) was "quite heavy" and "hard" for first-year students. Second, class size may account for the findings because the class size of FS subjects was usually much larger than other GUR components. This would hinder effective teacher-students interaction and make students feel difficult to concentrate [12].

The cross-year comparison on SFQ results further revealed that there were general improvements in students' ratings on GUR subjects in terms of students' learning experience and teaching of the staff after 1 year of GUR implementation. This suggests that with the consolidation of the GUR curriculum at PolyU, GUR subjects have become more and more attractive to students. Meanwhile, it should be noted that there was variation in the temporal changes across different GUR components. On the one hand, students in 2013–14 academic year gave higher ratings on GUR subjects as a whole in all six items than the respondents in 2012–13 academic year. This positive change may be due to the fact that several areas in the LCR arrangements and policies were refined in 2013–14 academic year.

On the other hand, there was a general decrease in SFQ ratings in SL subjects. There are several explanations for this observation. First, the number of classes and students in 2013–14 academic year were much higher than those in 2012–13 academic year. As large class was associated with less favorable course evaluation, this may explain the observed findings. Second, as SL subjects required students to provide direct service to different communities, different composition of client groups to which students provide service in the 2 years may explain different levels of satisfaction toward the subjects by the students. In 2013–14 academic year, more SL subjects involving non-local service recipients. Students may experience more difficulties in adapting to the culture and encounter more

difficulties in providing the service, which may adversely affect students' SFQ scores.

Despite the pioneer nature of this study and its unique features, there are several limitations of this study. First, the response rate was not high (i.e. overall 54% in academic year of 2013–14). However, similar response rates were observed in other studies across the world. Second, the number of common items in the current version of SFQ was limited. As academic departments usually add other discipline-specific items in the SFQ, there is not much space to include more common items. Hence, assessment of multiple dimensions in the teaching and learning process is not possible. Third, only students' perspective was assessed in the SFQ. As teachers are important stakeholders in the evaluation exercise, it would be illuminating if the views of the teachers can be collected. Fourth, qualitative comments in the SFQs were not analyzed in this study. If resource permits, such qualitative data should be analyzed and reported in future.

Before the effects of GUR on students' objective outcome indictors can be observed, subjective outcome evaluation offers timely and important information about the implementation of the curriculum, which contributes to the improvement of the GUR subjects [22]. In the present study, subjective outcome evaluation findings based on SFQ results provide important pointers for the implementation of GUR subjects at PolyU. The results suggest that the GUR implementation was successful, although further work should be done to understand cross-domain and cross-time variations in the course evaluation ratings.

References

1. Osterlind SJ. A national review of scholastic achievement in general education: how are we doing and why should we care? Washington: ERIC Clearinghouse on Higher Education, 1997.
2. Witt DH. Internationalization of higher education in the United States of America and Europe. Boston: Center for International Higher Education and the Program in Higher Education, 2009.
3. Marginson S. Education and public policy in Australia. New York: Cambridge University Press, 1993.
4. AAC&U. Trends and emerging practices in general education. In: Wehlburg CM, editor. Integrated general education: new directions for teaching and learning. San Francisco: Jossey-Bass, 2010.
5. Peterson S. GED success: 2003. Lawrenceville, NJ: Peterson's Guide, 2003.
6. Boyer EL, Arthur L. A quest for common learning: the aims of general education. A Carnegie Foundation Essay. Washington, DC: Carnegie Foundation for the Advancement of Teaching, 1981.
7. South China Morning Post. SCMP opinion leader survey [Internet]. Hong Kong. 2006. Available at: http://www.scmp.com/article/562719/opinion-leaders-survey. Accessed on June 20, 2014.
8. University Grant Committee. Hong Kong higher education to make a difference to move with the times, 2004. Available at: http://www.ugc.edu.hk/eng/doc/ugc/publication/report/policy_document_e.pd. Accessed on June 20, 2014.
9. Chen YH. The study of core competency learning and outcome assessment in general education. Paper presented at the proceeding at General Education and University Curriculum Reform: an International Conference in Hong Kong (pp 99–104). Hong Kong: City University of Hong Kong and The Hong Kong America Center, 2012.
10. Kember D, Leung DY. Establishing the validity and reliability of course evaluation questionnaires. Assess Eval High Educ 2008;33:341–53.
11. Educational Development Centre: The PolyU Student Feedback Questionnaire (SFQ) System. The Hong Kong Polytechnic University 2012. Available at: URL. http://edc.polyu.edu.hk/sfq-collect.htm. Accessed on June 21, 2014.
12. Kwan K. How fair are student ratings in assessing the teaching performance of university teachers? Assess Eval Higher Educ 1999;24:181–95.
13. Shek DT, Yu L, Wu FK, Chai WY. General university requirements at The Hong Kong Polytechnic University: evaluation findings based on student focus groups. Assess Eval High Educ. Epub 2014 October 1. Available from: http://dx.doi.org/10.1080/02602938.2014.960362.
14. Shek DT, Yu L, Wu FK, Chai WY. Teachers' views of a new general education program in Hong Kong: a qualitative study. Int J Adolesc Med Health, in press.
15. Shek DT, Yu L, Wu FK, Ng CS. General education program in a new 4-year university curriculum in Hong Kong: findings based on multiple evaluation strategies. Int J Disabil Hum Dev 2015;14:377–84.
16. Shek DT, Ma CM. Do university students change after taking a subject on leadership and intrapersonal development? Int J Disabil Hum Dev 2014;13:451–6.
17. Shek DT, Yu L. Post-course subjective outcome evaluation of a subject on leadership and intrapersonal development for university students in Hong Kong. Int J Disabil Hum Dev 2014;12:193–201.
18. Feldman KA. Course characteristics and college students' ratings of their teachers: what we know and what we don't. Res High Educ 1978;9:199–242.
19. Marsh H. Students' evaluations of university teaching: dimensionality, reliability, validity, potential biases, and utility. J Educ Psychol 1984;76:707–54.
20. Smith RA, Cranton PA. Students' perceptions of teaching skills and overall effectiveness across instructional settings. Res High Educ 1992;33:747–64.
21. Cashin WE. Students do rate different academic fields differently. In: Theall M, Franklin J, editors. Student ratings of instruction: issues for improving practice (new directions for teaching and learning). San Francisco: Jossey Bass, 1990.
22. Shek DT. Is subjective outcome evaluation related to objective outcome evaluation? Insights from a longitudinal study in Hong Kong. J Pediatr Adolesc Gynecol 2014;27:S50–6.