







Subjective Outcome Evaluation of Service-Learning by the Service Recipients: Scale Development, Normative Profiles and Predictors

D. T. L. Shek¹  · Z. Yang¹  · C. M. S. Ma¹  · C. W. Y. Chai¹ 

Accepted: 6 August 2020 / Published online: 18 August 2020

© The Author(s) 2020

Abstract

This study validated a subjective outcome evaluation scale based on the perceptions of service recipients and examined the normative profiles as well as correlates of client satisfaction under a corporate-university-community Service-Learning (SL) programs, where university students conducted the SL programs with high school students as the service recipients. Based on data collected over three years from high school service recipients ($N=1854$) who responded to a subjective outcome evaluation measure which assesses perceived program content, program implementers and benefits, exploratory and confirmatory factor analyses supported the three-factor model, and the three subscales showed good factorial validity, convergent validity, concurrent validity and internal consistency. Most of the service recipients had positive perceptions of the different aspects of the program. Regarding correlates of client satisfaction, the three aspects of satisfaction all had predictive effects on the overall client satisfaction. Client satisfaction ratings also differed significantly across different activities and grades.

Keywords Service-learning · Subjective outcome evaluation · Service recipients · Chinese · University students

✉ D. T. L. Shek
daniel.shek@polyu.edu.hk

Z. Yang
zi.yang@polyu.edu.hk

C. M. S. Ma
cecilia.ma@polyu.edu.hk

C. W. Y. Chai
wenyu.chai@polyu.edu.hk

¹ Department of Applied Social Sciences, The Hong Kong Polytechnic University, Hong Kong, People's Republic of China

1 Introduction

Ehrlich (1996) conceptualized service-learning (SL) as pedagogies which “link community service and academic study so that each strengthens the other. The basic theory of service-learning is Dewey’s: the interaction of knowledge and skills with experience is key to learning” (p. xi). Operationally, Bringle et al. (2006) pointed out that SL is “a course-based, credit-bearing educational experience” (p. 12) where students plan and organize service activities that can satisfy community needs as well as reflect on SL experience to consolidate the course experience, appreciate the related knowledge, and build up personal values and sense of social responsibility. According to Furco (1996), although there are different forms of service programs, service-learning benefits not only those who provide service but also those who receive service and emphasizes the occurrence of both service and learning. Felten and Clayton (2011, p. 78) outlined a conceptual framework for SL with three components (“academic material, relevant service, and critical reflection”) involving three partners (“students, faculty/staff, and community members”) resulting in three learning goals (“civic learning, personal growth, and academic learning”).

Different theories have been proposed with reference to the goals, process and outcomes of SL. Regarding the goals of SL, there is a continuum ranging from “charity-based” SL to “social justice” emphasis (Einfeld and Collins 2008; Kahne and Westheimer 1996; Morton 1995). In “charity-based” SL, the focus of SL is to provide community service for the needy with integration with academic learning. On the other extreme, the “social justice” model attempts to help students understand the ways to “transform” the society by understanding and possibly removing inequality and injustice. Based on critical theory principles, SL should help students understand the underlying issues of inequality, exploitation, and injustice, hence exploring ways of social reforms (Boyle-Baise and Langford 2004). A middle-of-the-road approach is “civic education”, “project” or “civic engagement” approach focusing on civic and community engagement where students learn to understand the community and its needs, and build up partnership with the community to solve community problems. Under this approach, SL requires students to learn empathy and develop passion about the community, focus on how to solve community problem and/or satisfying the community needs via application of the subject knowledge, and strengthen the linkage between the subject matter and the community needs. The civic engagement approach of SL has been commonly endorsed by researchers and practitioners in higher education.

Regarding the teaching and learning in SL, Giles and Eyerl (1994) reviewed the thoughts of John Dewey and discussed how these thoughts are related to experiential learning, critical thinking, citizenship, and community engagement. Cone and Harris (1996) stated that the theories of John Dewey, Paulo Freire, and David Kolb had been used as theoretical bases of SL and argued that additional perspectives based on cognitive psychology and social theory could help to enrich SL theories. Ash and Clayton (2004) also proposed a model of “critical reflection” in SL. As far as the impact of SL on developmental outcomes is concerned, positive youth development (PYD) approach can be used to understand the impact of SL on youth development (Shek et al. 2019a). With reference to the developmental assets proposed by the Search Institute (Benson et al. 2011), SL can help to promote the assets of “community valued youth”, “youth as resources”, “service to others”, “youth programs”, “positive values” such as

“caring, responsibility, equality and social justice”, social competencies, and positive identity. Similarly, SL can promote youth development based on the 5C/6C model of Richard Lerner (Lerner et al. 2011, p. 9) where SL can help to promote “connection”, “competence”, “confidence”, “character”, “contribution”, and “care” of the service recipients. Previous studies have showed that SL subjects were able to promote PYD attributes in students (e.g., Shek et al. 2019b).

As a form of innovative pedagogy, Service-Learning (SL) has sparked much interest among researchers, institutions, and policy makers (Billig and Furco 2002). Premised on experiential learning, SL provides students with a unique opportunity that integrates academic learning with structured service activities, and connects educational objectives with community needs (Deeley 2010). This pedagogy has been embraced by higher education sectors because it goes beyond the academic mission for students and encourages their civic engagement. Bringle and Hatcher (1996) pointed out that many national associations within the United States endorsed the value of SL (e.g., Campus Compact, American National Association for Higher Education, and Partnership for Service-Learning). In other parts of the world, many universities have incorporated credit-bearing SL programs into their curricula, including Canada (Kricsfalussy et al. 2016), Europe (European Observatory of Service-Learning in Higher Education 2020), Australia (Birkeck 2012), Africa (Pacho 2019), and Asia (e.g., Anorico 2019; Lin et al. 2014). There are also international associations promoting SL in the sector of higher education, such as the International Association of Research on Service Learning and Civic Engagement (IARSLCE) and Talloires Network.

With the proliferation of SL programs, there has been a steady increase in attempts to evaluate the effectiveness of such programs. There are two observations based on the existing evaluation literature. First, existing studies generally showed the positive impact of SL on the students who provided the service (i.e., service providers). Two decades ago, Billig (2000) concluded that “research, while limited, finds that students who help others help themselves academically and socially” (p. 1). In the past two decades, meta-analytic studies supporting the beneficial effects of SL on student development have gradually increased. Conway et al. (2009) reviewed 103 studies and showed that service learning promoted academic outcome (moderate effect size), personal and citizenship outcomes (small effect size) and social outcomes (small to moderate effect size). In a meta-analytic study based on 62 studies involving 11,837 participants, Celio et al. (2011) showed that students participating in SL performed better on “attitudes toward self, attitude toward school and learning, civic engagement, social skills and academic performance” (p.174–175). Besides, use of reflection was related to better outcome performance. Another meta-analytic review of 40 studies involving 5495 participants found that SL promoted understanding of community needs, personal understanding, and cognitive competence, although some moderating factors were also identified (Yorio and Ye 2012). There are also some recent studies on the impact of SL on civic engagement: Mason and Dunens (2019) reported that SL could promote student learning and achieving the foundational knowledge in public health; Trigos-Carrillo et al. (2020) highlighted the ability of service learning experience to transform students’ cultural humility and affective understanding in post-conflict setting. However, it is noteworthy that there are cautions about the “dark side” of SL (Eby 1998; Morin 2009) such as whether it promotes social justice or social stagnation. Overall speaking, there is support for the positive impact on different

domains of student development, although some researchers suggested that we should further clarify the moderators such as instructional design and reflection.

The second observation is that different evaluation methods have been used to assess the impact of SL. Primarily, objective outcome evaluation using psychosocial assessment tools has received much scholarly attention. Drawing on pre-experimental and experimental designs, many studies of objective outcome evaluation have demonstrated positive changes of the service implementers in psychosocial or behavioral domains (e.g. McBride et al. 2014; Strage 2000). For instance, results of a quasi-experimental research study revealed that, compared with the control group (seventh-grade students not joining the program), the experimental group (seventh-grade students who joined a SL program) exhibited significant improvement in civic awareness and academic performance. In addition, qualitative evaluation studies focusing on the subjective experiences and different views of various stakeholders have also been carried out (e.g., Meili et al. 2011; Rubio et al. 2018).

Besides objective outcome evaluation and qualitative evaluation, subjective outcome evaluation has also been widely used to examine the perceptions of service recipients. Grounded in the approach of client satisfaction, subjective outcome evaluation has been extensively used in the fields of education, health settings, and social service (Shek and Ma 2014). In the field of education, subjective outcome evaluation in the form of course evaluation questionnaire has been widely used in higher education (e.g., Kember and Leung 2008; Spooren et al. 2007). In the field of human and social services, program participants are commonly asked about their views of the program, implementer, and benefits after program completion (Burke and Bush 2013; Fraser and Wu 2016; Hsieh 2006; Lee et al. 2018). In the health settings, health professionals usually ask the patients about their satisfaction with the service received and the perceived benefits (Perneger et al. 2020).

Despite criticisms on subjective outcome evaluation such as biased results (O'Neal 1999), client satisfaction survey offers immediate access to different stakeholders' firsthand experiences in the programs (Shek and Ma 2014). Besides, it is easy to implement and can be fitted into routine practice of the implementers. In addition, empirical studies have demonstrated a strong correlation between subjective and objective measures, confirming that subjective measure can be used as a reliable subjective outcome indicator in different fields (Shek 2014; Sun and Richardson 2016).

Subjective outcome evaluation is commonly used in the field of SL. For example, Lee et al. (2018) showed that participants of a SL program serving rural community in the US perceived their experiences positively. They reported greater interest in applying knowledge into practice and having more clarity on their career development. In the education field, subjective outcome evaluation facilitates students' voices to be heard (Chen and Hoshower 2003). Their perceptions can bring great insight into teaching evaluation, as they provide meaningful feedback to the teaching process which can inform the course design and improvement. Maccio and Voorhies (2012) used the subjective outcome evaluation approach to understand the perceptions of students joining SL programs. Actually, routine course evaluation using the subjective outcome evaluation approach is commonly used in the higher education context.

A closer review of the literature in the field of SL program evaluation reveals several research gaps. First, existing research focuses primarily on the impact of SL on service providers, particularly on university students as service providers. It has been well-

documented in the literature that participation in SL benefits service providers in terms of their civic engagement (Kiely 2004), cognitive development and academic performance (Kearney 2004), positive youth development (Chung and McBride 2015), etc. Unfortunately, research on how service recipients benefit from SL, particularly those provided by university students, remains scant. Theoretically, reciprocity and mutual benefits are regarded as a defining feature of SL (Faber 2017). Therefore, knowledge of service recipients' perceptions and experiences contributes to the theoretical understanding of the extent to which reciprocal exchanges of SL are achieved. Practically, service recipients' perspectives provide an honest assessment of the effectiveness of a SL program (d'Arlach et al. 2009). Their voices should not be ignored, as they shed light on the critical concerns of SL such as whether service recipients truly need those services (Weah et al. 2000).

The second research gap is that few studies have adopted validated tools in subjective outcome evaluation in the context of SL. For instance, using a self-report questionnaire and descriptive statistics, Ma et al. (2018) found an overall positive response towards a SL subject in terms of curriculum content, lecturers, and perceived benefits. In a survey study in a US university, Chen (2015) examined the views of the service providers and community partners by asking them to complete a 64-item online questionnaire. Although such studies are helpful, it is unclear whether the related instruments have been validated or not. In addition, there has been no known validated subjective outcome evaluation tool for the service recipients in the context of SL implemented in the higher education sector.

Third, the lack of validated subjective outcome evaluation tools on SL is probably the result of a lack of well-articulated theoretical model on subjective evaluation of SL. Although there are well-conceived conceptual models on course evaluation (e.g., Spooen et al. 2007), they are rarely adopted in the context of SL. Similarly, although there are many tools in the context of social welfare (Fraser and Wu 2016), they have not been used to evaluate SL projects. Obviously, we can borrow concepts from education and social services to guide the assessment of subjective outcomes in SL. In the area of education, it is a common practice to assess the students' perceptions of the subject they have taken (such as subject content and course design), instructor (such as teacher caring about students and teacher involvement) and effectiveness (such as whether students find the subject beneficial to their learning). In the area of social service, quality of program, workers, and benefits are commonly focused upon (Fraser and Wu 2016). For example, this conceptual framework has been used in a positive youth development program (Project P.A.T.H.S.) in Hong Kong. With a conceptual framework with three dimensions, findings showed that the 36-item assessment tool (program quality, worker quality and benefits with 10, 10 and 16 items, respectively) possessed excellent construct validity (Shek and Ma 2014). This conceptual framework was also used to assess students' perception of a service leadership subject, including subject (10 items), teachers (10 items) and benefits (18 items) (Shek and Liang 2015). In this study, these three comprehensive dimensions were employed to evaluate the service recipients' views of SL in the area of program, workers (university students providing the service), and benefits of the program.

Fourth, in contrast to the development of SL in the Western contexts, SL remains an emerging concept in the Chinese context (Xiang and Luk 2012). With the well-documented benefits of SL in other contexts, the multi-faceted pedagogy may also

bring new ideas and philosophies to the traditional classrooms and benefit the huge number of adolescents in China. Moreover, as China entails unique culture and collectivistic values, SL in China may have distinct features of its own (Guo et al. 2016). For example, core in Confucianism and Buddhism is the concept that people are encouraged to show kindness and offer help to those in need. Therefore, SL programs in Chinese contexts may incorporate these traditional philosophies in their value and practice. Based on these reasons, a systematic study of SL with a re-contextualization to the Chinese context is needed.

The current study focused on a SL program conducted by university students in Hong Kong. SL has received increased attention in Hong Kong (Shek et al. 2019a). As a pioneer in SL, *** University has incorporated a series of SL subjects into the curriculum since the 2012/13 academic year. Among those subjects, two subjects entitled “Promotion of Children and Adolescent Development” and “Service Leadership through Serving Children and Families with Special Needs” have been developed. In these two subjects, students are offered the opportunity to implement services in a pioneer service project entitled Project *WeCan*. As described in the website of Project *WeCan*, the project was launched in 2011 which is “a Business-in-Community initiative providing students who are disadvantaged in learning with opportunities and care to empower them for pursuing higher studies and future careers. Through diversified programmes, WeCan strives to enhance students’ communication skills and basic competence, increase their exposure, cultivate their character and develop their common sense, and foster their innovativeness and creativity” (<http://www.projectwecan.org/about-us/overview>). Essentially, the project attempts to help secondary school students from lower socio-economic backgrounds, as findings showed a positive association between scholastic performance and socio-economic milieu (e.g. Becker and Luthat 2002; Considine and Zappalà 2002). As such, Project *WeCan* aims to empower those under-privileged students by enhancing their learning and career opportunities.

According to the American Psychological Association, civic engagement is “individual and collective actions designed to identify and address issues of public concern” (<https://www.apa.org/education/undergrad/civic-engagement>). As such, Project *WeCan* can be regarded as a civic engagement initiative. Besides, fostering collaboration between the business sector, higher education sector, and community sector (i.e., high schools) can help to bring more resources for the participating schools and students, such as social capital in the school and the community (Campbell 2000). For example, each participating secondary school collaborates with a corporate sponsor to provide financial aid (i.e., financial capital) and a university to offer academic support (i.e., social capital). Through different activities (including Service-Learning) based on the community engagement approach, it is expected that the service recipients (i.e., high schools students) will develop better with respect to the objectives of Project *WeCan*. From 2016/17 to 2018/19 academic years, students of *** University had provided customized service activities to 12 local secondary schools, including language tutorials, educational camp, career talk, campus visit, etc.

Regarding the goals of these two SL subjects, we adopted a civic engagement approach to help students to develop civic responsibilities and help the community to solve its problems or meet its needs. Adler and Goggin (2005) defined civic engagement as “the ways in which citizens participate in the life of a community in order to improve conditions for others or to help shape the community’s future” (p. 236).

Hence, we can regard these two SL subjects in Project *WeCan* as civic engagement programs. When we examine the intended learning outcomes of these two subjects, there are four generic learning outcomes. After taking the SL subjects, it is expected that the students are able to: a) apply the subject knowledge to meet the needs of the community via SL projects; b) reflect on citizenship and civic responsibilities; c) develop empathy; and d) appreciate the linkage between academic subjects and community needs. Obviously, these generic learning outcomes are expected to be the result of participation in the community to solve community issues (i.e., civic engagement). In terms of pedagogies in these two subjects, experiential, collaborative, and reflective learning approaches are used. It is expected that through experiential learning, collaborative learning, and reflective learning, university students providing the service and the high school students receiving the service would benefit from the SL activities.

To systematically fill the above-mentioned research gaps in the field of SL, this study attempted to validate a subjective outcome evaluation scale for the service recipients (i.e., high school students) under the Project *WeCan*, examine their perceptions, and identify the correlates of their subjective perceptions.

For the psychometric properties of the tool assessing subjective perceptions of the service recipients, we addressed the following research question:

Research Question 1: Does the subjective outcome evaluation scale possess acceptable psychometric properties, including construct validity (factorial validity and convergent validity), concurrent validity, and internal consistency? Regarding factorial validity, it was predicted that there are three aspects in the scale, including perceived content of program, service providers, and benefits (see Fig. 1). Concerning convergent validity, it was hypothesized that significant relationships exist between: a) perceived content of program and perceived qualities of service provider (Hypothesis 1a); b) perceived content of program and perceived benefits of program (Hypothesis 1b); c) perceived qualities of service provider and perceived benefits of program (Hypothesis 1c). For concurrent validity, it was expected that subjective outcome evaluation scale scores would be positively correlated to: a) the participants' willingness to participate in the program again (Hypothesis 2a), b) the participants' willingness to recommend the program to others (Hypothesis 2b) and c) overall satisfaction (Hypothesis 2c).

Regarding the profiles and predictors of the subjective outcome evaluation findings, three research questions were raised:

Research Question 2: What are the service recipients' perceptions of the services?

Research question 3: What are the predictors of the overall satisfaction with the program? Based on the previous studies that perceived qualities of service providers and content predicted the perceived satisfaction with the program (e.g. Shek and Sun 2014), it was hypothesized that perceived content, instructor qualities, and perceived benefits would predict the overall satisfaction of the program (Hypotheses 3–5).

Research question 4: Does subjective outcome evaluation differ across activity types (e.g. language workshops, educational camp) and grade levels (junior grades versus senior grades)? As adolescent prevention studies suggest that participants' positive changes were related to the amount of exposure to the intervention (i.e., dosage effect, Ferrer-Wreder et al. 2010; Reyes et al. 2012), it was predicted that participants' client satisfaction would be more positive if they joined more than one program (Hypothesis 6). Besides, as students in lower grades had more positive views of positive youth development programs than did students in the senior grades because of novelty effect

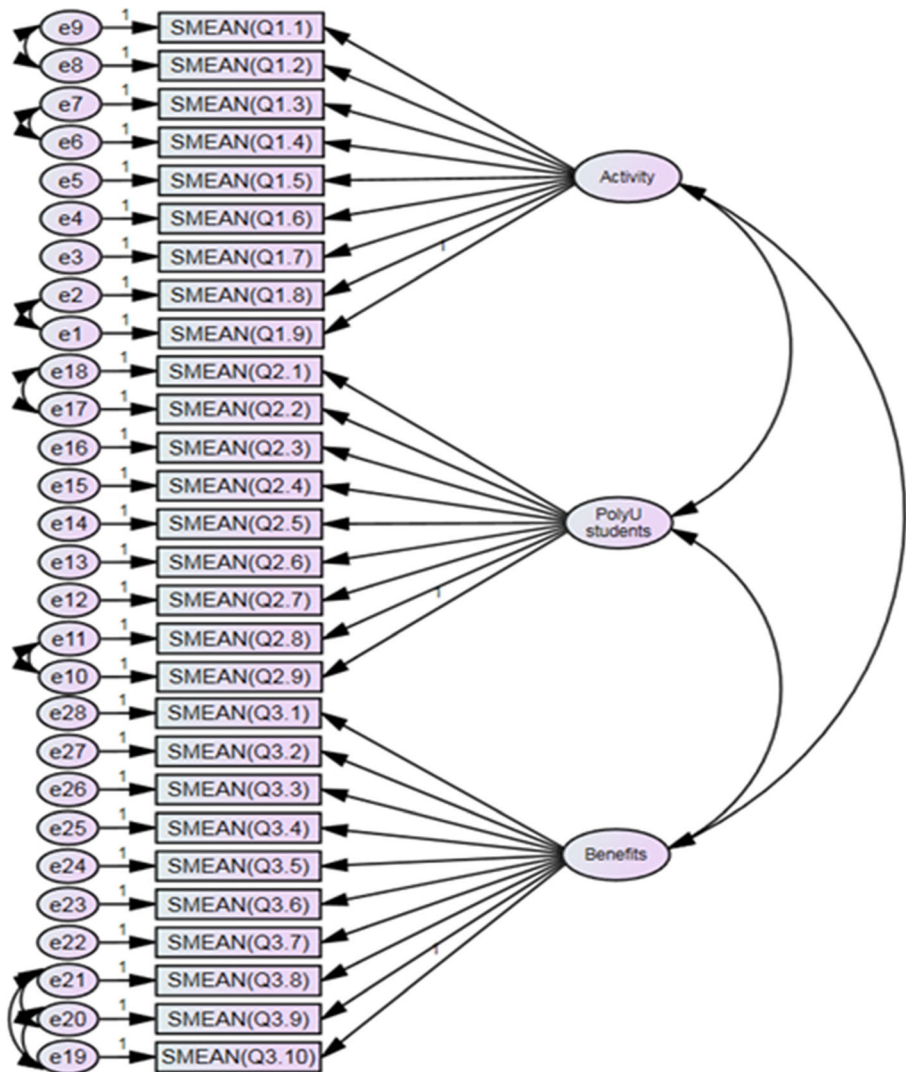


Fig. 1 Hypothesized Structure of SOES-SR

(Shek and Law 2014), it was expected that students in the junior grades would have relatively better subjective perceptions than did senior grades (Hypothesis 7).

2 Methodology

The current study was part of a larger project that examined the effectiveness of the SL subjects implemented in Project *WeCan* from 2016/17 to 2018/19 academic years. Data were collected from different stakeholders, including secondary school students (service recipients), secondary school teachers, students (service implementers) from *** University, teachers from *** University, and partners from The Wharf (Shek et al. 2019a). The present study focused on the service recipients' perceptions of the program.

2.1 Participants and Procedures

In these three academic years, a total of 18,250 secondary school students (3903 in 2016/17; 8621 in 2017/18; 5726 in 2018/19) participated in the SL services provided by the students of *** University. In total, 505 students have enrolled in the two subjects and provided services to 12 secondary schools across Hong Kong. Approximately 1500 h of service were provided to those secondary school students. Where appropriate and practical, the subjective outcome evaluation questionnaire was administered to the students in a voluntary manner at the end of the semesters. Research assistants explained the purposes and procedures of the research, and assured the participants of anonymity, confidentiality, and voluntary participation. In total, 1854 questionnaires (2016/17: $n = 471$; 2017/18: $n = 498$; 2018/19: $n = 885$) were collected.

2.2 Instrument

Modeling after similar validated measures in the field (e.g. Shek and Ma 2014; Shek and Sun 2014), we designed a questionnaire entitled “Subjective Outcome Evaluation Scale - Service Recipients (SOES-SR)”, with 28 closed-ended items assessing respondents’ views towards the service content/activities (9 items), implementers (9 items), and perceived benefits (10 items). Students were asked to rate on a 6-point Likert scale, in which “1” stands for “strongly disagree”, and “6” represents “strongly agree”. With particular reference to the objective of Project *WeCan*, it is obvious that the items on perceived benefits of the program are aligned with the objectives of Project *WeCan*, such as promotion of basic competence, aspirations, holistic development as well as increasing exposure.

Three additional items were included in the questionnaire, including: 1) willingness of the respondent to participate again, 2) willingness of the participant to suggest friends to participate, and 3) overall satisfaction towards the project. In this study, we used them to examine the concurrent validity of the 28-item SOES-SR. While we acknowledge the limitation of using these three items as external criteria, we used them for three reasons. First, based on our conceptual model (Shek and Ma 2014; Shek et al. 2014) and other conceptual models (Larsen et al. 1979), perceived program quality, instructor quality and program effectiveness are determinants of overall satisfaction indexed by whether one would participate in similar programs, whether one would recommend others to join, and global satisfaction. Second, in many subjective outcome evaluation scales, assessment of perceived quality of specific domains (such as quality of the program and workers) and global satisfaction were regarded as separate but theoretically related constructs (Bahia et al. 2000; Spiro et al. 2009). Third, previous studies showed that there was significant correlation between subjective outcome evaluation scores and objective outcome evaluation scores (Shek 2010; Shek 2014). In the questionnaire, we also collected information on activity type and service recipients’ grade level.

2.3 Data Analytic Strategy

The psychometric properties of the scale were assessed through a systematic approach, including exploratory factor analyses (EFA), confirmatory factor analyses (CFA), and

multigroup CFA for measurement invariance tests. To perform these analyses, the whole sample was divided into two subsamples based on “odd” and “even” case numbers (Subsample 1: $n = 927$; Subsample 2: $n = 927$). First, we conducted EFA based on Subsample 1 via SPSS Statistics 25.0 to detect the factorial structure of the scale. Second, we performed CFA based on Subsample 2 via AMOS 25.0 to confirm the factor structure obtained from EFA (Fig. 1). Third, multiple group CFA was conducted based on Subsample 2 via AMOS 25.0 to examine the factorial validity of the scale. Specifically, a series of factorial invariance tests were conducted on a series of nested CFA models across “odd”-and-“even”-case-number groups in Subsample 2, including configural invariance, “metric invariance” (i.e., weak factorial invariance), “scalar invariance” (i.e., strong factorial invariance), equality of factorial variance/covariance, and “strict factorial invariance” (Gregorich 2006). Multiple indices were adopted to indicate the goodness of model fit, including “comparative fit index (CFI)”, “root-mean-square error of approximation (RMSEA)”, “Tucker-Lewis Index (TLI)”, and “standardized root-mean-square residual (SRMR)”. The following criteria of goodness-of-fit indices were used: for CFI and TLI, a value ≥ 0.90 suggests an adequate model fit (Bentler and Bonett 1980); the value of SRMR ≤ 0.10 and RMSEA ≤ 0.80 respectively imply an acceptable fit to the data (Hirsh 2010). Given the large sample size, the changed values in CFI ($\Delta\text{CFI} \leq 0.01$) were adopted as the main indicator in invariance test (Cheung and Rensvold 2002). We also examined the inter-relationship among the three subscales to measure the convergent validity of the scale. Besides, we examined the relationship between the total scale score and three external criteria (i.e., whether the participants would like to join again, whether the participants would like to recommend friends to take the program, and overall satisfaction). Finally, internal consistency was measured via SPSS 25.0 to examine reliability of the scale.

Concerning the profiles of the scale, descriptive statistics were performed to analyze service recipients’ perception of their experiences (Research Question 2). Regression analyses were also performed to assess how well each of the subjective outcome dimensions predicts the overall satisfaction (Research Question 3). In addition, one-way ANOVAs were used to test the students’ perceptions in relation to activity types and their grade levels respectively (Research Question 4). For activity types, “Language workshops” include English tutorials, Korean workshops, etc., and “Workshop on STEM, art or craft” refers to workshops that aim to improve science literacy, critical thinking ability or creativity. “Two activities” refer to those attended two service activities (See Table 6). Grade levels were grouped into junior high school (Secondary 1 to 3) level and senior high school (Secondary 4 to 6) level.

3 Results

3.1 Psychometric Properties of the Scale

To examine whether the three facets of the 28-item scale and the dimension of the three external criteria existed, an EFA was performed based on Subsample 1. As the data did not meet the assumption of multivariate normal distribution, Principal Axis Factoring (PAF) followed by Promax rotation was conducted assuming that the factors are

correlated with each other. Based on the criteria of Eigenvalue ≥ 1 , results showed that a four-factor model emerged from the data with items significantly loaded on the respective factors. The four factors in total explained 70.482% of the item total variance. Table 1 shows the related findings.

Regarding the factor structure of the 28-item SOES-SR, EFA was also performed based on Subsample 1. The KMO value of 0.964 and $p < 0.001$ for Bartlett's test of sphericity suggested the suitability for factorial detection of the scale. EFA using PAF and Promax rotation showed a clear three-factor structure which explained 70.775% of the total variance in the scale. The three-factor structure matched with the hypothesized factor structure. All factor loadings ranged between 0.65 and 0.89 (Table 1), indicating good representation of their respective factors.

Based on the results of EFA, CFA was then conducted on the 3-factor model of the scale based on Subsample 2. Results revealed that the model yields a good fit ($\chi^2(339) = 1543.268$, $p < .001$; CFI = .96; TLI = .95; RMSEA = .06; SRMR = .04). The factor loadings were all above 0.77 and statistically significant (Table 2). The satisfactory fit provided basis for performing multi-group CFA to examine measurement invariance across groups.

Subsample 2 was further split into two subgroups based on "odd" and "even" case numbers to perform measurement invariance tests. In the configural invariance model (Model 1), no parameters were constrained to be equal. As revealed in Table 3, the obtained fit indices suggest configural invariance across groups ($\chi^2(678) = 2218.25$, $p < .001$; CFI = .94; TLI = .94; RMSEA = .05; SRMR = .04). In Model 2, equal constraints were applied to factor loadings (i.e., metric invariance). No significant difference was found between Model 1 and Model 2 ($\Delta\chi^2 = 32.48$, $p > .05$; $\Delta\text{CFI} = .000$). In other words, factor loadings were equivalent across groups. In Model 3, the item intercepts were further constrained to be equal based on Model 2. The insignificant $\Delta\chi^2$ and the ΔCFI being 0.001 suggested strong factorial invariance across groups. In Model 4, factorial variance/covariance was set to be equal based on Model 3. The equality of factorial variance/covariance was supported based on insignificant $\Delta\chi^2$ and ΔCFI of 0.000 between Model 3 and Model 4. Based on Model 4, the residual variances were constrained to be equal (i.e., Model 5). Although there was a significant difference in $\Delta\chi^2$ ($p < .001$), the changes in CFI ($\Delta\text{CFI} = .004$) signaled that residual variance held equal in both groups. In summary, the multi-group CFA suggests the factorial invariance across groups with "even" and "odd" case numbers. Overall, the results well supported the three-factor model.

For the inter-relationships among the three facets of SOES-SR, there were significant correlations between perceived program content and implementer quality ($r_s = .82$, $p < .001$), perceived program content and perceived benefits ($r_s = .55$, $p < .001$), and perceived implementer quality and perceived benefits ($r_s = .48$, $p < .001$), thus giving support to the convergent validity of the scale (Hypotheses 1a to 1c). Using the three items as external criteria, the total score of the 28-item SOES-SR was significantly correlated to willingness to join similar program again ($r_s = .49$, $p < .001$), willingness to recommend friends to join ($r_s = .50$, $p < .001$), and overall satisfaction ($r_s = .55$, $p < .001$), which provided support for Hypotheses 2a to 2c. The Cronbach's α values also showed excellent reliability of the total scale and subscales (see Table 4).

Table 1 Standardized factor loadings in EFAs of the 28-item scale and the 28-item scale plus three items on external criteria ($n = 927$)

Scale	Item	Factor loading in EFA of the scale	Factor loading in EFA of the scale and the dimension of three external criteria
Service activity	1	0.78	0.74
	2	0.75	0.71
	3	0.76	0.73
	4	0.66	0.63
	5	0.76	0.78
	6	0.71	0.73
	7	0.79	0.80
	8	0.81	0.83
	9	0.74	0.74
Service implementers	1	0.79	0.79
	2	0.88	0.89
	3	0.73	0.73
	4	0.83	0.84
	5	0.82	0.83
	6	0.81	0.82
	7	0.87	0.89
	8	0.65	0.66
	9	0.75	0.77
Perceived benefits	1	0.75	0.73
	2	0.89	0.89
	3	0.88	0.89
	4	0.79	0.79
	5	0.86	0.88
	6	0.84	0.85
	7	0.82	0.83
	8	0.86	0.85
	9	0.87	0.86
	10	0.86	0.84
Dimension of the three external criteria	1		0.92
	2		0.78
	3		0.59

3.2 Descriptive Profiles and Correlates

The percentage responses of the service recipients are shown in Tables 5. More than 90% of secondary school students rated the service activities positively (e.g. 94.0% of them thought the atmosphere of the activities was pleasant). Similarly, the majority of students perceived the service implementers positively with over

Table 2 Standardized factor loadings for CFA of the 28-item scale (n=927)

Scale	Item	Factor Loading	R ²
Service activity	1	0.88	0.77
	2	0.88	0.77
	3	0.87	0.76
	4	0.79	0.62
	5	0.79	0.62
	6	0.77	0.59
	7	0.84	0.71
	8	0.83	0.69
	9	0.84	0.71
Service implementers	1	0.82	0.67
	2	0.86	0.74
	3	0.82	0.67
	4	0.88	0.77
	5	0.87	0.76
	6	0.88	0.77
	7	0.88	0.77
	8	0.77	0.59
	9	0.81	0.66
Perceived benefits	1	0.82	0.67
	2	0.85	0.72
	3	0.89	0.79
	4	0.84	0.71
	5	0.83	0.69
	6	0.84	0.71
	7	0.80	0.64
	8	0.79	0.62
	9	0.79	0.62
	10	0.79	0.62

90% students giving positive responses for all the 9 items. With regard to the perceived benefits, a high percentage (92.5%) agreed that this experience broadened their horizons and 91.3% reported the usefulness of it. Overwhelmingly positive responses (97.7%) were found for the item of overall satisfaction.

3.3 Predictors of Subjective Outcome Evaluation Measures

Multiple regression analyses showed that perceptions of the program, implementer, and benefits showed statistically significant positive relationships with the overall satisfaction ($F=222.74$, $p<.001$, with the related regression coefficients = 0.18, 0.20 and 0.24, respectively, $p<.001$). The findings supported Hypotheses 3 to 5.

Table 3 Model fit of various measurement invariance tests for groups based on “Odd” ($n = 464$) and “Even” ($n = 463$) case numbers in subsample 2

Model	Description	<i>df</i>	χ^2	$\Delta\chi^2$	CFI	Δ CFI	TLI	RMSEA (90% CI)	SRMR
M1	Configural invariance	678	2218.25***		0.94		0.94	0.050	0.04
M2	Weak factorial invariance (metric invariance)	703	2250.73***	32.48	0.94	0.000	0.94	0.049	0.04
M3	Strong factorial invariance (scalar invariance)	731	2281.60***	30.86	0.94	0.001	0.94	0.048	0.04
M4	Equality of factorial variance/covariance	737	2287.92***	6.32	0.94	0.000	0.94	0.048	0.04
M5	Strict factorial invariance	773	2458.81***	170.89***	0.94	0.004	0.94	0.049	0.05

*** $p < .001$

CFA, Confirmatory factor analysis; RMSEA, Root mean square error of approximation; CI, Confidence interval; CFI, Comparative fit index; TLI, Tucker-Lewis index; SRMR, Standardized root mean square residual. $\Delta\chi^2$, Change in goodness-of-fit; Δ CFI, Change in comparative fit index

3.4 Service Recipients' Evaluation by Activity Type

As shown in Table 6, activity types showed a statistically significant effect on the mean scores of perceived program content ($F = 9.34$, $p < .001$, $\eta^2 = .03$), implementer qualities ($F = 10.38$, $p < .001$, $\eta^2 = .03$), and benefits ($F = 3.40$, $p = .002$, $\eta^2 = .01$). Post-hoc comparisons employing the Least Significant Difference (LSD) test were further performed to examine the related differences across service types.

For the satisfaction of activity, students who attended language workshops ($M = 5.18$, $SD = 0.83$) or two types of activities ($M = 5.15$, $SD = 0.82$) had highest scores, which were significantly higher than those who joined workshops for life planning or personal development ($M = 4.78$, $SD = 0.94$). Similarly, students participating in language workshops ($M = 5.36$, $SD = 0.68$) and two types of activities ($M = 5.34$, $SD = 0.79$) reported the highest scores on the satisfaction of service providers, while students who took part in the workshops for life planning or personal development had the lowest means scores ($M = 4.97$, $SD = 0.92$). For perceived benefits, students who joined two types of activities scored the highest ($M = 5.17$, $SD = 0.79$), followed by those who engaged in a volunteer or

Table 4 Descriptive statistics and internal consistency of SOES-SR

Variables	Total ($n = 1854$)			Sample 1 ($n = 927$)			Sample 2 ($n = 927$)		
	Mean	SD	α	Mean	SD	α	Mean	SD	α
Views towards the service activities (9 items)	4.92	.89	.95	4.90	.88	.95	4.94	.90	.95
Views towards the service providers (9 items)	5.11	.84	.96	5.08	.84	.96	5.15	.83	.96
Perceived benefits (10 items)	4.95	.91	.96	4.94	.93	.96	4.95	.90	.96
Total scale	4.95	.78	.97	4.97	.77	.97	5.01	.73	.96

Table 5 Aggregated results of service learning evaluation

	M	SD	N	%	Strongly Disagree (1)	Disagree (2)	Slightly Disagree (3)	Negative Response (Aggregate 1-3)	Slightly Agree (4)	Agree (5)	Strongly Agree (6)	Positive Response (Aggregate 4-6)						
Part 1: Students' views towards the activity																		
1.1. The content design of the activity is very good (n = 1848).	4.91	1.00	32	1.7	16	0.9	70	3.8	118	6.4	385	20.8	816	44.2	529	28.6	1730	93.6
1.2. The format of the activity is appropriate (n = 1844).	4.95	0.98	24	1.3	20	1.1	72	3.9	116	6.3	356	19.3	808	43.8	564	30.6	1728	93.7
1.3. The atmosphere of the activity was pleasant (n = 1846).	5.03	1.00	27	1.5	16	0.9	68	3.7	111	6.0	330	17.9	721	39.1	684	37.1	1735	94.0
1.4. There was much peer interaction amongst the students (n = 1846).	4.93	1.06	29	1.6	30	1.6	87	4.7	146	7.9	384	20.8	684	37.1	632	34.2	1700	92.1
1.5. I participated in the activity actively (n = 1843).	4.94	1.07	40	2.2	26	1.4	63	3.4	129	7.0	366	19.9	734	39.8	614	33.3	1714	93.0
1.6. I was encouraged to do my best (n = 1808).	4.77	1.09	36	2.0	37	2.0	96	5.3	169	9.3	456	25.2	704	38.9	479	26.5	1639	90.7
1.7. The learning experience enhanced my interests towards the service (n = 1836).	4.86	1.09	37	2.0	34	1.9	90	4.9	161	8.8	398	21.7	709	38.6	568	30.9	1675	91.2
1.8. Overall speaking, I have a very positive evaluation on the activity (n = 1833).	4.97	1.06	34	1.9	25	1.4	66	3.6	125	6.8	371	20.2	674	36.8	663	36.2	1708	93.2
1.9. On the whole, I like this activity very much (n = 1823).	4.94	1.09	37	2.0	23	1.3	97	8.6	157	8.6	342	18.8	684	37.5	640	35.1	1666	91.4
Part 2: Students' views towards PolyU's students:																		
2.1. PolyU student(s) was (were) well prepared for the activity (n = 1833).	5.11	0.95	20	1.1	8	0.4	70	3.8	98	5.3	284	15.5	730	39.8	721	39.3	1735	94.7
2.2. PolyU student(s) showed professional attitudes (n = 1837).	5.17	0.92	17	0.9	9	0.5	48	2.6	74	4.0	275	15.0	704	38.3	784	42.7	1763	96.0
2.3. PolyU student(s) understood my needs and potentials (n = 1836).	4.92	1.03	21	1.1	22	1.2	103	5.6	146	8.0	398	21.7	682	37.1	610	33.2	1690	92.0

Table 5 (continued)

	M	SD	N	%	Very Unhelpful (1)	N	%	Unhelpful (2)	N	%	Unhelpful (2)	N	%	Slightly Unhelpful (3)	N	%	Negative Response (Aggregate 1-3)	N	%	Slightly Helpful (4)	N	%	Helpful (5)	N	%	Strongly Agree (6)	N	%	Positive Response (Aggregate 4-6)	N	%
2.4. PolyU student(s) was (were) actively involved (n = 1830).	5.22	0.91	17	0.9	9	0.5	37	2.0	63	3.4	271	14.8	649	35.5	847	46.3	1767	96.6													
2.5. PolyU student(s) encouraged me to participate in the activity (n = 1830).	5.12	0.97	18	1.0	18	1.0	55	3.0	91	5.0	307	16.8	672	36.7	760	41.5	1739	95.0													
2.6. PolyU student(s) cared about me (n = 1832).	5.02	1.00	21	1.1	16	2.0	80	4.4	117	6.4	354	19.3	672	36.7	689	37.6	1715	93.6													
2.7. PolyU student(s) showed readiness to offer help to me when needed (n = 1838).	5.16	0.96	21	1.1	13	0.7	49	2.7	83	4.5	293	15.9	658	35.8	804	43.7	1755	95.5													
2.8. PolyU student(s) had much interaction with us (n = 1830).	5.09	1.03	22	1.2	24	1.3	73	4.0	119	6.5	317	17.3	609	33.3	785	42.9	1711	93.5													
2.9. Overall speaking, I have a very positive evaluation on the PolyU students (n = 1825).	5.20	0.98	24	1.3	17	0.9	49	2.7	90	4.9	245	13.4	625	34.3	863	47.3	1733	95.1													
Part 3: The extent to which the activity has helped students:																															
3.1. It has strengthened my resilience in adverse conditions (n = 1747).	4.84	1.13	26	1.5	36	2.1	140	8.0	202	11.6	369	21.1	591	33.8	585	33.5	1545	88.4													
3.2. It has helped me to face the future with a positive attitude (n = 1729).	4.92	1.13	31	1.8	28	1.6	127	7.3	186	10.8	325	18.8	570	33.0	648	37.5	1543	89.2													
3.3. It has enhanced my self-confidence (n = 1730).	4.94	1.09	24	1.4	24	1.4	121	7.0	169	9.8	344	19.9	573	33.1	644	37.2	1561	90.2													
3.4. It has broadened my horizon (n = 1690).	5.07	1.05	23	1.4	16	0.9	87	5.1	126	7.5	288	17.0	557	33.0	719	42.5	1564	92.5													
3.5. It has enhanced my interests towards study (n = 1695).	4.98	1.09	24	1.4	22	1.3	116	6.8	162	9.6	309	18.2	557	32.9	667	39.4	1533	90.4													
3.6 It has enhanced my ability of caring. (n = 1694).	4.94	1.11	25	1.5	26	1.5	123	7.3	174	10.3	325	19.2	552	32.6	643	38.0	1520	89.7													

Table 5 (continued)

3.7. It has helped me to develop a good relationship with adults (for example, teachers, parents etc.) (<i>n</i> = 1662).	4.84	1.22	43	2.6	33	2.0	147	8.8	223	13.4	310	18.7	517	31.1	612	36.8	1439	86.6
3.8. It has promoted my aspiration (<i>n</i> = 1672).	4.93	1.15	32	1.9	28	1.7	132	7.9	192	11.5	290	17.3	547	32.7	643	38.5	1480	88.5
3.9. It has promoted my holistic development (<i>n</i> = 1673).	4.99	1.09	24	1.4	18	1.1	109	6.5	151	9.0	320	19.1	537	32.1	665	39.7	1522	91.0
3.10. On the whole, I think the activity is very useful to me (<i>n</i> = 1654).	5.02	1.10	25	1.5	25	1.5	94	5.7	144	8.7	303	18.3	509	30.8	698	42.2	1510	91.3
				Definitely Will Not (1)			Will Not (2)		Negative Response (Aggregate 1-2)		Will (3)		Definitely will (4)		Positive Response (Aggregate 3-4)			
	Mean	SD	N	%	N	%	N	%	N	%	N	%	N	%	N	%		
4.1. Hope to participate again in the future (<i>n</i> = 1830).	3.19	0.59	20	1.1	120	6.6	140	7.7	1175	64.2	515	28.1	1690	92.3				
4.2. Suggest friends to take this course (<i>n</i> = 1830).	3.16	0.61	18	1.0	159	8.7	177	9.7	1165	63.7	488	26.7	1653	90.3				
				Very dissatisfied (1)		Moderately dissatisfied (2)		Negative Response (Aggregate 1-2)		Satisfied (3)		Moderately satisfied (4)		Very satisfied (5)		Positive Response (Aggregate 3-5)		
	Mean	SD	N	%	N	%	N	%	N	%	N	%	N	%	N	%		
4.3. Overall satisfaction (<i>n</i> = 1812)	4.02	0.85	12	0.7	29	1.6	41	2.3	479	26.4	687	37.9	605	33.4	1771	97.7		

educational camp ($M = 5.11$, $SD = 0.88$). Both groups scored significantly higher than did the participants who joined workshops for life planning or personal development ($M = 4.82$, $SD = 1.10$). The present findings support Hypothesis 6.

3.5 Service Recipients' Evaluation by Grade Level

Regarding perceptions towards the service activities, junior grades showed significantly higher scores ($F = 43.53$, $p < .001$, $\eta^2 = .03$) than did senior counterparts (Table 7). Similar findings were found for perceptions of service providers ($F = 68.02$, $p < .001$, $\eta^2 = .04$). However, students' grade level did not show a significant effect on the perceived benefits ($F = .92$, $p = .34$). The present findings provide support for Hypothesis 7 with reference to perceived program content and implementer qualities.

4 Discussion

The current study represents a pioneer attempt to validate a subjective outcome evaluation scale and examined the perceived effectiveness of a SL program by focusing on program recipients. Several features of the study should be highlighted. First, this is the first known study that validated a subjective outcome evaluation measure based on SL service recipients in Chinese contexts. It provides a validated measure that can be used by practitioners and researchers in the future. Second, this study reinforces the findings that the corporate-university-community partnership model may have a positive influence on both the service providers and service recipients (Shek et al. 2019b). Thirdly, this study improves our understanding of the correlates of satisfaction with SL and how perceptions of SL interplay with a variety of activity types and students from different grade levels. Fourth, as the study was conducted in the context of Hong Kong, it enriches the Chinese database on SL.

In terms of scale validation, the findings strongly support the conceptual model with the three dimensions of the subjective outcome. Besides, analyses of the scale

Table 6 Students' perceptions by activity types

	Views towards Activity Mean (SD)	View towards Service Providers Mean (SD)	Perceived Benefits Mean (SD)
Workshop on life planning or personal development ($n = 430$)	4.78 (.94)	4.97 (.92)	4.82 (1.10)
Language workshop ($n = 175$)	5.18 (.83)	5.36 (.68)	4.90 (.93)
PolyU visit ($n = 384$)	4.92 (.82)	5.06 (.85)	4.85 (.91)
Volunteer or educational camp ($n = 221$)	5.03 (.82)	5.31 (.73)	5.11 (.88)
Workshop on STEM, art or craft ($n = 208$)	5.02 (.85)	5.19 (.83)	4.98 (.91)
Two activities ($n = 133$)	5.15 (.82)	5.34 (.79)	5.17 (.79)
Others ($n = 225$)	4.70 (.95)	4.97 (.81)	4.89 (1.03)
F	9.34***	10.38***	3.40**

*** $p < .001$, ** $p < .01$

and three subscales suggest good convergent validity and high reliability. This finding is important because there is a lack of validated subjective outcome evaluation tools in the area of SL in different Chinese contexts. The present study also suggests that the conceptual model with tripartite attributes is applicable in the field of SL.

The effectiveness of the SL program was evidenced by the overwhelmingly positive responses from the service recipients. Positive feedback was predominant in the evaluation of service content, service providers, and the perceived benefits of service participation. Two implications can be drawn. First, the findings add to the literature that service-learning benefits service recipients in different aspects. Consistent with the previous studies on service providers (e.g. Chung and McBride 2015; Kiely 2004), the present study contributes to the understanding of the theoretical underpinning of service-learning that reciprocity and mutual benefits can be achieved during the process. It confirms the notion that both service implementers and recipients can be “engaged as co-learners and co-creators of knowledge” (Vogel and Seifer 2011, p. 186). Second, this study suggests that the corporate-university-community partnership model may be a viable vehicle for teachers to implement SL projects.

The present study also enhances our understanding of the predictors of overall satisfaction in the SL context. As hypothesized, perceived program content, program implementer quality and benefits significantly contributed to the overall satisfaction of the program. This finding echoes several previous studies (e.g., Shek and Sun 2014) in similar contexts where program quality, instructor quality, and perceived benefits had a significantly predictive effect on the perceived overall effectiveness.

For the correlates of activity types, students who joined language workshops and two service activities reported the highest mean scores in their satisfaction level. The results strike a similar chord with students’ views towards service providers, where students attended language workshops and two activities provided significantly better feedback. The results also showed that those who participated in two activities and volunteer/educational camp scored higher in terms of the perceived benefits from the service participation. In general, this study provides empirical evidence that language workshops and volunteer/educational camp might be the most well-received service activities. The popularity of language workshops might be a result of the limited access to foreign/second languages for

Table 7 Students’ perceptions by grade Level

	Views towards Activity	View towards Service Providers	Perceived Benefits
	Mean (SD)	Mean (SD)	Mean (SD)
Junior secondary students ($n = 669$)	5.07 (.91)	5.29 (.82)	4.95 (1.05)
Senior secondary students ($n = 743$)	4.79 (.84)	4.97 (.83)	4.90 (.91)
F	43.53***	68.02***	.92

*** $p < .001$

those underprivileged students, and language workshops therefore addressed their long-term needs. There are also possible explanations for students' favorable attitude towards volunteer/educational camps, as research has documented its capability in promoting socialization and improving emotional climate (Silliman and Schumm 2013).

In addition, the fact that the highest scores were consistently reported by students who attended two activities may be related to the concept of dosage in prevention programs. This result supports the connection between the dosage and positive perception of SL participation. It is consistent with many previous studies (e.g. Ferrer-Wreder et al. 2010; Reyes et al. 2012) which suggested that higher dosage (i.e. more program instruction, training or participation) can lead to more optimal outcomes. An important practical implication drawn from this type of dosage effect studies is that program partners and implementers would be motivated to step up effort to implement SL programs for the benefit of service recipients (Zhai et al. 2010).

Significant associations were also found between students' perceptions of SL and their grade levels. Despite the overall positive feedback, junior grades scored significantly higher than did senior grades in terms of their feedback on the SL content and implementers. These findings give support to the conjecture that SL may have greater influences on the younger cohorts, which echoes the earlier studies (e.g., Shek and Law 2014) of the positive youth development program in Hong Kong, where they found that lower-grade students perceived the program in a more favorable light. In general, the interrelationships between the views of SL and the correlates contribute to our understanding of SL in a Chinese context, providing a basis for future studies.

Despite the above-mentioned contributions of the present study, several limitations should be noted. First, although the sample was large in this study, it only represents secondary school students joining the Project *WeCan*. It would be illuminating to replicate the findings in different contexts and with different cohorts. Second, this study yields findings on subjective outcomes based on the client satisfaction approach only. It would be more insightful to include objective outcome evaluation as a method for triangulation, as it would be able to demonstrate findings such as students' transformation before and after SL participation (Strage 2000). The collection of pretest and posttest data as a routine practice would be helpful (Li and Shek 2019). Additionally, qualitative evaluation could have shed more light on the subjective experiences underlying the statistics, such as why some activities were better received by service recipients. Finally, although the present findings give support to the concurrent validity of the measure using measures on whether the participants will join similar programs, whether the participants will recommend others to join, and overall satisfaction, it would be helpful to use other external criteria not related to overall satisfaction of the programs. In particular, researchers can use measures of positive youth development (PYD) such as psychosocial competence (Zhou et al. 2020) and academic performance as external criteria for the concurrent validity of the measure. Despite these limitations, the present study contributes to the limited SL literature on the value of SL learning activities from the perspective of the service recipients.

Funding This paper and the evaluation project are financially supported by The Wharf (Holdings) Limited.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Adler, R. P., & Goggin, J. (2005). What do we mean by “civic engagement”? *Journal of Transformative Education*, 3(3), 236–253.
- Anorico, H. C. (2019). Service-learning in the Philippines: The University of Santo Tomas’ National Service Training Program. *International Journal of Community Research and Engagement*, 12(1), 1–8. <https://doi.org/10.5130/ijcre.v12i1.5944>.
- Ash, S. L., & Clayton, P. H. (2004). The articulated learning: An approach to guided reflection and assessment. *Innovative Higher Education*, 29(2), 137–154.
- Bahia, K., Paulin, M., & Perrien, J. (2000). Reconciling literature about client satisfaction and perceived services quality. *Journal of Professional Services Marketing*, 21(2), 27–43.
- Becker, B., & Luthat, S. (2002). Social-emotional factors affecting achievement outcomes among disadvantaged students: Closing the achievement gap. *Educational Psychologist*, 37(4), 197–214. https://doi.org/10.1207/s15326985ep3704_1.
- Benson, P. L., Scales, P. C., & Syvertsen, A. K. (2011). The contribution of the developmental assets framework to positive youth development theory and practice. In R. M. Lerner, J. V. Lerner, & J. B. Benson (Eds.), *Advances in child development and behavior* (pp. 197–230). New York: Elsevier.
- Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin*, 88, 588–606.
- Billig, S. H. (2000). The effects of service learning. Resource document. <https://digitalcommons.unomaha.edu/cgi/viewcontent.cgi?article=1038&context=slceslgen>. Accessed 8 June 2020.
- Billig, S., & Furco, A. (2002). *Service-learning through a multidisciplinary lens*. Greenwich: Information Age.
- Birbeck, D. (2012). Differentiating service learning in an Australian higher education context. *The Journal of the Education Research Group of Adelaide*, 2(3), 27–32.
- Boyle-Baise, M., & Langford, J. (2004). There are children here: Service-learning for social justice. *Equity and Excellence in Education*, 37(1), 55–66.
- Bingle, R. G., & Hatcher, J. A. (1996). Implementing service learning in higher education. *The Journal of Higher Education*, 67(2), 221. <https://doi.org/10.2307/2943981>.
- Bingle, R., Hatcher, J., & McIntosh, R. (2006). Analyzing Morton’s typology of service paradigms and integrity. *Michigan Journal of Community Service Learning*, 13(1), 5–15.
- Burke, A. S., & Bush, M. D. (2013). Service learning and criminal justice: An exploratory study of student perceptions. *Educational Review*, 65(1), 56–69. <https://doi.org/10.1080/00131911.2011.638138>.
- Campbell, D. E. (2000). Social capital and service learning. *Political Science and Politics*, 33(3), 641–645.
- Celio, C. I., Durlak, J., & Dymnicki, A. (2011). A meta-analysis of the impact of service-learning on students. *The Journal of Experimental Education*, 34(2), 164–181.
- Chen, S. (2015). Faculty members’ perception of service-learning courses in curriculum. *International Journal of Education and Human Developments*, 1(3), 9–19.
- Chen, Y., & Hoshower, L. (2003). Student evaluation of teaching effectiveness: An assessment of student perception and motivation. *Assessment & Evaluation in Higher Education*, 28(1), 71–88. <https://doi.org/10.1080/02602930301683>.
- Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, 9(2), 233–255. <https://doi.org/10.1097/NNR.0b013e3182544750>.

- Chung, S., & McBride, A. M. (2015). Social and emotional learning in middle school curricula: A service learning model based on positive youth development. *Children and Youth Services Review*, *53*, 192–200. <https://doi.org/10.1016/j.childyouth.2015.04.008>.
- Cone, D., & Harris, S. (1996). Service-learning practice: Developing a theoretical framework. *Michigan Journal of Community Service Learning*, *3*, 31–43.
- Conway, J. M., Amel, E. L., & Gerwien, D. P. (2009). Teaching and learning in the social context: A meta-analysis of service learning's effects on academic, personal, social, and citizenship outcomes. *Teaching of Psychology*, *36*, 233–245.
- Considine, G., & Zappalà, G. (2002). The influence of social and economic disadvantage in the academic performance of school students in Australia. *Journal of Sociology*, *38*(2), 129–148. <https://doi.org/10.1177/144078302128756543>.
- d'Arlach, L., Sánchez, B., & Feuer, R. (2009). Voices from the community: A case for reciprocity in service-learning. *Michigan Journal of Community Service Learning*, *16*(1), 5–16.
- Deeley, S. J. (2010). Service-learning: Thinking outside the box. *Active Learning in Higher Education*, *11*(1), 43–53. <https://doi.org/10.1177/1469787409355870>.
- Eby, J. (1998). Why service-learning is bad. Resource document. <https://digitalcommons.unomaha.edu/slceslgen/27>. Accessed 8 June 2020.
- Ehrlich, T. (1996). Foreword. In B. Jacoby & Associates (Eds.), *Service-learning in higher education: Concepts and practices* (pp. xi–xvi). San Francisco: Jossey-Bass.
- Einfeld, A., & Collins, D. (2008). The relationships between service-learning, social justice, multicultural competence, and civic engagement. *Journal of College Student Development*, *49*(2), 95–109.
- European Observatory of Service-Learning in Higher Education (2020). Mapping of the service-learning experiences in higher education across Europe. Resource document. <https://www.eoslhe.eu/>. Accessed 7 June 2020.
- Faber, K. A. (2017). Learning by doing: Service learning as a means of personal growth in the middle grades. *Current Issues in Middle Level Education*, *22*(1), 1–9.
- Felten, P., & Clayton, P. H. (2011). Service-learning. *New Directions for Teaching and Learning*, *128*, 75–84.
- Ferrer-Wreder, L., Cadely, H. S., Domitrovich, C. E., Small, M. L., Caldwell, L. L., & Cleveland, M. J. (2010). Is more better? Outcome and dose of a universal drug prevention effectiveness trial. *The Journal of Primary Prevention*, *31*(5–6), 349–363. <https://doi.org/10.1007/s10935-010-0226-4>.
- Fraser, M. W., & Wu, S. Y. (2016). Measures of consumer satisfaction in social welfare and behavioral health: A systematic review. *Research on Social Work Practice*, *26*(7), 762–776.
- Furco, A. (1996). Service-learning: A balanced approach to experiential education. In B. Taylor and Corporation for National Service (Eds.), *Expanding boundaries: Serving and learning* (pp. 2–6). Washington, DC: Corporation for National Service.
- Giles, D. E. Jr., & Eyler, J. (1994). The theoretical roots of service-learning in John Dewey: Toward a theory of service learning. Resource document. <http://digitalcommons.unomaha.edu/slceslgen/150>. Accessed 8 June 2020.
- Gregorich, S. E. (2006). Do self-report instruments allow meaningful comparisons across diverse population groups? Testing measurement invariance using the confirmatory factor analysis framework. *Medical Care*, *44*(11), S78–S94.
- Guo, F., Yao, M., Wang, C., Yan, W., & Zong, X. (2016). The effects of service learning on student problem solving: The mediating role of classroom engagement. *Teaching of Psychology*, *43*(1), 16–21. <https://doi.org/10.1177/0098628315620064>.
- Hsieh, C. M. (2006). Incorporating perceived importance of service elements into client satisfaction measures. *Research on Social Work Practice*, *22*(1), 93–99.
- Hirsh, J. B. (2010). Personality and environmental concern. *Journal of Environmental Psychology*, *30*, 245–248. <https://doi.org/10.1016/j.jenvp.2010.01.004>.
- Kahne, J., & Westheimer, J. (1996). In the service of what? The politics of service learning. *Phi Delta Kappan*, *77*(9), 593–599.
- Kearney, K. R. (2004). Students' self-assessment of learning through service-learning. *American Journal of Pharmaceutical Education*, *68*(1), 29. <https://doi.org/10.5688/aj680129>.
- Kember, D., & Leung, D. Y. P. (2008). Establishing the validity and reliability of course evaluation questionnaires. *Assessment and Evaluation in Higher Education*, *33*(4), 341–353.
- Kiely, R. (2004). A chameleon with a complex: Searching for transformation in international service learning. *Michigan Journal of Community Service Learning*, *10*(2), 5–20.
- Kricsfalussy, V., Zecevic, A., Assanand, S., Bigelow, A., & Gaudet, M. (2016). The frontiers of service-learning at Canadian universities. *Community-Engaged Research, Teaching, and Learning*, *2*(2), 87–102.

- Larsen, D. L., Attkisson, C. C., Hargreaves, W. A., & Nguyen, T. D. (1979). Assessment of client/patient satisfaction: Development of a general scale. *Evaluation and Program Planning*, 2(3), 197–207.
- Lee, S. J., Wilder, C., & Yu, C. (2018). Exploring students' perceptions of service-learning experiences in an undergraduate web design course. *Teaching in Higher Education*, 23(2), 212–226. <https://doi.org/10.1080/13562517.2017.1379486>.
- Lerner, R. M., Lerner, J. V., Lewin-Bizan, S., Bowers, E. P., Boyd, M. J., Mueller, M. K., Schmid, K. L., & Napolitano, C. M. (2011). Positive youth development: Processes, programs, and problematics. *Journal of Youth Development*, 6, 38–62. <https://doi.org/10.5195/JYD.2011.174>.
- Li, X., & Shek, D. T. L. (2019). Objective outcome evaluation of a leadership course utilising the positive youth development approach in Hong Kong. *Assessment & Evaluation in Higher Education*, 45, 741–757. <https://doi.org/10.1080/02602938.2019.1696944>.
- Lin, C. Y., Wu, S. W., Wu, S. F., Pan, B. S., & Liao, H. C. (2014). Effect of service learning in a university in Taiwan. *Procedia. Social and Behavioral Sciences*, 116, 902–906.
- Ma, C. M. S., Shek, D. T. L., Li, P., & Shek, V. (2018). Promotion of service leadership: An evaluation of a service-learning subject in Hong Kong. *Journal of Service-Learning in Higher Education*, 8, 61–72.
- Maccio, E., & Voorhies, R. A. (2012). Social work students' perceptions of service-learning. *Journal of Service-Learning in Higher Education*, 1, 50–69.
- Mason, M. R., & Dunens, E. (2019). Service-learning as a practical introduction to undergraduate public health: Benefits for student outcomes and accreditation. *Frontiers in Public Health*, 7. <https://doi.org/10.3389/fpubh.2019.00063>.
- McBride, A. M., Robertson, A., & Chung, S. (2014). *Assessing the impacts of service learning on middle school students: Wyman's Teen Outreach Program*. St. Louis: Center for Social Development.
- Meili, R., Fuller, D., & Lydiate, J. (2011). Teaching social accountability by making the links: Qualitative evaluation of student experiences in a service-learning project. *Medical Teacher*, 33(8), 659–666.
- Morin, E. (2009). Service learning pitfalls: Problems you didn't see coming. *College Teaching Methods & Styles Journal*, 5(1), 43–51.
- Morton, K. (1995). The irony of service: Charity, project and social change in service-learning. *Michigan Journal of Community Service Learning*, 2, 19–32.
- O'Neal, P. D. (1999). Methodological problems associated with measuring consumer satisfaction in the mental health field. *Australian Social Work*, 52(3), 9–15. <https://doi.org/10.1080/03124079908414130>.
- Pacho, T. O. (2019). *Service-learning in higher education in Africa*. UK: Cambridge Scholars Publishing.
- Perneger, T. V., Peytremann-Bridevaux, I., & Combescur, C. (2020). Patient satisfaction and survey response in 717 hospital surveys in Switzerland: A cross-sectional study. *BMC Health Service Research*, 20, 1–8. <https://doi.org/10.1186/s12913-020-5012-2>.
- Reyes, M. R., Brackett, M. A., Rivers, S. E., Elbertson, N. A., & Salovey, P. (2012). The interaction effects of program training, dosage, and implementation quality on targeted student outcomes for the RULER approach to social and emotional learning. *School Psychology Review*, 41(1), 82–99.
- Rubio, G., Chung, H. L., & Reyes, C. A. (2018). Qualitative approach to exploring short-term service learning and civic engagement. *Social Behavior Research and Practice*, 3(1), 10–19. <https://doi.org/10.17140/SBRPOJ-3-111>.
- Shek, D. T. L. (2010). Subjective outcome and objective outcome evaluation findings: Insights from a Chinese context. *Research on Social Work Practice*, 20(3), 293–301.
- Shek, D. T. L. (2014). Is subjective outcome evaluation related to objective outcome evaluation? Insights from a longitudinal study in Hong Kong. *Journal of Pediatric and Adolescent Gynecology*, 27, 50–56. <https://doi.org/10.1016/j.jpjag.2014.02.012>.
- Shek, D. T. L., Dou, D., Zhu, X., & Chai, W. (2019a). Positive youth development: Current perspectives. *Adolescent Health Medicine Therapeutics*, 10, 131–141. <https://doi.org/10.2147/AHMT.S179946>.
- Shek, D. T. L., & Law, M. (2014). Evaluation of the project P.a.T.H.S. based on the client satisfaction approach: View of the students. *Journal of Pediatric and Adolescent Gynecology*, 27, 2–9. <https://doi.org/10.1016/j.jpjag.2014.02.006>.
- Shek, D. T. L., & Liang, J. (2015). Subjective outcome evaluation of a university subject on service leadership. *International Journal on Disability and Human Development*, 14(4), 385–392. <https://doi.org/10.1515/ijdh-2015-0460>.
- Shek, D. T. L., & Ma, C. M. S. (2014). Validation of a subjective outcome evaluation tool for participants in a positive youth development program in Hong Kong. *Journal of Pediatric and Adolescent Gynecology*, 27(supplement), S43–S49.
- Shek, D. T. L., Ma, C. M. S., & Siu, A. M. H. (2014). Validation of a subjective outcome evaluation tool for participants in a positive youth development program in Hong Kong. *Journal of Pediatrics and Adolescent Gynecology*, 27(supplement), S26–S31.

- Shek, D. T. L., Ma, C. M. S., & Yang, Z. (2019b). Transformation and development of university students through service-learning: A corporate-community-university partnership initiative in Hong Kong (project WeCan). *Applied Research in Quality of Life*. <https://doi.org/10.1007/s11482-019-09738-9>.
- Shek, D. T. L., & Sun, R. C. F. (2014). Positive youth development programs for adolescents with greater psychosocial needs: Subjective outcome evaluation over 3 years. *Journal of Pediatric and Adolescent Gynecology*, 27(supplement), S17–S25.
- Silliman, B., & Schumm, W. (2013). Youth program quality survey: Youth assessment of program quality. *Marriage & Family Review*, 49(7), 647–670.
- Spiro, S. E., Dekel, R., & Peled, E. (2009). Dimensions and correlates of client satisfaction: An evaluation of a shelter for runaway and homeless youth. *Research on Social Work Practice*, 19(2), 261–270.
- Spooren, P., Mortelmans, D., & Denekens, J. (2007). Student evaluation of teaching quality in higher education: Development of an instrument based on 10 Likert-scales. *Assessment and Evaluation in Higher Education*, 32(6), 667–679.
- Strage, A. A. (2000). Service-learning: Enhancing student learning outcomes in a college-level lecture course. *Michigan Journal of Community Service Learning*, 7(1), 5–13.
- Sun, H., & Richardson, J. (2016). Students' perceptions of the academic environment and approaches to studying in British postgraduate business education. *Assessment & Evaluation in Higher Education*, 41(3), 384–399. <https://doi.org/10.1080/02602938.2015.1017755>.
- Trigos-Carrillo, L., Fonseca, L., & Reinoso, N. (2020). Social impact of a transformative service-learning experience in a post-conflict setting. *Frontiers in Psychology*, 11. <https://doi.org/10.3389/fpsyg.2020.00047>.
- Vogel, A., & Seifer, S. (2011). Impacts of sustained institutional participation in service-learning: Perspectives from faculty, staff and administrators. *International Journal of Community Research and Engagement*, 4, 186–202. <https://doi.org/10.5130/ijcre.v4i0.1789>.
- Weah, W., Simmons, V., & Hall, M. (2000). Service-learning and multicultural/multiethnic perspectives: From diversity to equity. Resource document. <https://digitalcommons.unomaha.edu/slcediversity/3>. Accessed 8 June 2020.
- Xiang, R., & Luk, T. (2012). Service learning as a social work pedagogy of China. *China Journal of Social Work*, 5(3), 223–236. <https://doi.org/10.1080/17525098.2012.721179>.
- Yorio, P. L., & Ye, F. (2012). A meta-analysis on the effects of service-learning on the social, personal, and cognitive outcomes of learning. *Academy of Management Learning & Education*, 11(1), 9–27. <https://doi.org/10.5465/amle.2010.0072>.
- Zhai, F., Raver, C., Jones, S., Li-Grining, C., Pressler, E., & Gao, Q. (2010). Dosage effects on school readiness: Evidence from a randomized classroom-based intervention. *Social Service Review*, 84(4), 615–655. <https://doi.org/10.1086/657988>.
- Zhou, Z., Shek, D. T. L., Zhu, X., & Dou, D. (2020). Positive youth development and adolescent depression: A longitudinal study based on mainland Chinese high school students. *International Journal of Environmental Research and Public Health*, 17, 4457. <https://doi.org/10.3390/ijerph17124457>.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.