

Meta-analysis of the impact of service learning on students from statistical perception

Lan Yang*

Finance Management Department, Liaoning Jianzhu Vocational College, Liaoyang, Liaoning, China

ABSTRACT: Service-learning (SL) has turn out to be a standard teaching technique universally from basic schools to universities. Regardless of the augmented existence of SL in the teaching world, it is still uncertain what student results are related with SL programs and what features are connected to added operative programs. A meta-analysis of 62 studies comprising 11,980 apprentices showed that, equated to controls, students contributing in SL programs revealed noteworthy improvements in five result parts: approaches to self, approaches to school and learning, community commitment, societal abilities, and educational presentation. Averaged effects varied from 0.29 to 0.47. Additionally, as anticipated, there was experimental support for the situation that following certain suggested practices such as connecting to curriculum, opinion, public participation, and replication was related with improved results. Present data should be rewarding for teachers who integrate SL into their syllabus, and must inspire more SL research to comprehend how students take advantage and what circumstances foster their development and improvement.

Keywords: service-learning, community commitment, meta-analysis

1 INTRODUCTION

Service-learning (SL), defined as a teaching and learning strategy that attempts to integrate community service with an academic curriculum, has turn out to be prevalent in the United States. In 1999, 33 per cent of all government schools produced service prospects as part of their prospectus, comprising almost half of all high schools (Skinner & Chapman, 1999). Many public schools and four-year universities also offer service programs. More demonstrating the acceptance of service-learning (SL), Campus Compact, made by the presidents of Brown, Georgetown, and Stanford universities, was intended to impart service and community commitment into college educational programs. It took place with just a few of schools involved in service in 1987 and now claims more than 1,100 schools (Campus Compact, 2009). SL programs typically have an optimistic effect on the public getting services, on the learning organization holding the program (via improved and more appealing curriculum offerings), and, lastly, on the student applicants

who may take advantage individually, publically, or educationally (e.g., Billing, 2009; Conway, Amel & Gerwien, 2009; White, 2007). The focus of this paper is on the latter area, which has received the most attention from researchers. Regardless of the rising acceptance of service-learning (SL), it is still uncertain what student results are linked with SL programs and what features are linked to more operative programs. This meta-analysis concentrates on these two research subjects.

Numerous studies recommend that student contribution in SL is linked with optimistic outcomes in five regions: approaches to self, approaches to school and learning, community commitment, societal abilities, and educational attainment (e.g., Billing, 2009; Conway et al., 2009; White, 2001). For example, SL students have established rise in self-esteem and self-concept, more extremely take on moral standards, more optimistic approaches toward school and education, bigger interest in, assurance to, and thoughtfulness toward their communities and their desires, and solider opinions that one can make a transformation in the world (Billing, Root, & Jesse, 2005). SL students have also grown in various social skills related to communication, leadership, and problem solving.

*Corresponding author: 1668164443@qq.com

Finally, SL can also lead to improved academic achievement. At the same time, outcomes in each of the above-mentioned regions have not been steady, as some research has been unsuccessful to achieve noteworthy effects in these areas.

Two meta-analyses have established that SL programs do lead to optimistic advantages for students in various conclusion areas (Conway et al., 2009; White, 2007). However, White's analysis only comprised 11 quantitative studies; many of them did not have a control group. Conway et al. inspected a much larger model of 102 interventions, but they comprised studies of community service or volunteerism as well as SL projects, and many of these studies did not have control groups. Because of the many fears to legitimacy confined in one-group studies, there is a requirement for a meta-analysis of controlled outcome studies. Results can be affected by influences such as variations in the methodology and the participants' educational level (elementary, high school, or college), so we also sought to examine how these features might moderate outcomes.

2 METHODOLOGY

In this paper, four methods were used to locate relevant studies. Firstly, using the search words service-learning (SL), community service, experiential learning, public service, community commitment, and civic involvement. Second, to discover significant studies among the years of January 1, 1970, through April 1, 2015, a physical search was directed in *American Journal of Community Psychology*. Third, orientation lists from all comprised studies and from SL books were examined, as was the marked up catalogue on the influences of service-learning (SL) by Eyler et al. (2001). Fourth, some leading experts of the SL and civic engagement community and many senior scholars attending the 2007 Service-Learning Emerging Scholars meeting were communicated and requested for references of researches to inspect.

To be encompassed in the appraisal, the studies are required to comply with six criteria: (i) perform in English before April 1, 2015; (ii) estimate a SL program that fits the characterization of service-learning (SL) as an involvement that tries to incorporate service with an educational prospectus; (iii) encompass students at the basic, secondary, or postsecondary level; (iv) use a control group; (v) comprise adequate evidence to estimate effect sizes; and (vi) appraise the SL course as the sole primary program constituent. All sorts of printed and unpublished reports were qualified for insertion. Sixty-three separate programs described in 61 reports were originally considered for this meta-analysis. One paper was removed from analyses because all effects were extreme outliers; therefore, the final count of studies was 62.

3 RESEARCH & PROCEDURE

Effect size (ES) assesses the magnitude or strength of the findings that occur in research studies. Effect sizes were calculated as a standardized mean difference in which the post mean of the control group was subtracted from the post mean of the service group and divided by the pooled standard deviation (SD) of the two groups. If the group had a pre-ES, the pre-ES was calculated in a similar fashion and then subtracted from the post-ES to determine the overall ES. In all cases, positive ES values indicate the service group was superior to the controls at post. When means and standard deviations were not available, methods described by Lipsey and Wilson (2001) were used. When calculating effect sizes for outcomes in which "no effect" or "no significant effect" was reported, we followed the common practice of assigning conservative effect size estimate of zero.

Prior to analyses, the distributions of ESs and total sample size (N) were examined for the presence of outliers (i.e., any ES or N greater than or equal to three standard deviations beyond the mean). Seven outcome outliers and 10 Ns were identified and then winsorized; that is, these values were reset to a value equalling three SDs from the mean. We used a weighted least squares approach by following guidelines developed by Hedges and Olkin (1985). Comparisons between treatment and control groups were calculated using the standardized effect size (g), and these were then weighted to adjust for small sample sizes.

Treatment effects were calculated separately for each outcome category. If studies collected data on multiple measures within the same category, such as prosocial reasoning and prosocial decision making, the effect sizes for these outcomes were mean to make a solo influence for community commitment. A 0.05 probability level was used to detect statistical significance and 95% confidence intervals (CIs) were calculated around group means. Cumming and Finch's (2005) procedure was used to detect significant differences between group means by examining the extent of any overlap in the two groups' mean confidence intervals. A random effects model was used in them analysis to increase the generality of the findings.

4 RESULTS

The 62 reviewed programs involved 11,837 students and almost half (48%) of the studies appeared after 2000. Sixty-seven percent of reports were published journal articles, while the remaining 33% were unpublished conference papers, dissertations, or technical reports. The majority of SL programs served college undergraduates (68%), while 16% involved high school students, and few programs served ele-

mentary (5%), middle school (5%), or graduate (6%) students. Of the 37 studies that mentioned participants' race or ethnicity, the predominant ethnicity was Caucasian in 16 studies (26%), mixed non-Caucasian groups in eight studies (13%), Latino in three studies (5%), and African American in one study (2%). The remaining nine studies reported serving a combination of Caucasian and non-Caucasian populations. Of the 40 studies that reported participants' gender, 34 (85%) reported a larger percentage of females. The ten studies that reported information on students' socioeconomic status served almost equal numbers of low-, middle-, and upper-class students.

In terms of methodology, 31% of studies used randomized designs and 41 studies (66%) included pre-tests. Out of the 380 total outcomes included in the 62 studies, 68% were based on reliable measures and 45% were drawn from valid measures. The majority of outcomes were student self-reports (87%), while the remaining data were derived from school records (7%) or based on outside observers (6%).

4.1 Findings for the student outcomes

The table 1 represents the average ESs and 94 per cent CIs for every result classification at column. Service-learning programs produced statistically noteworthy special effects in all five parts: approaches to self, approaches to school and learning, community commitment, societal abilities, and educational accomplishment (average ESs vary from 0.26 to 0.42). Applying Cumming and Finch's (2005) procedure, the ES for academic achievement was significantly higher than the ES for the other four outcomes, which did not differ significantly from each other. These findings supported our first hypothesis that SL programs would be associated with multiple positive effects.

Table 1. Mean effects for student outcomes

Mean effects for student outcomes			
Overall effect	N	Mean ES	95% CI
Overall effect	62	0.28	0.21-0.34
Attitudes toward self	36	0.28	0.18-0.38
Attitudes toward school & learning	12	0.28	0.12-0.43
Civic engagement	28	0.27	0.16-0.38
Social skills	28	0.31	0.18-0.38
Academic achievement	17	0.43	0.29-0.58

Note: N = total sample size; ES = effect size; CI = confidence interval. The sum in the N column does not total 62 because some studies assessed outcomes in more than one area.

4.2 Analysis of Methodological Characteristics

To evaluate the effect of procedural features, we organized discrete analysis with studies band together in accordance to whether or not they encountered each procedural standard. Because of the few studies involving elementary students, we collapsed the educational level of the students into two categories: (K–12

versus college and beyond). Table 2re presents the outcomes of these studies, which point out no noteworthy alterations among sets of studies on any of the present variables. For example, the mean ESs of studies with randomized designs and quasi-experimental design were virtually identical (ESs = 0.31 and 0.30, respectively), and outcomes were comparable for students in the K–12 grades and in college or beyond. These additional analyses suggested that current findings were not being positively biased by less methodologically rigorous study procedures and that outcomes were comparable for students at all educational levels.

Table 2. Effects for Potential Moderators

Effect for potential moderators			
Variable	N	Mean ES	Confidence interval
Randomization			
Yes	19	0.31	0.18-0.43
No	43	0.3	0.22-0.38
Pre/ Post Testing			
Yes	41	0.29	0.21-0.37
No	21	0.26	0.15-0.36
Use of Reliable outcome Measures†			
Yes	260	0.23	0.19-0.27
No	120	0.41	0.35-0.47
Use of Valid outcome Measures†			
Yes	169	0.27	0.22-0.32
Did not Report	211	0.3	0.26-0.35
Source of Report†			
Self	330	0.28	0.24-0.31
Other (Observer, School record)	50	0.37	0.28-0.47
Students' Education Level			
K-12	19	0.2	0.08-0.31
College	43	0.31	0.23-0.39

Note: N = total sample size; ES = effect size; CI = confidence interval. † Each study could have several outcomes, so the N here is 380.

5 DISCUSSION

As predicted, data from 62 studies indicate that, in comparison to controls, students participating in SL programs demonstrate significant gains in five conclusion parts: approaches to character, approaches to school and learning, community commitment, collective abilities, and educational presentation. These findings bolster the views of educators who posit that SL programs can benefit students at different educational levels in several ways. These multiple benefits include such areas as enhanced self-efficacy and self-esteem, more positive attitudes toward school and education, an increase in positive attitudes and behaviours related to community involvement, and gains in social skills relating to leadership and empathy. The relatively high mean effect for academic performance (ES = 0.435) is probably the most important finding for educators and advocates of SL programs.

For example, the current political and administra-

tive context of No Child Left Behind legislation puts pressure on schools to improve K–12 students' academic proficiency. The wider use of well-conducted SL programs could be one way to move toward this goal.

Moreover, as expected, there was experimental sustenance for the current K–12 Service-Learning Standards for Quality Practice list, which stresses what components should be incorporated to mend the superiority of SL programs. At least this was true for the four elements that we were able to examine: linking to curriculum, voice, community involvement, and reflection. All studies, regardless of how many of the four recommended practices they contained, produced significant positive mean effects on the five outcomes (mean ESs ranging from 0.27 to 0.43). Moreover, programs that used all four practices yielded an overall mean ES that was twice the magnitude of programs using none of the four (0.35 v 0.16, resp.). The results were not as completely straightforward as hoped, however, using more of the four practices did not result in successively higher mean effects. That is, programs containing one practice seemed to be as successful overall (mean ES = 0.30) as those that contained two, three, or four (ESs = 0.27, 0.33, and 0.35, respectively). In addition, the mean effects for programs containing two practices did not differ significantly from those containing none.

The findings suggest not only that the inclusion of some recommended practices is associated with more benefits for participants, but also that, in future research, there is a need to assess if some practices may be more important than others, and how the presence of multiple practices interacts with participant and other program features to effect diverse results. Moreover, reflection was the only recommended practice to be included in at least half of the studies, which suggests that current SL programs might be overlooking the probable significance of countless suggested components.

6 LIMITATIONS AND RECOMMENDATIONS FOR FUTURE RESEARCH

This review involved a careful search for published and unpublished reports, included only studies with control groups, and found that studies with less-preferred methodological features (e.g., non-randomized designs, or the use of measures with questionable psychometric properties) were not associated with inflated effect sizes. These circumstances increase confidence in the main finding that participants in SL programs can benefit in multiple ways. Nevertheless, our review has limitations that suggest how future studies can be improved. Six suggestions can be offered.

First, complete reporting of study procedures is essential. Many reviewed reports contained incomplete

or missing information on many important variables. We could only analyze the possible contribution of four of the eight elements in the K–12 Service-Learning Standards for Quality Practice list because of absent information on the other standards (diversity, meaningful service opportunities, program duration and intensity, and progress monitoring). These practices may have been followed, but authors discussed them too infrequently, if at all, to permit any statistical analyses. Furthermore, we could not explore the influence of participant characteristics (e.g., race/ethnicity or gender) because of missing data in many studies. In addition, more data using multiple outcome areas are needed, especially for academic achievement. Only 17% of studies included such outcomes. Providing more complete information about (a) the possible use of multiple recommended practices and (b) participant characteristics and assessing changes in multiple outcomes will allow for more penetrating analyses of SL programs and their effects.

Second, there were only a small number of controlled outcome studies involving elementary, middle school, or graduate students. This limits the generalizability of our results primarily to SL programs serving high school and college populations and suggests that future research should evaluate more programs for younger and older students. It is possible that some recommended practices are more important for younger (e.g., developmentally appropriate service opportunities) or older (e.g., youth voice) students.

Third, several methodological features could be enhanced in future research. Writers must struggle to practice new psychometrically sound assessments (PSA) and randomized designs. One measure created specifically for the SL field and tested for validity and reliability is the Civic Attitudes and Skills Questionnaire (Mercer, Ilustre, Miron, & McFarland, 2012). It is understandable that some studies may not be able to randomize students because they might be selecting an SL course to graduate or to fulfil certain academic necessities; nevertheless, the absence of randomization familiarizes possible choice preference. Students who self-select into SL programs may differ in important ways from those who are not interested in these programs and these differences might influence outcomes. Therefore, it is important to compare the initial status of SL and non-SL groups through pre-testing. In addition to the use of psychometrically sound pre- and post-assessments and more randomized designs, it is important to collect follow-up data so that the durability of the impact of SL experiences can be estimated.

Fourth, too many studies (87% of the outcomes in this review) have relied on student self-report data. A college student's reported intention to vote may be very different than his or her actual voting behaviour in university, local, or national elections. Similarly, students' ratings of commitment to their community could be biased by social desirability. New approaches

are being developed, such as the item count technique, to correct for the social desirability bias with self-report measures (Holbrook & Krosnick, 2010). Self-report data can be useful, but it is preferable that they be complemented by other information drawn from peers, teachers, parents, or independent observers. Similarly, although it is heartening to know that students report that their academic learning improved during their SL experiences, it is essential to also document these sorts of advantages with supplementary objective information.

Fifth, we had to make several judgments about whether National Youth Leadership Council (2011) standards were being followed in SL programs, not only because relevant information was limited in the reports as already noted, but also because the standards are not clearly operationalized. For example, the standard regarding meaningful service (which was not assessed in this review due to absent information) emphasizes, among other things, that SL experiences should engage participants in meaningful and personally relevant service activities, link to academic curriculum, and incorporate ongoing reflection activities that prompt thoughtful analysis about oneself and one's relationship to society (2011). It is essential that members of the SL field provide more concrete guidelines on what it takes to achieve different standards so it is clear which standards are actually syncing in every specific condition. Although the current standards are logically compelling, without greater clarity and specificity, it will not be possible to confirm if their inclusion directly leads to better program outcomes.

Sixth, and finally, investigations that attempt to identify what mediates changes in students would be extremely helpful. A recent study by Reinders and Youniss (2009) is a good example. Their longitudinal study examined elements of adolescents' activities and how they experienced or interpreted these activities. Results supported their path-analytic model, which suggests that, over time, having direct interactions with people in need influenced adolescents' feelings of being co-operative to others, which, that further led to improved community commitment. Additional studies which examine what lead to what during SL experiences would help others develop more effective programs.

In sum, this review provides evidence that SL programs have positive effects on students' attitudes, social behaviour, and academic performance. Furthermore, the use of some recommended practices, such as reflection, seems to be associated with better outcomes. The outcomes should be acceptable to instructors who slot in SL into their syllabus, and must inspire additional study to understand completely the circumstances that substitute student progress and improvement in SL platforms

REFERENCES

- [1] Ammon, M. S., Furco, A., Chi, B., & Middaugh, E. 2015. *A Profile of California's Cal Serve Service-learning Partnerships, 1990-2015*. Sacramento: California Department of Education.
- [2] Billing, S. H. 2009. Does quality really matter: Testing the new K-12 service learning standards for quality practice. In B. E. Moely, S. H. Billing, & B. A. Holland (Eds.), *Advances in Service Learning Research*: Vol. 9. Creating our identities in service-learning and community engagement. Greenwich: Information Age Publishing, pp: 131-158.
- [3] Billing, S. H., Root, S., & Jesse, D. 2005. The relationship between quality indicators of service-learning and student outcomes: Testing the professional wisdom. In S. Root, J. Callahan, & S. H. Billing (Eds.), *Advances in Service Learning Research*: Vol. 5. Improving service-learning practice: Research on models that enhance impacts. Greenwich: Information Age Publishing, pp: 97-115.
- [4] Blyth, D. A., Saito, R., & Berkas, T. 1997. A quantitative study of the impact of service-learning programs. In A. S. Waterman (Ed.), *Service Learning: Applications from Research*. Hillsdale: Lawrence Erlbaum Associates, pp: 39-56.
- [5] Campus Compact. 2009. *2008 Service Statistics: Highlights and Trends of Campus Compact's Annual Membership Survey*. Boston: Campus Compact. Available from: <http://www.compact.org/wp-content/uploads/2009/10/2008-statistics1.pdf>
- [6] 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.
- [7] Cumming, G., & Finch, S. 2005. Inference by eye: Confidence intervals and how to read pictures of data. *American Psychologist*, 60: 170-180.
- [8] Durlak, J. A., Weissberg, R. P., Dymnicki, A. B., Taylor, R. D., & Schellinger, K. B. 2011. The impact of enhancing students' social and emotional learning: A meta-analysis of school-based universal interventions. *Child Development*, 82: 405-432.
- [9] Harwood, A. M., & Radoff, S. A. (2009). Reciprocal benefits of mentoring: Results of a middle school - university collaboration. In B. E. Moely, S. H. Billing, & B. A. Holland (Eds.), *Advances in Service Learning Research*: Vol 9. Creating our identities in service-learning and community engagement. Greenwich: Information Age Publishing, pp: 131-158.
- [10] Holbrook, A. L., & Krosnick, J. A. 2010. Social desirability bias in voter turnout reports: Tests using the item count technique. *Public Opinion Quarterly*, 4: 37-67.
- [11] Larkin, R., & Mahoney, A. 2006. Empowering youth to change their world: Identifying key components of a community service program to promote positive development. *Journal of School Psychology*, 44: 513-531.
- [12] Learn and Serve America. 2010. *What is Service-learning?* Retrieved March 16, 2010, from <http://www.servicelearning.org/what-service-learning>
- [13] Meyer, S. 2006. *Texas Center for Service-learning: Evaluation of K-12 SCP and CHESP programs*. Denver: RMC Research.
- [14] National Youth Leadership Council. 2011. *K-12 Service-learning Standards for Quality Practice*.

- <http://www.nylc.org/k-12-service-learning-standards-quality-practice>.
- [15] Reinders, H., & Yourmiss, J. 2009. School-based required community service and civic development in adolescents. *Applied Developmental Science*, 10: 2-12.
 - [16] Shadish, W. R., Navarro, A. M., Matt, G. E., & Phillips, G. 2000. The effects of psychological therapies under clinically representative conditions: A meta-analysis. *Psychological Bulletin*, 126: 512-529.
 - [17] Skinner, R., & Chapman, C. 1999. Service-learning and community service in K-12 public schools. Retrieved from the National Center for Education Statistics website: <http://nces.ed.gov/surveys/frss/publications/1999043/index.asp>.
 - [18] S. White. 2007. Inference by eye: Confidence intervals and how to read pictures of data. *American Psychologist*, 60: 170-180.
 - [19] Holbrook, A. L., & Krosnick, J. A. 2010. Social desirability bias in voter turnout reports: Tests using the item count technique. *Public Opinion Quarterly*, 4: 37-67.
 - [20] Moely, B. E., Mercer, S. H., Ilustre, V., Miron, D., & McFarland, M. 2012. Psychometric properties and correlates of the Civic Attitudes and Skills Questionnaire (CASQ): A measure of students' attitudes related to service learning. *Michigan Journal of Community Service-learning*, 8(2): 15-26.
 - [21] Reinders, H., & Yourmiss, J. 2009. School-based required community service and civic development in adolescents. *Applied Developmental Science*, 10: 2-12.