

Helping Others and Helping Oneself:
A Meta-Analysis of Service-Learning Programs

Christine I. Celio
Loyola University Chicago

Abstract

Service Learning (SL), the connection of community service with academic curriculum, has increased in popularity in recent years as exemplified by increased youth participation in service activities and increased funding for these programs. Previous research has found that SL programs have promoted several competencies, often focusing on civic, social, personal and academic engagement, but the results have been difficult to interpret because of methodological inconsistencies in the literature and the small sizes of the studies. Through meta-analysis, this paper evaluates SL research, examines the benefits derived by the providers, evaluates the effectiveness of these programs, and points to future directions in the field. This review finds that, overall, SL programs yield a small, positive and significant effect size, and that personal, civic, social, and academic competencies all produce small, positive, significant effect sizes, but were not significantly different from each other. Further, the “essential elements,” proposed in the SL literature, including having a logic model, community involvement, student planning and engagement, and reflection, did not yield significantly different effects than programs without these elements. These findings are explained and suggestions and limitations are discussed.

Introduction

Service-learning (SL) has become widespread in the United States. Although many definitions exist in the literature, in general, SL is a way to engage students in the learning process by having them provide meaningful service to others, and to connect this service experience with the students' academic curriculum (Fenzel & Leary, 1997; Giles, Honnet, & Migliore, 1991). In 1999, the National Center for Education Statistics of the U. S. Department of Education found that 32% percent of all public schools organized SL as part of their curriculum, including nearly half of all high schools (Skinner & Chapman, 1999). Beyond primary and secondary schools, community colleges and four-year universities also have a high level of involvement in SL programs (Campus Cares, n.d.). The popularity of SL programs has increased so much that in 2003, US News and World Reports started to list universities with SL programs in their annual "Best Colleges" issue (U.S. News and World Report, 2006). Further exemplifying the popularity of service learning, Campus Compact, an organization designed to infuse service and civic engagement into college academics, started in 1987 with just a handful of schools interested in service and now boasts 935 schools (Campus Compact, 2006).

Assisting in the popularity of these programs is the ubiquity of funding. Money is available through universities, corporations, private foundations and the federal government for these programs (e.g., Illinois Campus Compact, n.d.; ETR Associates, 2004; Learn and Serve America, 2006). Advice on how to start programs and examples of curricula are available on websites funded by groups ranging from the Disney Corporation to the federal government (Disney, n.d.; Learn and Serve, 2006). In January of 2003, President George W. Bush created the President's Council on Service and Civic

Participation which was established to “recognize the important contributions Americans of all ages are making within their communities through service and civic engagement.” (Learn and Serve, 2006). Administered by the Corporation for National and Community Service, the council encourages more Americans to get involved in their communities. This group offered \$40 million in Learn and Serve America funds to implement and support school-based, community-based, and higher education SL programs in 2006 (Learn and Serve, 2006).

However, despite SL programs’ popularity and the availability of funding, it is unclear what effects SL programs produce, and what contributes to the effectiveness of a program. To address these concerns, this paper will: 1) present the promising findings in the SL field; 2) describe the components presumed to be essential for an effective program; 3) discuss the need for a review of these programs; 4) explain the methodology of the current meta-analysis; 5) present the overall effectiveness of SL programs, specifically which outcomes are most affected, and the effectiveness of the theoretical essential elements of SL; and 6) discuss these findings and their limitations.

Promising outcomes from previous studies

Previous studies have focused primarily on the benefits derived by service providers personally, socially and civically, and academically.

Personal benefits. In reviewing SL programs, Conrad and Hedin (1989) noted that personal development is one of the “best documented” outcomes of secondary school sponsored community service programs. The major personal benefits accruing to students participating in SL programs have been increases in self esteem, empathy, positive attitudes toward the self, more highly internalized moral standards, beliefs that

one can make a differences in the world, and increases in self concept (e.g., Conrad & Hedin, 1982; Allen & Rushton, 1983; Berger 2002; Giles & Eyler, 1994).

Civic engagement. The impact of SL on students' civic engagement has also been promising. Involvement in SL has been found to increase social responsibility and moral reasoning (Conrad & Hedin 1982; Newmann & Rutter 1983) and increase positive attitudes toward the populations served (Conrad & Hedin, 1982). Finally, youth with SL experience are more likely to understand social issues and how to approach them than students who did not participate in SL (Markus, Howard & King, 1993; Eyler et al 1997).

Social Outcomes. SL has positively affected students' attitudes toward those others. Research suggests that following their service activities, college students are less likely to blame clients for their misfortune, more likely to stress a need for equal opportunity and to report more awareness of societal problems (Eyler et al., 1997). In addition, compared to non-SL classmates, SL students report increases in their tolerance and appreciation of others following a SL course (Markus, Howard & King, 1993).

Academic Outcomes. The effect of SL on academic outcomes has been similarly positive. Being involved in a SL program has been linked to increased knowledge about the subject and academic development (Hamilton & Zeldin 1987; Conrad & Hedin 1982). Markus, Howard & King (1993) found that course grades were significantly higher for those in the SL condition than those in the control group and that students did better academically when the service is integrated with traditional classroom instruction.

In summary, these SL programs suggest that SL shows promise in increasing social and personal competencies, promoting civic engagement and enhancing academic achievement for those providing the service.

Presumed essential elements of service-learning

In 1998, the National Service-Learning Cooperative created a list of “Essential Elements of Service-Learning” which outlines the basis for good SL programs (National Leadership Council, 1998). These “essential elements” are consistent with much of the literature in the field regarding components for effective SL programs and emphasize a few issues (National Service-Learning Clearinghouse, 2007). First, SL programs should be guided by some form of a logic model that contains, at the minimum, clear goals for the students and specifies the activities whereby students will reach these goals. Second, there should be student involvement, input, and engagement. That is, students should have a role in deciding what the SL project will be and how it will be conducted. Third, community input is important. Useful SL involves service that strengthens community ties, and creates strong partnerships with community groups based on mutually agreed upon goals, roles, and responsibilities (National Service-Learning Clearinghouse, 2007). Fourth, there should be a reflection component in the program so that students have adequate time to process the service experience. Reflection can involve asking students to keep a journal, having classroom discussions or small group reflections. These activities need to be planned and not left to chance (National Service Learning Clearinghouse, 2007). Presumably, the existence of these elements will lead to effective SL programs; however these assumptions have never been tested by comparing programs that do and do not contain these features.

Beyond these presumed essential elements, other information needs to be collected so that the findings can be generalized and the programs can be replicated. Some basic demographic information, like the ethnicity, gender and socio-economic

status of those providing the service, is helpful. In addition, details about the sites of service, the population being served, and the length, duration and intensity of the program, can help in evaluating the program. Finally, the magnitude of change achieved across studies, and the extent to which methodological limitations have affected the findings are all empirical questions worthy of attention. SL has the ability to be a useful tool to help youth integrate and connect their academics with practical service to others, but this information cannot be fully understood or appreciated if it is not collected and reported.

The Need for a Meta-analysis

Although some promising findings have appeared for SL programs, Billig (2000) has noted: “Research in the field of SL has not caught up with the passion that educators feel for it.” Despite the thousands of students that have been involved in SL at various ages and educational levels and the financial support that SL has received from local sources and the federal government, the empirical support for SL is unclear. Small sample sizes, unreliable measures, and inconsistent collection of data make findings in SL research difficult to generalize.

The purpose of this study is to conduct a meta-analysis of outcome studies on SL in order to evaluate the effectiveness of these programs, evaluate the empirical support for the recommended practices of SL, describe current SL programs, and discuss the implications of current research and practice.

Method

Identifying Studies

To find studies for the current meta-analysis, a literature search was performed in psychology (Psychinfo), medical (PUBMED) and education (ERIC) computer databases using the search words, “service-learning,” “community service,” “experiential learning,” “public service,” “civic engagement” and “civic involvement.” Further, reference lists from SL studies and books were inspected. Finally, leading experts of the SL community (Eyler, Ferrari) were contacted via email and asked for recommendations of studies to examine.

The 65 studies collected for this meta-analysis are believed to be a representative, non-biased sample of current evaluations of SL programs. To ultimately be included in the review, the studies met the following seven inclusion criteria: (1) appeared before May 1, 2006, (2) used a control group or one-group pre-post design, (3) contained sufficient information to calculate effect sizes, (4) involved students at the elementary, secondary or postsecondary level, (5) was reported in English, (6) fit the definition of SL defined earlier (the combination of service and academic curriculum) and (7) evaluated the SL course as the primary program (e.g., SL was not a component of another bigger program). All types of reports were eligible for inclusion, including published studies in academic journals, unpublished program reports, conference presentations, unpublished dissertations and theses, unpublished technical reports, and program evaluations available on the internet.

Coding Procedure

A coding manual was created to identify methodological and program characteristics of each study (see Appendix B). Each study had its identifying features recorded such as author, year of report, title, and source. Information about the service

provider and the population being served was coded, including demographics regarding gender, ethnicity and SES, as well as the location of the intervention and the grade of the students in the program. The dosage of the SL program was recorded, based on duration of project in weeks and hours. The studies were examined for the recording of pre and post data and if a comparison or randomized control group was used. Each outcome measure was coded according to which construct it was targeting (e.g., social learning, personal competency, academic performance, civic engagement), if its reliability and validity was reported, and information about the name of the measure, whether it was created for the study or taken from somewhere else. Further, the sample sizes were recorded as well as the methods used to calculate effect sizes. Data regarding the programs were assessed and the components of logic model, student engagement, community involvement and reflection were collected.

Description of coding manual

Outcome definitions. Study outcomes were organized in four main categories: social competencies, personal competencies, academic competencies and civic competencies. Social competencies were defined as anything having to do with how a person treats, thinks about, or works with other people, which included empathy, motivation, leadership skills, cultural competence, etc. Personal competencies were defined as anything having to do with how a person thinks about him or herself, or attitudes or thoughts on his or her own skills, which included self-esteem, self-efficacy, knowing how to do things, feeling in control, etc. Academic competencies included grades, test scores, academic engagement, or amount of learning taking place. Civic competencies were defined as any outcome measure having to do with helping the

“community” not encompassed by social competency, like civic values, civic attitudes, current and future voting behaviors, etc.

Essential elements definitions. In addition, the components of the “essential elements of SL” were defined in the following ways. To be counted as having a logic model, the study needed only to report having goals for the program and have corresponding activities to match those goals (e.g., the goal of having students volunteer at a home for the elderly was to allow students to bond with this group, and the goal was reported as “bonding with community,” or becoming more comfortable with the elderly, etc). Reflection activities included using journals, having discussions in class or in small groups regarding the service, writing essays, presenting to the class or reflecting with the teacher or supervisor at the site. Student engagement and planning were difficult to assess; when students were involved in the planning process of the program (e.g., students brainstormed to figure out where to do service), they were considered as being engaged. If other information was given to suggest that students were engaged in the process, like if they were required to do a certain amount of service and exceeded it, that was also considered “engagement.” However, this element was not very well defined in the literature and even more nebulous in its reporting. The fourth “essential element,” community involvement, was considered being fulfilled if the community was part of the planning or evaluation process or was indicated to have some part in the program besides providing a place for students to serve.

Index of Effect Size

Whenever possible, the Effect Size (ES) was calculated from the means and standard deviations reported in evaluations; otherwise, methods described by Lipsey and

Wilson (2001), and Durlak, Meerson, and Foster (2003), were used (e.g., p values, or t or F statistics). In cases involving multiple outcomes in one study, effects were averaged so that each program contributed only one effect per analysis when examining the overall ES (Durlak, Meerson, & Foster 2003). A random effects model was used, as recommended by Lipsey and Wilson (2001, p. 121). The distribution of effect sizes, and sample sizes, were inspected and outliers were windsorized (explained below). That is, values greater or equal to three standard deviations from the mean were trimmed to less extreme values.

Distribution of effects

Sixty-six studies (64 programs) were originally considered for this meta-analysis. Two papers included evaluation of two different populations, thus the study total was 66. However, under further analysis, one study was more than four standard deviations in both positive and negative effects. Leming's (2001) "Integrating a structured ethical reflection curriculum into high school community service experiences: Impact on students' socio-moral development" had outcome effect size outlier values of -4.13 and 14.6 and was not considered to be a representative study, so it was removed from final analysis. Thus, the final count of studies is 65. As recommended by Lipsey (2001), these data were windsorized. This process is one where the program effect sizes and study samples were evaluated and the means and standard deviations were assessed.

Under first analysis, the mean of the study effect sizes was 0.343, with a standard deviation of 0.836. The study effect sizes were windsorized so that no effect sizes were more than 3 standard deviations from the mean in either direction (larger than 2.51 or smaller than -0.343). For the studies larger than 2.51, they were given the value 2.51, and

the studies smaller than -0.343 were given the value -0.343. A similar process was done with the study sample sizes for the comparison and experimental groups. The experimental groups mean sample size was 106.23, with a standard deviation of 190.15. The range, with three standard deviations that the experimental N needed to be within, was between 676.68 and -464 (i.e., $106.23 \pm 3 \times 190.15$). Again, the larger values were given 676.68 as their value and there were no studies smaller than zero. The comparison group sample sizes had a mean of 99.05 with a standard deviation of 125.73. Thus, the largest study samples became 476 and there were no studies smaller than zero. The next step was to recalculate the mean and standard deviations of the new values and re-windsorize. After the first windsorizing procedure (with a range of 2.51 and -0.343), the mean effect size became 0.295, with a standard deviation of 0.474. The new range of mean effect sizes was 1.72 and -1.13. Any values larger than 1.72 became 1.72 and any below -1.13 became -1.13. The new experimental group mean study sample became 88.15 with a standard deviation of 102.47. No values fell below zero, so the largest values were windsorized to 395.57. The new comparison group mean became 71.14 with a standard deviation of 75.6. Again, there were no values below zero, so the largest study comparison group sample size was capped at 297.94.

Results

Characteristics of study sample

SL is only a recently studied field. Of the 65 programs included in this review, a majority of them were released in the last several years (Table 1.). Thirty- three (51%) were released since 2000, 28 (42%) in the 1990s, and only four (6%) in the 1980s. A majority of these studies (43, 66%) were published journal articles (Table 2.). However, the rest of the studies varied; 11 (16%) were unpublished conference papers, 9 (14%)

were unpublished dissertations, and the rest (2, 3%) were unpublished reports or book chapters (Table 2.).

The study designs of these programs need improvement. Although 46 (71%) of these studies measured the students before and after the service experience and 44 (68%) reported using a comparison group, just 14 (22%) of the studies reported using some sort of randomization procedure (Table 3.). Further, only 10 studies reported having both pre/post measures and a randomized control group (Table 3.). Twenty-nine (45%) reported some sort of reliability of the measures in their studies, and very few (9, 14%) reported any sort of validity. No studies reported follow-up data (Table 3.).

The data was reported in a variety of ways. Most of the outcomes were reported in mean and SD (201 outcomes, 51%) (Table 4.). Ascertaining effect sizes from p-values came from 106 outcomes (27%), f-values were used for 40 outcomes (10%), 7 outcomes (2%) were from chi-squares, 4 (1%) were from T and mean, 4 (1%) were from a d-value, and mean change, percent ratio, and non-significant findings resulted in one outcome each, totaling less than 1% each (Table 4.). The number of outcomes reported per study varied from 1 to 26 outcomes, with a mean a little less than 6 outcomes per study and a standard deviation of 5.2 (Table 5).

Reporting on measures was fairly good. All measures and outcomes were reported in 43 (69%) of the studies. Ten (15%) of studies had measures mentioned in the methods section that were left out of the results, 11 (14%) had no way to tell whether some results were left out and only 2 studies (3%) reported just the significant findings. Twenty-two of the studies reported results from individual items or questions (e.g., “Do you plan to vote in the next election?”), 20 (31%) reported using a full measure (e.g, Coopersmith

Self-esteem Scale), 19 (29%) reported using an entire subscale of a measure, 10 (15%) used an academic or behavioral rating (e.g., grades, truancy, etc), 8 (12%) reported using single items from a subscale of a measure and 2 (1.5%) reported using raters to evaluate the content of qualitative journals (Table 6.).

Program characteristics

Very little information was given about the type of service being done in these programs. Twenty-four (37%), reported that multiple types of service (like mentoring, working with people with developmental disabilities, working at a soup kitchen, etc) occurred in their programs but none of the studies separated effects of the program by type of service (e.g., comparing students who worked with the elderly with students who worked with children) (Table 7.). Eight (12%) programs did not report any information about the type of service in which students were engaged. Of the programs that did include more specific information, 9 (14%) programs served younger students in either a tutoring or mentoring capacity, 9 (14%) programs worked with elderly people, 8 (12%) programs were in a health setting (hospitals, mental health programs, etc), 2 (3%) programs served the school community, 2 (3%) programs served individuals with developmental disabilities and one program worked with a governmental agency (Table 7.). Beyond naming the population with whom they worked, very little additional information was given about the type of service done (e.g., games played in mentoring programs, details of visits, etc). These details could greatly aid replication.

Almost half of the studies (30, 46%) indicated an interpersonal type of service, where students appeared to directly interact with the people they served (Table 8.). Unfortunately, many (24, 37%) of the studies did not report the type of service in which

the students were engaged, or had varied settings (9, 14%) so the interaction style could not be assessed (Table 8.). The two remaining programs were either organizational or environmental and did not work directly with people (Table 8.). Overall, most (38, 59%) studies did not report enough information to assess whether service was done individually or in groups (Table 9.). Of those that did report, 16 (25%) of the studies served individually, 7 (11%) programs involved both individual and group experiences, and just 4(6%) programs did service in groups.

Sample characteristics

According to this sample, SL programs are most often evaluated in university settings. Most of the programs involved college students (41, 63%) (Table 10.). Nine (13.8%) programs were done with high school students, 8 (12.3%) with professional/graduate students (pharmacy students, medical students, etc), 3 (4.6%) with junior high students, 2 (3.1%) with middle school students (grades 4-6), 1 (1.5%) with elementary school students, and one (1.5%) study had a mix of students (Table 10.).

The rest of the demographic characteristics were difficult to assess because of reporting. The characteristic best reported upon, gender, showed up in 43 (66%) of the studies. Of the studies that did report gender, the programs were predominantly female (mean = 62.4%) (Table 11.).

Information on race/ethnicity was very poorly reported upon (Table 12.). Over half (33, 51%) did not present information on race/ethnicity. Only 31 (48%) of studies reported how many white students were involved, and of those studies that did report, the programs were predominantly white (mean=63.4%). Only 22 (34%) reported how many black students were involved, with the studies reporting having little involvement of

black students involved (mean= 12.5%). Similarly few studies(21, 32%) reported participation of Latino students, with those studies reporting having a mean of 15.4%. Involvement Asian students was 17 (26%), with the studies reporting having a mean of 5.8%, and only 16 (25%) reported on the amount of Native American students (mean= 1.5%).

Reporting on other demographic characteristics was also poor. Only 9 (14%) of the studies reported anything about the socio-economic background of the students serving. Of those reporting, most (5, 56%) were students from middle or upper class communities (Table 13.). In terms of location, only 20 studies (30%) mentioned the location of the programs (rural, suburban, urban, etc) (Table 14.). In light of poor reporting, it is difficult to generalize many of these findings to any particular population.

Program Effectiveness

Overall, SL programs appear to be effective. SL programs produce a small, positive overall significant effect size (mean= 0.28, $p= 0.00$) when multiples outcomes from each study are averaged to yield one effect size per study (Table 15.). In addition, SL programs produce significant effects in four different outcome categories: social competencies, civic competencies, personal competencies and academic competencies. Each of these mean effects was small in magnitude, but positive and statistically significant from zero. Personal competencies (self-esteem, self efficacy, etc) were measured in 39 (60%) studies and had a mean effect size of 0.26 ($p=0.00$). Social competencies (empathy, leadership skills, etc) were measured in 32 (49%) studies and had a mean effect size of 0.26 ($p=0.00$). Civic outcomes (civic engagement, altruism, etc.) had a similar effect size (0.28, $p=0.00$) and were measured in 32 (49%) studies.

Finally, academic outcomes (class grades, test scores, etc) were measured in 23 (35%) of the studies and has an effect size of 0.27 ($p=0.00$). The mean effects for these 4 outcome categories did not differ significantly from each other (Table 16).

Essential element effectiveness overall

SL programs rarely report following the proposed “essential” elements of SL. With the exception of using reflection, programs did not apparently follow recommended practices (i.e., contain essential elements) with any regularity. When measuring programs containing a single element against programs not containing that same element, there was no significant difference between the two. That is, although having an “essential” element, in all cases but one, resulted in a significant, small, positive effect size, not having the element produced just as significant and strong of an effect size.

Logic model. To be considered to have a logic model, a program needed to clearly state the goals of the program and report activities that logically met these goals. Only 24 (36.9%) of the studies included enough information to be considered having a logic model. The studies that did include this information had a mean effect size of 0.28 ($p=0.00$). However, the studies that did not include enough information to be considered as having a logic model had the same mean effect size (0.28, $p=0.00$) and were not significantly different from those studies that reported a logic model (Table 17.).

Student engagement and planning. Student engagement and planning were both difficult to assess in these studies. Studies that reported having students plan the service were considered as having student engagement. In addition, engagement was defined by demonstrating that the students were involved in the planning process of the program. A little over one-third of the studies (23, 35.4%) reported some sort of student engagement,

however student planning was not often mentioned (6 studies, 9.2%). Both student planning and the combination of student planning and engagement had a mean effect sizes of 0.29 and were not significant ($p=0.06$). Although programs with student engagement had a significant mean effect size of 0.30 ($p=0.00$), this was not significantly different from programs without student engagement (Table 18.).

Community involvement. Communities were rarely reported to be involved in these programs. Only 10 (15.4%) of the studies reported that communities were involved in the planning, implementation or evaluation of the programs. Although programs with community involvement had a significant positive effect size (0.23, $p=0.04$), they were not significantly different from programs that did not report community involvement (Table 19).

Reflection. Reflection, by far, was the ‘essential element’ that was most often reported. Forty-seven (72.3%) of the programs reported reflection of some type, whether it was in-class discussion, reflection essays, journals, or small group reflection. The effect of reflection was small, positive and significant (0.24, $p=0.00$) but, again, not significantly different from programs that did not report reflection activities (0.38, $p=0.00$) (Table 20).

When looking at the individual outcome categories, the trend continued. No program with the individual elements of having a logic model, student engagement and planning, community involvement, or reflection were significantly different from programs without those individual elements. When separated into social, personal, civic and academic outcomes, the programs that had the individual ‘essential’ elements (just

logic model, etc) were not significantly different from the programs that did not report these elements (Tables 21-24).

Discussion

Because this is the most comprehensive review of the empirical studies pertaining to service-learning, and the only meta-analysis of SL programs, this review considerably contributes to the literature. It provides evidence for the effectiveness of these programs and supports the assumptions made by those in the field that these programs are worth funding. The findings suggest that SL programs are not only effective in helping students connect to others evidenced through increases in social competencies, but also supports the notion that helping others does, indeed, help oneself, by assisting in the increase of varied personal competencies, like self esteem and self efficacy. Additionally, the ramifications for society are positive: since these programs increase civic competencies, it is possible to make students feel a sense of responsibility toward their communities through the implementation of these programs. And, finally, since the No Child Left Behind Act is such a strong force in education today, this meta-analysis is especially important, as it provides support for the idea that focusing on service has the potential of helping students increase their scores and grades. It is the author's hope that high quality SL programs will continue to be implemented in schools and these positive effects will persist.

Limitations

At least four limitations exist in the current review qualify the conclusions:

Methodology. First, study methodology could have contributed to the effect sizes. An examination of methodological factors, like the types of measures used, the validity and reliability of these measures, and use of control groups and pre/post measures, is

important to examine. For example, in this review, only 10 (15.4%) studies had pre/post measures and used a randomized experimental design. These data alone suggest that the lack of methodological rigor of these studies makes it difficult to generalize the results. However, it is promising to note that the studies with randomized control groups and pre/post measures did not significantly differ from those groups without both characteristics, and that the groups without pre/post measures or randomized control groups actually had a lower mean effect size (0.27 vs 0.38, $p < 0.05$). In addition, when attempting a regression using methodology variables (pre/post measures, randomized control groups, reliability and validity), nothing was found to be significant.

Reporting. Second, low rates of reporting these program characteristics make it difficult to conduct a full assessment of many programs. Most importantly, details were simply not reported to gauge the application of presumably essential components. Whenever an element or feature of the program was not directly commented upon, it was rated as not having it, as opposed to “not reported.” It is possible that these programs did have logic models, student engagement and planning, community involvement or reflection, but the authors did not describe these features in their report.

Program Characteristics. Third, the difference in effect sizes could have been based on differences in other program characteristics, including duration and intensity of the program, with whom the students worked, who the students were, where the program was held, or the type of service done. For example, the programs varied in terms of program intensity and duration (Table 25). Only 40 (61.5%) programs reported the length of the program in hours, which averaged to be 33.9 hours per program, with a standard deviation of 36.4 hours and ranged between 1 and 180 hours. The duration of program in

weeks was similarly varied, with 50 programs reporting the amount of weeks involved, or that the program was a year (coded as 36 weeks) or semester (coded as 18 weeks) long. The mean number of weeks that the program was run was 16 with a standard deviation of 10 weeks. Although the program might have reported a certain number of hours or weeks, it is not clear if those hours were service or preparation hours, or if service was done each week. This variability of duration and intensity of program could have influenced the effectiveness of these programs.

Types of Studies. Fourth, as is the nature of meta-analysis, the only studies that were included had quantitative measures. It is possible that these quantitative studies are not capturing the full value of SL programs. Many good qualitative studies exist in the SL literature and should be examined for the lessons they might suggest.

Effective Practices

The impact of what has been suggested as effective practices or “essential elements” is difficult to assess. As mentioned earlier, the lack of reporting on the practices of student planning and engagement, community involvement, reflection, and inclusion of a logic model creates a problem when attempting to assess the usefulness of these practices. However, of the four elements, reflection did show up the most often and does seem to contribute to the effectiveness of the programs.

Conclusions and Future Recommendations

Overall, this review provides evidence that SL programs are effective, specifically in affecting social, civic, personal, and academic outcomes. The evidence for the value of some presumably essential elements is lacking, however. It is unclear at this point if the findings in this meta-analysis are because of the variability of reporting across studies,

insufficient methodological rigor, differences in program characteristics, or the types of studies being used.

Service-learning is a worthwhile field, but it can be improved. Programs continue to be created and run throughout the nation, and to increase their effectiveness, practitioners need to evaluate their programs. Recommendations for SL evaluations exist (see Celio & Durlak, 2007), and these evaluations can greatly enhance the field. Only through more complete reporting and evaluation can practitioners and researchers be certain about what programs are most effective and how those effective programs can be shared and replicated.

Appendix A.

Meta-Analysis Coding For Service-learning Program Evaluations 4.2.2007

95 = Not Applicable

97 = not reported

1= yes

2=no

SECTION ONE: Report Characteristics

- 1) ID
- 2) Author
- 3) Title
- 4) Year
- 5) Are there pre post measures?
- 6) Is there a control group?
- 7) Is the Control Group Randomized?
 - a. Describe randomization procedures
- 8) Type of study
 - 1= intervention
 - 2= survey of students involved in SL programs
 - 3=archival data of SL intervention
 - 96=Other
- 9) Source of report
 - 1 = Published journal article
 - 2 = Unpublished conference paper
 - 3 = Unpublished dissertation
 - 4 = Unpublished masters thesis
 - 6 = Unpublished technical report, program evaluation
 - 7 = Unpublished report available on Internet
 - 8= Article in a book
- 10) Who was being served?
 - 1= varied
 - 2=tutoring/school programs
 - 3=developmental disability
 - 4=elderly
 - 5= health care/hospital/Mental Health
 - 6= government agency
 - 7= school community (like recycling projects)
 - 97 = not reported
- 11) How many outcomes were being measured/reported?
- 12) What was the location of the study (which university/area of the country)
 - 1=urban
 - 2=rural
 - 3= suburban

4= varied
97 =not reported

- 13). What number of interpretable sessions are involved in the service experience?
- 14). What is the average length of each session (in minutes)
- 15). What is the total length of the service experience (in hours)? If there is a range, take the average. If they say an hour a week for a semester, give 18 hours of credit.
- 16). What is the duration/course of the service experience from beginning to end (in weeks) Note. A school semester is 18 weeks; A school year is 36 weeks. If they say it is a “semester long course” give credit for 18 weeks. If they indicate that it might not last the entire semester, give a “97”
- 17) Reflection component
 - 1) journals
 - 2) class discussion
 - 3) class presentation
 - 4) small group discussion
 - 5) talk with teacher/supervisor
 - 6) combination of reflection activities
 - 7) essay
 - 97) not reported
- 18) Was their collaboration with the community about the service project?
 - 1) Yes
 - 2) No
- 19) Were the tasks reported to challenge the students in some way/engage them?
 - 1) Yes
 - 2) No
- 20) Were there explicit goals for the students in the service projects (eg, the goal of having students read to Latino children in Spanish was to allow them to bond with Latino culture, the goal is reported as bonding with community) [logic model part 1]
 - 1) Yes
 - 2) No
- 21) Were activities specified to meet these goals? [logic model part 2]
 - 1) Yes
 - 2) No
- 22) Were both parts of the logic model met?
 - 1) Yes
 - 2) No
- 23) Were students involved in the planning of this program?
 - 1) Yes
 - 2) No

24) Was the service done individually?

- 1) yes
- 2) no; as a group
- 3) Both

25) Service interaction?

- 1) interpersonal (tutoring, elderly, developmental disability, working in a pharmacy)
- 2) environmental (cleaning up trails, etc)
- 3) organizational (food drive)
- 4) administrative (soup kitchen)
- 5) Varied

25a) Description of program? (string)

26)) Who developed the measure(s)?

- 1=not created for this study
- 2=created for this study
- 3=grades
- 4=combination
- 5= other (behavioral measure, etc)
- 97 = not reported

26a) If not created by the researcher, what is the name of the measure?

SECTION TWO: Study Characteristics

27) Total N

28) Age group/grade

- 1=elementary (grades 1-3)
- 2=middle school (grades 4-6)
- 3=junior high (grades 7-9)
- 4=high school (9-12)
- 5=undergrads
- 6=grad students/professional school students (school of nursing, dental school, med school)
- 7=mixed
- 97 = not reported

29) Gender (% female)

- 51=majority
- 97 = not reported

30). General Program Descriptor : who were they working with? (string)

Indicate the percentage of the sample by race/ethnicity

31)Caucasian

- 51=majority
- 97 = not reported

32) African American

- 51=majority
- 97 = not reported

- 33) Latino/Hispanic
 - 51=majority
 - 97 = not reported
- 34) Asian American/ Pacific Islander
 - 51=majority
 - 97 = not reported
- 35) Native American, Native Alaskan/Eskimo
 - 51=majority
 - 97 = not reported
- 36) International
 - 51=majority
 - 97 = not reported
- 37) Ethnic Minority Groups Total
- 38) SES Status (Use author's descriptions of this variable)
 - 1= Low SES (at or < poverty line)
 - 2 = Working class or lower middle class
 - 3 = Middle or upper class
 - 4 = Mixed- including lowest
 - 5 = Mixed- excluding lowest SES
 - 96= Other
 - 97= Not reported

OUTCOME CODING

- 39a) outcome number (which outcome it is)
- 39b) all measures reported in the study?
 - 1) all items reported/ all measures
 - 2) only significant findings reported
 - 3) no way to tell; measures either not explained or not all items shown, some not used for no explained reason
 - 4) no, this measure was left out
 - 5) just took items from measures (or it was adapted in some way), no way to know why other items not taken
- 39c) Is outcome an individual item or the entire measure?
 - 1) item; name of measure
 - 2) measure; name
 - 3) subscale/group of items of measure
 - 4) something academic or behavioral (grades or dropping a class)
 - 5) rating of content in journal
 - 6) item(s) from an adapted subscale of a measure
- 39d) The measure is created by researcher or published somewhere else?
 - 1) researcher made
 - 2) somewhere else
 - 3) something academic
 - 4) something behavioral (dropping classes)
- 39d) Reported reliability?

- 1) Yes; good
- 2) Yes: acceptable (according to author)
- 3) Yes; bad
- 4) not applicable
- 5) reported but varied
- 6) reported, but used only a subscale or adapted the measure, so the reliability may not be the same
- 7) interrater reliability
- 97) Not reported

39e) Reported validity?

- 1) Yes; good
- 4) Not Applicable
- 6) reported, but used only a subscale or adapted the measure, so the validity may not be the same
- 97) Not reported

39) Outcome Name (string)

40) Outcome Category version 1

- 1) Social competency skills:
 - anything having to do with how you treat, think about, or work with other people
 - (social responsibility; concern for other's welfare; attachment to others; motivation; empathic reasoning, prosocial reasoning; perspective taking, prosocial reasoning and cultural competence/comfort with other populations, analytical skills/decision making with social problems)Note: the “personal and social responsibility” measure ends up being more social responsibility, so that is how it was coded.
- 2) Personal Competency:
 - anything having to do with how you think about yourself, attitudes about yourself or thoughts on your own skills (self-esteem, self-efficacy, knowing how to do things, feeling in control)
- 3) Civic Engagement:
 - anything having to do with helping the “community” not encompassed by social competency (civic values, civic attitudes, current behaviors)
- 4) Academic
 - (course or assignment grades, standardized tests, academic engagement/ “doing well”, amount of learning taken place)
- 5) Problem behaviors (absences, sexual behavior, etc)
- 6) Other
- 9) Future behaviors
 - (intending to do something in the future, like “intention to vote in next election”)
- 11) Attitudes toward service learning

(SL should be done in the future, importance of SL, we should do service (shoulds)/volunteer/ experiential learning, etc)

- 12) knowledge/new awareness
(role taking, attitudes toward/appreciation for others, community awareness, perspective taking, learning)(encorporate # 7, which is meaningfulness/satisfaction prosocial attitudes about others (I like old people)
- 15) Teacher/administrator opinion
(eg, instructor satisfaction) and others ratings of the student's service
- 16) Peer ratings
- 17) Thoughts about course
(I liked this class, would recommend it to others)
- 19) Career related skills
(I know how to apply for a job, how to plan a project, etc).

When reduced for “newoutcomes2” others opinions were combined into the category they were talking about, most “11”s were encorporated into civic engagement/attitudes.

Big group Codes

- 1) Social competencies
(now includes, civic, new awareness about others, career skills, etc, attitudes toward SL, future behaviors concerning other people)
- 2) Personal competencies
(new awareness toward the self, career skills related to personal skills, etc)
- 3) Academic (includes behaviors that influence academics, like dropping a class)
- 4) Other
- 5) Civic (knowledge of social programs, altruism, etc)

41. Type of data used in analysis

- 1=f
- 2=mean and sd
- 3=T
- 4= Beta
- 5= p-value
- 6=% ratio
- 7=arcsine
- 8= d value
- 9= ns value
- 10= mean change
- 11= mean with no sd
- 12= t and mean
- 13=chi square

42. Pre Experimental N

43. Post Experimental Group N

44. Pre Control Group N

45. Post Control Group N

- 46) Post D
- 49) Follow up data? (yes/no)
- 50) How long after (weeks)

Appendix B.

Table 1.

Date of Programs

Date of Program	N	%
1980-1989	4	6
1990-1999	28	43
2000-2006	33	51

Table 2.

Source of Report

Variable	N	Mean ES	95% CI
Published journal article	43	0.28*	0.19 - 0.36
Unpublished conference paper	11	0.31*	0.13 – 0.48
Unpublished dissertation	9	0.19*	0.00 - 0.38
Unpublished technical report, program evaluation	1	1.70*	1.00 - 2.44
Article in Book	1	-0.20	-0.61 - 0.21

*p<0.05

Table 3.

Methodology

	N	Mean ES	95% CI
Had Pre/Post measure	46	0.27*	0.17 – 0.36
Did not have Pre/Post measure	19	0.31*	0.17 – 0.46

Had comparison group	44	0.31*	0.21 - 0.41
Had no comparison group	21	0.22*	0.08 - 0.36
Randomized control group	14	0.36*	0.18 - 0.54
Not randomized	30	0.29*	0.17 - 0.41
Not applicable (no control group)	21	0.22*	0.08- 0.36
Randomized with pre/post	10	0.38*	0.162 – 0.588
Having either randomized design or pre/post	55	0.27*	0.177 – 0.352
Reported reliable measures	29	0.31*	0.18 - 0.44
Not reported/not reliable	40	0.23*	0.15 - 0.37
NA (school grades, etc)	9	0.15	-0.06 - 0.36
Reported valid measures	9	0.40*	0.17 - 0.63
Not reported/not Valid	55	0.28*	0.20 - 0.37
NA (school grades, etc)	9	0.15	-0.04 - 0.34
Reported follow-up data	0	0	0

*p<0.05

Table 4.

Data characteristics

<i>Type of Data</i>	N	Mean ES	95% CI
F	7	0.38*	0.17- 0.59
Mean and SD	31	0.18*	0.08 - 0.28
T	7	0.53*	0.32 - 0.74
P-value	14	0.18*	0.04 - 0.31

% ratio	1	0.32	-0.10 - 0.77
Mean change	1	1.72*	1.01 - 2.43
T and Mean	2	0.65*	0.17 - 1.13
Chi square	4	0.30*	0.08 - 0.52

*p<0.05

Table 5.

Number of outcomes

	N	Mean ES	95% CI
2 or under	29	0.34*	0.22 - 0.46
3-5	11	0.15	-0.04 - 0.35
6-10	14	0.26*	0.09 - 0.44
11-30	11	0.26*	0.08 - 0.45

*p<0.05

Table 6.

Use of Item or measure

	N	Mean ES	95% CI
Item	22	0.32*	0.19 - 0.45
Measure	20	0.21*	0.07 - 0.35
Subscale	19	0.21*	0.06 - 0.36
something academic or behavioral	10	0.18*	0.01 - 0.36
Rating from content of journal	2	0.84*	0.36 - 1.33
Items from an adapted subscale of a measure	8	0.17	-0.07 - 0.40

*p<0.05

Table 7.

Population being served

	N	Mean ES	95% CI
Varied	25	0.24*	0.12 - 0.35
Younger	9	0.28*	0.09 - 0.48
Developmental Disability	2	0.88*	0.40 - 1.35
Elderly	9	0.28*	0.06 - 0.49
Health	8	0.32*	0.10 - 0.54
Government agency	1	0.56	-0.08 - 1.20
School community	2	0.03	-0.52 - 0.59
Not reported	9	0.24*	0.04 - 0.45

*p<0.05

Table 8.

Type of service interaction

	N	Mean ES	95% CI
Interpersonal	30	0.33*	0.21 - 0.45
Organizational/Environmental	2	0.33	-0.18 - 0.85
Varied	9	0.21	-0.02- 0.44
Not reported	24	0.24*	0.11 - 0.36

*p<0.05

Table 9.

Individual versus group service

	N	Mean ES	95% CI
--	---	---------	--------

Individual	16	0.20*	0.03- 0.37
Group	4	0.13	-0.17- 0.43
Both	7	0.39*	0.16 - 0.63
Not reported	38	0.30*	0.20 – 0.41

*p<0.05

Table 10.
Age/year in school

	N	Mean ES	95% CI
Elementary (grades 1-3)	1	0.01	-0.76 - 0.77
Middle school (grades 4-6)	2	0.09	-0.27 - 0.45
Junior high (grades 7-9)	3	0.17	-0.13 - 0.46
High school (9-12)	9	0.42*	0.22 - 0.61
Undergraduates	41	0.28*	0.18 - 0.37
Professional/Graduate students	8	0.34*	0.12 - 0.56
Mixed group of students	1	-0.21	-0.66 - 0.25

*p<0.05

Table 11.

Gender representation in studies

	Number of programs reporting gender	Mean % of N of females in studies when included	Range	Standard Deviation
Female	46 (71%)	64.2	40-100	15.67

Table 12.

Race/Ethnicity

Ethnicity	Number of programs reporting this race/ethnicity (%)	Mean % of N of ethnic group in studies where included
White	31 (48%)	63.4%
Black	22 (34%)	12.5%
Latino	21 (32%)	15.4%
Asian	17 (26%)	5.8%
Native Americans	16(25%)	1.5%
Any	32 (49%)	N/A

information on race/ethnicity

Table 13.

Socio-economic status

SES	N	%
Low SES	1	1.5
Lower Middle Class	1	1.5
Middle or upper class	5	7.7
Mixed, including lowest	1	1.5

Unspecified 56 86

Table 14.

Location of service

Location

Not specified	45	69
Suburban	7	11
Urban	6	9
Varied locations	4	6
Rural	3	4.5

Table 15.

Overall Effect Size of Service-learning Programs

Mean ES	N	95% Confidence Interval	P	Range
0.2773	65	0.20 - 0.35	0.00	-0.70 – 1.72

Table 16.

Outcome Category Effect Sizes

Outcome Category	Mean ES	N	95% Confidence Interval
Personal	0.32*	39	0.19 – 0.46
Civic	0.28*	32	0.14 – 0.42

Academic	0.27*	23	0.10 – 0.43
Social	0.26*	32	0.11 – 0.41

*p<0.05

Table 17.

Effect of Logic Model on Overall Effect Size

Logic model?	Mean ES	N	95% Confidence Interval
Yes	0.28*	24	0.14 – 0.42
No	0.28*	41	0.18 – 0.38

*p<0.05

Table 18.

Effect of Student Planning and Engagement on Overall Effect Size

Variable	Mean ES	N	95% Confidence Interval
Student planning	0.29	6	-0.01 – 0.59
No Student planning	0.28*	59	0.20 – 0.36
Student Engagement	0.30*	23	0.16 – 0.44
No Student Engagement	0.27*	42	0.1698 – 0.3716
Both Engagement and Planning	0.29	6	-0.01 – 0.59

*p<0.05

Table 19.

Effect of Community Involvement on Overall Effect Size

Variable	Mean ES	N	95% Confidence Interval
----------	---------	---	-------------------------

Community involvement	0.23*	10	0.01 – 0.45
No community involvement	0.29*	55	0.20 – 0.38

*p<0.05

Table 20.

Effect of Reflection on Overall Effect Size

Variable	Mean ES	N	95% Confidence Interval
Reflection	0.24*	47	0.15 – 0.33
No Reflection	0.38*	18	0.23 – 0.53

*p<0.05

Table 21.

Logic model by outcome category

Outcome category	N	Mean ES	Confidence Interval
Social with logic model	15	0.40*	0.19 – 0.62
Social without logic model	17	0.14	-0.06 – 0.34
Personal with logic model	15	0.30*	0.05 – 0.54
Personal without logic model	24	0.35*	0.16 - 0.55
Civic with logic model	10	0.24*	0.05 – 0.42
Civic without logic model	22	0.27*	0.15 - 0.39
Academic with logic model	8	0.25	-0.06 – 0.57
Academic without logic model	15	0.28*	0.07 – 0.50

(*p<.05)

Table 22.

Student planning and engagement by outcome category

Outcome category	N	Mean ES	Confidence Interval
Social with planning/engagement	3	0.06	-0.13 – 1.13
Social without planning/engagement	29	0.24*	0.08 – 0.39
Personal with planning/engagement	4	0.13	-0.35 – 0.62
Personal without planning/engagement	35	0.35*	0.19 – 0.51
Civic with planning/engagement	4	0.27	-0.05 – 0.60
Civic without planning/engagement	28	0.26*	0.15 – 0.36
Academic with planning/engagement	3	0.47	-0.08 – 1.03
Academic without planning/engagement	20	0.25*	0.07 – 0.44

(*p<.05)

Table 23.

Community involvement by outcome category

Outcome category	N	Mean ES	Confidence Interval
Social with community involvement	6	0.39*	0.01 – 0.77
Social without community involvement	26	0.23*	0.07 – 0.40
Personal with community involvement	6	0.15	-0.25 – 0.55
Personal without community involvement	33	0.36*	0.20 – 0.53
Civic with community involvement	7	0.23	-0.02 – 0.48
Civic without community involvement	25	0.27*	0.16 – 0.38
Academic with community involvement	3	0.45	-0.07 – 0.97

Academic without community involvement 20 0.25* 0.06 – 0.44

(*p<0.05)

Table 24.

Reflection by outcome category

Outcome category	N	Mean ES	Confidence Interval
Social with reflection	25	0.25*	0.08 – 0.42
Social without reflection	7	0.30	-0.04 – 0.64
Personal with reflection	31	0.34*	0.17 – 0.52
Personal without reflection	8	0.29	-0.05 – 0.62
Civic with reflection	23	0.29*	0.17 – 0.41
Civic without reflection	9	0.20*	0.03 – 0.38
Academic with reflection	18	0.23*	0.03 – 0.42
Academic without reflection	5	0.45*	0.07 – 0.84

*p<0.05

Table 25.

Program characteristics

Variable	N	Mean	SD
Number of sessions	10	17.4	18.0
Average length of session (minutes)	13	138.6	77.8
Total length of program (hours)	40	33.9	36.4
Duration of program (weeks)	50	16	10.0

Table 26.

Program Size

Program Size	Mean	Standard Deviation	Range
Amount of Participants	187.06	249.42	25-1374

References

- Allen, N. J., & Rushton, J. P. (1983). Personality characteristics of community mental health volunteers: A review. *Journal of Voluntary Action Research, 12*(1), 36-49.
- *Angiullo, L., Whitbourne, S. K., & Powers, C. (1996). The effects of instruction and experience on college students' attitudes toward the elderly. *Educational Gerontology, 22*(5), 483-495.
- *Barbee, P. W., Scherer, D., & Combs, D. C. (2003). Prepracticum service-learning: Examining the relationship with counselor self-efficacy and anxiety. *Counselor Education and Supervision, 43*(2), 108-119.
- *Batchelder, T. H., & Root, S. (1994). Effects of an undergraduate program to integrate academic learning and service: Cognitive, prosocial cognitive, and identity outcomes. *Journal of Adolescence, 17*, 341-355.
- Berger, J. B., & Milem, J. F. (2002). The impact of community service involvement on three measures of undergraduate self-concept. *NASPA Journal, 40*(1), 85-103.
- *Berson, J. S., & Younkin, W. F. (1998). *Doing well by doing good: A study of the effects of a service-learning experience on student success. ASHE annual meeting paper.* U.S.; Florida.

- Billig, S.H. (2000). Research on K-12 school based service-learning: The evidence builds. *Phi Delta Kappa*, 81(9), 650-664.
- *Blieszner, R., & Artale, L. M. (2001). Benefits of intergenerational service-learning to human services majors. *Educational Gerontology*, v27(n1), p71.
- *Blyth, D. A., Satino, R., & Berkas, T. (1997). A quantitative study of the impact of service learning programs. In A. S. Waterman (Ed.), *Service learning: Applications from the research* (pp. 39-56). Mahwah, NJ: Lawrence Erlbaum Associates.
- *Bringle, R. G., & Kremer, J. F. (1993). *Evaluation of an intergenerational service-learning project for undergraduates*
- *Brody, S. M. (2003). *Serving a higher purpose: How service learning leads to improved intergroup relations*. Unpublished Ph.D., University of California at Santa Cruz.
- *Brush, D. R., Markert, R. J., & Lazarus, C. J. (2006). The relationship between service learning and medical student academic and professional outcomes. *Teaching and Learning in Medicine*, 18(1), 9-13.
- Campus Cares (n.d). Retrieved June 16, 2006, from <http://www.campuscares.org/resources/national.html#service>
- Campus Compact. (2006). *2004 Service statistics: The engaged campus*. Retrieved June 16, 2006 from <http://www.compact.org/about/statistics/2004/>

Celio, C.I., & Durlak, J.A. (in press) Service-learning: Learning by doing for others.

Chapter to appear in T. P. Gullotta & M. Bloom (Eds.). *A blueprint for promoting academic and social competencies in after school programs*. New York: Springer.

*Chowdhary, U. (2002). An intergenerational curricular module for teaching aging appreciation to seventh-graders. *Educational Gerontology*, v28(n7), p553.

*Cohen, J., & Kinsey, D. F. (1994). "Doing good" and scholarship: A service-learning study. *Journalism Educator*, v48(n4), p4.

Conrad, D. & Hedin, D. (1982). The impact of experiential education on adolescent development. *Child and Youth Services*, 4(3-4), 57-76.

Conrad, D., & Hedin, D. (1989). *High school community service: A review of research and programs* (Information Analyses. Madison, WI: National Center on Effective Secondary Schools.

*Conner, D. B. (2004). The effects of course-related service projects in a child development course. *College Student Journal*, 38(3), 462.

*Curran, J. M. (1999). Changing attitudes toward people with mental retardation: Effects of extended contact within a changing social context. *Biennial Meeting of the Society for Research in Child Development*, Albuquerque, NM.

*Daugherty, M., & Compton, J. (1996). Community service academics and at-risk high school. *High School Journal*, 80(1), 39-48.

*Deiger, M. (2005). *The impact of service learning on cognitive complexity*. Unpublished Ph.D., Loyola University Chicago.

Disney. (n.d). *Welcome to the Disney Channel and Cable in the Classroom learning to serve website!* Retrieved June 16, 2006, from <http://psc.disney.go.com/disneychannel/learningtoserve/index.html>;

*Dorfman, L. T., Murty, S. A., Ingram, J. G., Evans, R. J., & Power, J. R. (2004). Intergenerational service-learning in five cohorts of students: Is attitude change robust? *Educational Gerontology*, 30(1), 39-55.

*Drake Dones, T. R. (1999). *Effects of a service learning program on sense of community and perceptions of control in elementary school students*. Unpublished Ph.D., Indiana University.

Durlak, J. A., Meerson, I., & Foster, C. J. E. (2003). Meta-analysis. In J. C. Thomas, & Hersen (Eds.), *Understanding research in clinical and counseling psychology*. (pp. 243-267). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.

ETR associates. (2004). *Funding Sources - Awards, Scholarships, Funders, Grants*. Retrieved June 16, 2006, from http://servicelearning.org/resources/funding_sources/index.php#category294

- *Eyler, J. S. (1997). Service-learning and the development of expert citizens. *Annual Meeting of the American Educational Research Association*, Chicago, IL.
- *Eyler, J. S., Giles, D. E., Jr., & Braxton, J. S. (1997). The impact of service-learning on college students. *Michigan Journal of Community Service Learning*, 4, 5-15.
- *Eyler, J., Giles, D. E., Lynch, C., & Gray, C. (1997). Service-learning and the development of reflective judgment.
- *Fenzel, M. L., & Leary, T. P. (1997). Evaluating outcomes of service-learning courses at a parochial college. *Annual Meeting of the American Educational Research Association*, Chicago, IL.
- *Forte, J. A. (1997). *Calling students to serve the homeless: A project to promote altruism and community service*
- *Fox, R. A., & Lopuch, W. R. (1984). Using volunteer work to teach undergraduates about exceptional children. *Teaching of Psychology*, 11(2), 113-115.
- *Fredricksen, P. J. (2000). Does service learning make a difference in student performance? *Journal of Experiential Education*, 23(2), 64-74.
- *Freyss, S. F. (2003). Learning political engagement from the experts: Advocacy groups, neighborhood councils, and constituency service. *99th Annual Meeting of the American Political Science Association*,

- *Gadbury-Amyot, C. C., Simmer-Beck, M., McCunniff, M., & Williams, K. B. (2006). Using a multifaceted approach including community-based service-learning to enrich formal ethics instruction in a dental school setting. *Journal of Dental Education*, 70(6), 652-661.
- *Giles, D. E., & Eyler, J. (1994). The impact of a college community service laboratory on students' personal, social, and cognitive outcomes. *Journal of Adolescence*, 17(4), 327-339.
- Giles, D., Honnet, E. P., & Migliore, S. (Eds.). (1991). *Research agenda for combining service and learning in the 1990s*. Raleigh, NC: National Society for Internships and Experiential Education.
- *Gorman, M., Duffy, J. (1994). Service experience and the moral development of college students. *Religious Education*, 89(3), 422-431.
- *Haines, D. L. (2002). A study of community college student attitudes related to service learning. (Ed.D., Baylor University).
- *Hamilton, S. F., & Zeldin, S. R. (1987). ^{Learning civics in the community}. *Curriculum Inquiry*, 17(4), 407-420.
- *Hecht, D., & Fusco, D. R. (1996). *Student perceptions of the opportunities available from participating in service learning*. U.S.; New York:

*Johnson, A. M., & Notah, D. J. (1999). Service learning: History, literature review, and a pilot study of eighth graders. *Elementary School Journal*, 99(5), 453-467.

*Johnson, S. D., & Bozeman, M. Service learning and the development of social responsibility. *Paper presented at the Annual Convention of the Central States Communication Association (66th, Chicago, IL, April 2-5, 1998).*,

*Kim, Y., Clasen, C., & Canfield, A. (2003). Effectiveness of service learning and learning through service in dietetics education. *Journal of Allied Health*, 32(4), 275-278.

*Knapp, J. L., & Stubblefield, P. (2000). Changing students' perceptions of aging: The impact of an intergenerational service learning course. *Educational Gerontology*, 26(7), 611-621.

*Knutson Miller, K., & Yen, S. C. (2004). Group differences in academic achievement: Service learning in a child psychology course. *Teaching of Psychology*, 32(1), 56.

Learn and Serve America. (2006a).*Homepage*. Retrieved June 16, 2006, from <http://www.learnandserve.org/>

Learn and Serve America. (2006b). *Tools, Training and Information*. Retrieved June 16, 2006, from http://www.learnandserve.org/for_organizations/ta/index.asp

Leming, J. S. (2001). Integrating a structured ethical reflection curriculum into high school community service experiences: Impact on students' sociomoral development. *Adolescence*, 36(141), 33-45.

*Lies, J. M. (2005). *Effects of service learning on college students' moral reasoning*.

Unpublished Ph.D, University of Minnesota.

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks, CA,

US: Sage Publications, Inc.

*Markus, G. B., Howard, J. P., & King, D. C. (1993). Integrating community service and classroom instruction enhances learning: Results from an experiment. *Educational Evaluation and Policy Analysis, 15*(4), 410-419.

*Marullo, S. (1998). Bringing home diversity: A service-learning approach to teaching race and ethnic relations. *Teaching Sociology, 26*(4), 259-275.

*McKinney, K. G. (2002). Engagement in community service among college students: Is it affected by significant attachment relationships? *Journal of Adolescence, 25*(2), 139-154.

*Miller, J. (1994). Linking traditional and service-learning courses: Outcome evaluations utilizing two pedagogically distinct models. *Michigan Journal of Community Service Learning, 1*(1), 29-36.

The National Service-Learning Clearinghouse. (2007). "Students in service to America."

Retrieved March 17, 2007, from:

<http://www.studentsinservicetoamerica.org/guidebook/classroom.html>

National Service Learning Cooperative (1998). *Essential elements of service learning*. St. Paul, MN: National Youth Leadership Council.

Newmann, F. M., & Rutter, R. A. (1983). The effects of high school community service programs on students' social development. final report. Madison, WI: Wisconsin Center for Education Research.

*Nokes, K. M., Nickitas, D. M., Keida, R., & Neville, S. (2005). Does service-learning increase cultural competency, critical thinking, and civic engagement? *The Journal of nursing education*, 44(2), 65-70.

*O'Donnell, L., Stueve, A., Doval, A. S., Duran, R., Haber, D., & Atnafou, R., et al. (1999). The effectiveness of the reach for health community youth service learning program in reducing early and unprotected sex among urban middle school students. *American Journal of Public Health*, 89(2), 176-181.

*Olm-Shipman, C., Reed, V., & Jernstedt, G. C. (2003). Teaching children about health, part II: The effect of an academic-community partnership on medical students' communication skills. *Education for Health: Change in Learning & Practice*, 16(3), 339-347.

*Osborne, R. E., Hammerich, S., & Hensley, C. (1998). Student effects of service-learning: Tracking change across a semester. *Michigan Journal of Community Service Learning*, 5, 5-13.

- *Ostheim, P. Q., & Jewell, L. R. (1995). Effects of community service and service learning on multidimensional self concept of secondary school students. In V. O. Arnold (Ed.), *North carolina council of vocational teacher educators. proceedings of the annual summer workshop (10th, raleigh, north carolina, august 1995)* (). U.S.; North Carolina: ERIC.
- *Parker-Gwin, R., & Mabry, J. B. (1998). Service learning as pedagogy and civic education: Comparing outcomes for three models. *Teaching Sociology*, 26(4), 276-291.
- *Perry, C. M. (1998). Community service learning: Goals and outcomes. *Annual Meeting of the National Rural Education Association*, Buffalo, NY.
- *Piper, B., DeYoung, M., & Lamsam, G. D. (2000). Student perceptions of a service-learning experience. *American Journal of Pharmaceutical Education*, 64(2), 159-165.
- *Ramia, P. N. (2005). *Developing altruism and empathy in ecuadorian college students: Impact of a mandatory service-learning course*. Unpublished Ed.D, Boston University.
- *Reed, V. A., Jernstedt, G. C., Hawley, J. K., Reber, E. S., & DuBois, C. A. (2005). Effects of a small-scale, very short-term service-learning experience on college students. *Journal of Adolescence*, 28(3), 359-368.

- *Reinders, H., & Youniss, J. (2006). School-based required community service and civic development in adolescents. *Applied Developmental Science, 10*, 2-12.
- *Rossi, B. R. (2002). *Impacts and effects of service learning on high school students*. Unpublished Ph.D, University of Minnesota.
- *Scales, P. C., Blyth, D. A., Berkas, T. H., & Kielsmeier, J. C. (2000). The effects of service-learning on middle school students' social responsibility and academic success. *Journal of Early Adolescence, 20*, 332-358.
- *Shastri, A. (2001). Examining the impact of service-learning among preservice teachers. *53rd Annual Meeting and Exhibits of the American Association of Colleges for Teacher Education*, Dallas, TX.
- *Sherman, A. R. (1982). Psychology fieldwork: A catalyst for advancing knowledge and academic skills. *Teaching of Psychology, 9*(2), 82-85.
- *Singer, J. A., King, L. A., Green, M. C., & Barr, S. C. (2002). Personal identity and civic responsibility: "rising to the occasion" narratives and generativity in community action student interns. *Journal of Social Issues, 58*(3), 535-556.
- Skinner, R., & Chapman, C. (1999). *Service-Learning and Community Service in K-12 Public Schools*. National Center for Education Statistics. Retrieved June 16, 2006 from <http://nces.ed.gov/surveys/frss/publications/1999043/index.asp>

*Sperling, R., Wang, V. O., Kelly, J. M., & Hritsuk, B. (2003). Does one size fit all?: The challenge of social cognitive development.. *Michigan Journal of Community Service Learning, 9*(2), 5-14.

*Stafford, J. R., Boyd, B. L., & Lindner, J. R. (2003). The effects of service learning on leadership life skills of 4-H members. *Journal of Agricultural Education, 44*(1), 10-21.

*Strage, A. (2004). Long-term academic benefits of service-learning: When and where do they manifest themselves? *College Student Journal, 38*(2), 257-261.

*Sugar, J., & Livosky, M. (1998). Enriching child psychology course with preschool journal option. *Teaching of Psychology, 15*(2), 93-95.

*Switzer, C. L. (1999). *Service learning in a medical school: Psychosocial and attitudinal outcomes*. Unpublished Ph.D, University of Pittsburgh.

U.S. News and World Reports. (2006). *Academic programs: Service learning*. Retrieved on June 16, 2006, from http://www.usnews.com/usnews/edu/college/rankings/brief/acadprogs05_brief.php

*Yamauchi, L. A., Billig, S. H., Meyer, S., & Hofshire, L. (2006). Student outcomes associated with service-learning in a culturally relevant high school. *Journal of Prevention & Intervention in the Community, 32*(1-2), 149-16

