Educational Psychology & Higher Education Faculty Publications

Educational Psychology & Higher Education

Spring 2008

The Impact of Service Learning on Moral Development and Moral Orientation

Matthew L. Bernacki PhD
University of Nevada, Las Vegas, matt.bernacki@unlv.edu

Elizabeth A. Jaeger
St. Josephs University

Follow this and additional works at: https://digitalscholarship.unlv.edu/edpsych_fac_articles

Part of the Educational Assessment, Evaluation, and Research Commons, Educational Methods Commons, and the Educational Psychology Commons

Repository Citation

Exploring the Impact of Service-Learning on Moral Development and Moral Orientation

Matthew L. Bernacki and **Elizabeth Jaeger**
Saint Joseph's University

Research on Service-learning's (SL) impact on students' moral development has been "mixed." In this study, 46 students in SL and non-SL sections of comparable courses offered at a northeastern Catholic university completed the Defining Issues Test, the Moral Justification Scale, and the SL Outcome Scale at the beginning and end of a semester. Although scores on moral development and orientation did not change significantly, SL students reported becoming more compassionate and more sensitive, having a greater understanding of and ability to solve social problems, and possessing a greater efficacy to make the world better. While a single-semester exposure to SL may be too limited to affect moral development, participants' self-reported changes may be precursors to such developmental changes. Unfortunately, existing measures of moral orientation may preclude a thorough examination of change associated with SL. Future research would benefit from using tools that measure moral thinking and action, and understanding of hypothetical moral principles.

When John Dewey discussed the linkage of education and experience in a book by the same name (1938), he simultaneously established experiential learning as pedagogy and embedded in it the central idea that education includes a civic and moral imperative. In doing so, he established a field that would later encompass our modern conception of service-learning (SL) (Giles, 1991). Although the field of SL has continued to evolve through the work of additional theorists and practitioners, some in the discipline argue that the moral imperatives first discussed by Dewey should not be lost (Hatcher, 1997). SL practitioners in the vein of Dewey who see moral development as an implicit goal of the pedagogy do not shy away from assessing the construct and believe that SL can play an integral role in character development in undergraduate students (Hatcher, 1997; Kohlberg, 1971; Kohlberg, Higgins & Power, 1989; Rest & Narvaez, 1991).

Similar to the way Dewey saw moral responsibility as implicit in experiential education, Kohlberg suggested that moral judgment is a developmental phenomenon fostered by individual experiences (Kohlberg, Higgins & Power, 1989). Further, Kohlberg (1971) defined community service as an important out-of-classroom element of moral education which forces students to confront moral issues. He believed that, when presented with a new way of thinking, people may be coaxed into reasoning at higher levels by reorganizing the way they think about a given situation. Rest and Narvaez (1991) have identified community service

as a means of promoting moral development, especially at the college level.

Despite the belief that SL implicitly teaches moral reasoning, researchers have had difficulty supporting this idea through research. In a compendium of outcomes attributed to SL, Eyler, Giles, Stenson, and Grey (2001) state that "the impact of SL (SL) on student cognitive moral development is mixed" (p. 5). This decision was rendered based on the evidence of four published studies. The purpose of this study was to further examine the SL context in which moral development has been evaluated as an outcome and to expand the definition of moral development to include the dimension of moral orientation.

Research on Moral Development as an SL Outcome

Table 1 describes the major design features and results of the published studies examining moral development as an outcome of SL. As can be seen in the table, no consistent pattern of results emerges.

Two studies lend some support to the theory that SL coursework impacts moral development (Boss, 1994; Gorman, Duffy, & Heffernan, 1994). Boss compared gains in students' moral development in two sections of an ethics course using the Defining Issues Test (DIT; Rest, 1986), a measure of moral development based on Kohlberg's theory. One of the two sections of the course was randomly assigned an additional service and reflection component and compared to the section which employed a non-SL format. Analyzing pretest-

Table 1
Characteristics of Studies Examining Impact of Service-Learning (SL) on Moral Development of College Students

Authors	Sample	Course/Discipline	SL Context (treatment strength)	Measurement tool & Analysis	Results
Boss (1994)	71 (34 male, 37 female); Mean age= 20.3	Ethics, including moral development & orientation	20 hrs service; unknown hrs reflection; reflection at individual & group level;	Defining Issues Test (DIT) (P score); t-test comparing (T2-T1 P scores)	SL students display sig. greater gains in P scores than non-SL students
Fenzel & Leary (1997)	57 (16 male, 41 female); Mean age = unknown; were college sophomores	Philosophy	15 hour service; 10 hour reflection; individual level only	DIT (P, D score); 2x2 Group x Time ANOVA	No significant differences by group or over time for either score
Gorman, Duffy & Heffernan (1994)	70 (unknown breakdown by gender); mean age = 18.23 years	Theology & Philosophy; SL course focused on moral questions	12 credit, 2 semester course; 10-12 hours service per week for 2 semesters; reflection at individual and small group level; (no total service or reflection hours stated)	DIT (P, D scores); t-test comparing (T2-T1 P scores); T1 was middle of first semester	SL group P scores increased significantly; P scores of the non-SL group did not increase significantly
Greene (1997)	52 (6 male, 46 female); Mean age = 25.43	Occupational Therapy	6 hours of service; 2 hours reflection; reflection at individual level only	Sociomoral Reflection Measure- Short Form; 2x2 Group x Time ANOVA	No significant differences by group or over time

posttest score differences, Boss found that students in the SL group increased their use of principled moral reasoning more than the students in the non-SL group over the course of the semester. Gorman et al., also using the DIT, assessed moral reasoning change in students in a two-semester philosophy and theology course. Moral reasoning scores increased significantly in the SL group, but not in the non-SL group.

Two additional studies on the influence of SL on moral development found no significant differences in moral reasoning of SL students compared to non-SL students (Fenzel & Leary, 1997; Greene, 1997). Fenzel and Leary completed a study similar in method to Boss (1994). Philosophy students self-selected into two sections of a course. One section was later assigned as a SL course. Students were compared over the course of a semester using the DIT. In contrast to Boss' results, Fenzel and Leary did not find significant differences between the two groups in moral reasoning. Students, however, reported benefits of SL in interviews and journals. Finally, Greene found that moral reasoning, as assessed by the Sociomoral Reflection Measure—Short Form (Gibbs et al., 1984), actually declined over time within a group of SL students, even though the students showed an increase in their psychosocial maturity (defined as a greater appreciation for dignity, equality, and justice).

Variations in the design of these four studies make reconciliation of findings difficult. As can be seen in Table 1, such variations between studies include sample size, SL context, and the tool used to measure moral development.

The Boss (1994) and Gorman et al. (1994) studies employed larger samples than the studies which did not find SL students to significantly increase their moral reasoning ability over the semester. This suggests that finding significance could be a matter of power of the analysis. Variations also exist in measurement of moral development. Gains in moral development were significant in two of three studies which employed the Defining Issues Test. Gains were not significant (and negative) when measured by the Sociomoral Reflection Measure—Short Form. Finally, the strength of the treatment (SL context) also varied by study. The Gorman et al. study endured for two semesters and the service required can be extrapolated to approximately 300 hours. The Boss study included 20 hours of service. Both of these found significant impact of SL on moral development. The studies that found no significant impact of SL required only 15 hours (Fenzel & Leary, 1997) and 6 hours (Greene, 1997) of service. This suggests that the intensity of the SL context had an impact on stu-

dents' moral development (Eyler & Giles, 1997).

The Boss (1994) and Gorman et al. (1994) studies also included course content that specifically focused on moral issues. Though each researcher argues that the content of the course implies a moral development objective (as does any inclusion of service according to Dewey, 1938; Hatcher, 1997; Kohlberg, 1971), Boss explicitly stated that "[d]iscussion of moral dilemmas and moral development were part of the curriculum in both classes" and that "[t]he stages of moral development were included in the reading with special attention being given to the theories of Gilligan and Kohlberg" (p. 186). This inclusion, though it occurred in both the experimental and control groups, also suggests that the objective of moral growth was more germane to this course than the others that measured SL and moral development.

In Gorman et al. (1994), the SL course is described as including "explor[ation] of basic moral questions and their relation to the student, to society and in particular to the student's own project" (p. 425-426). It can be argued that such content is an artifact of the SL context of the course, but the difference in titles of the examined courses support the idea that these courses contained different content. They were entitled *Personal & Social Responsibility* (SL) and *Perspectives on Western Culture* (non-SL). Neither Fenzel & Leary (1997) nor Greene (1997) made mention of an inclusion of specific discussion of moral issues in the courses studied. It is possible that significant impact of SL on moral development only may be seen in courses that include discussion of moral issues and a service component. Differences may be caused by an interaction between such discussion and service immersion in SL sections of courses.

In each study thus far, the respective researcher(s) chose to measure moral development outcomes using an instrument that reports gains in moral reasoning through stage scores (Sociomoral Reflection Objective Measure) or percentages of time a respondent uses principled moral reasoning (P Score of the Defining Issues Test). While these tools attempt to measure the hierarchically different reasoning strategies people employ when making decisions, they do not take into account the breadth of factors a person may consider when employing a moral reasoning strategy. To consider the ethics a person employs in a given moral scenario, a tool should also account for the moral orientation of the reasoner.

According to Kohlberg's (1971) theory, a person makes decisions using either a moral orientation that focuses on issues of justice or care in a given scenario. Gilligan (1982) added to this theory by

suggesting gender as a moderator in the use of one ethic. According to Gilligan, the moral reasoning of men is more justice-oriented based on their greater detachment and preference for objectivity, while the moral reasoning of women is more care- and relationship-oriented due to greater sensitivity and perceived interdependence with others.

Traditional curricula in the social sciences and humanities tend to employ logic and abstract thinking as a means of evaluating course materials. Courses such as philosophy and sociology focus on issues of justice and equality, but attempt to do so in a theoretical environment. If one were to make an assumption about how curriculum would influence the adoption of an ethic of care or of justice, one might anticipate that objective approaches to course content would motivate students to adopt a justice orientation. In contrast, courses that draw upon SL experiences necessarily involve a humanizing experience involving individuals with whom students build relationships. Students' personal connection to individuals who face real moral dilemmas should then spur students to consider issues of emotional content and require a focus on response to individuals' needs. Because of this additional community experience, it could be hypothesized that SL students should be more likely to adopt an ethic of care than students in non-SL courses. It is curious that Boss (1994), who includes Carol Gilligan's theory in her course readings, did not assess changes in moral orientation of her students.

The Present Study

The first purpose of this study was to further investigate whether SL courses can promote increases in moral development above and beyond comparable courses with no service component. The definition of the SL context in this study included a combination of direct service (30 hours) and reflection conducted through journaling, individual reflection, and student-led group discussions in and out of class in the presence and absence of an instructor. This study also expands on previous research by incorporating a more representative selection of courses from multiple disciplines (English, philosophy, and sociology) to examine the impact of SL on moral development across disciplines. This study also examined whether SL courses would promote more frequent use of an ethic of care than would analogous non-SL courses. It is theorized that an ethic of justice would be employed at similar levels across groups.

Finally, this study also examined whether SL students would report more positive changes in themselves as a result of their experience, as has been found previously (Astin, Vogelgesang, Ikeda, &

Yee, 2000; Boss, 1994; Rowe & Chapman, 1999). Specifically, it was hypothesized that SL students would report a larger impact of their coursework on their understanding of social problems, their ability to act as problem solvers, their efficacy to make the world a better place, their ability as future leaders, their sensitivity and compassion than students in non-SL courses. Such self-reported changes are important to investigate as they may represent precursors to increased levels of prosocial variables like moral reasoning and orientation.

The National Survey on SL and Transitioning to Adulthood (Harris Interactive, 2006) reported that K-12 SL students felt more positive than their non-SL peers about their lives and their academic abilities. They later develop into leaders dedicated to community service and become active citizens during college and beyond. Given the trajectory of self-reported SL outcomes leading to future civic engagement outcomes, a similar trajectory can be hypothesized where SL outcomes such as an increased understanding of social problems, greater problem solving ability, leadership skills, and sensitivity can precede future maturation in moral reasoning and moral orientation. A similar trajectory also has been suggested by Greene (1997).

Method

Participants

Forty-six students at a Catholic university in the northeastern United States who enrolled in selected SL courses and analogous courses were involved in the study. These students self-selected into courses labeled as either having a SL component or into traditional courses without such a requirement or label. Twenty-five SL students and 21 non-SL students participated in both Time 1 and Time 2 data collections. This represents a participation rate of 38% of the total number of students enrolled in the six sampled courses. Students included in the sample were enrolled in freshman English, philosophy, or sociology classes. See Table 2 for a summary of students by class and gender.

These students included 16 males and 30 females. The sample was 82.6% White, 8.7% Latino, 2.2% African American, 2.2% Asian American, and 4.3% of mixed descent. The average age was 18.49 years ($SD = 0.38$). Students, on average, completed 98.56 hours ($SD = 131.24$) of community service since the beginning of high school.

Students in SL sections spent at least three hours outside the classroom per week (30 hours per semester) immersed at a community service site where they were exposed to issues complementing those they were learning about in the classroom. For instance,

students in a philosophy class discussing human dignity and personhood were given the opportunity to experience homelessness issues through a service project involving companionship, outreach, and advocacy with homeless men. Students in English classes tutored English language learners. Students in sociology classes provided companionship to the homeless and the elderly while others provided tutoring and learning support for inner-city elementary and middle school students.

Students also were required to keep a journal of their experiences and spent time inside the classroom, on discussion boards, and at scheduled reflection dinners discussing the content of their service commitment.

The non-SL courses contained similar curricula to the SL courses, with the exception of the service immersion and reflection components. As evidenced by the required readings and the assignments in course syllabi, students in SL and non-SL courses completed identical reading and the same number of writing assignments. Objectives of writing assignments varied only to include reflection on the service component in the SL courses. The English courses were taught by the same instructor; the sociology and philosophy courses were not. Students completing both pre-test and post-test packets were given the option of receiving curricular enhancement requirement credits as required by their academic department, and were entered into a drawing for a $50 gift certificate. An honorarium was paid to all participants after the second packet was collected.

The objectives for SL classes, as stated on the three course syllabi, included reflection on their work at the service site and how their understanding of service developed over the semester (English). Courses also aimed to develop in students an understanding of connections between individual experiences of social inequality at weekly service and larger structural issues of inequity (sociology). They challenged students to understand some of the central philosophical questions about being a human being and to put these universal and philosophical questions in a larger context than one's own experience (philosophy).

Procedure

Data were collected twice. The first data collection occurred in the first two weeks of the semester. Participants were asked to participate in the study by a researcher who visited their classroom. Participants were given a packet of questionnaires, including the DIT, the Moral Justification Scale, and a survey of personal characteristics, and briefed on the measures inside. Packets were completed outside of the classroom. Participants were instructed that the packet should take approximately one hour to complete, and to return them to the experimenter outside the classroom on determined dates. The second data collection occurred during the last two weeks of the semester. Participating students were given a second packet and were instructed to return them to the experimenter by the end of the semester. This packet contained copies of the DIT, the Moral Justification Scale, the SL Outcomes Scale (Rowe & Chapman, 1999), and a second survey assessing personal characteristics.

Measures

The Defining Issues Test (DIT), developed by Rest and Thoma (1979), gauges a person's level of moral reasoning according to the Kohlberg model of moral development. The test includes six dilemmas assessed by making judgments on 12 statements regarding importance to the dilemmas. Participants receive P scores on the DIT, which gauge their stage of moral development. The P score is a value representing the percentage of time participants use principled moral reasoning in assessing moral dilemmas. A P score can range from 0% to 96% by increments of 1.67% when using the 6-story DIT. The 6-story DIT has a test-retest reliability of .77 and a Cronbach's alpha of .77 (Rest, 1986).

The Moral Justification Scale (MJS; Gump, 1994) is a 6-story, 72-item test used to assess moral orientation. It includes six dilemmas similar to those in the DIT. However, instead of ranking item importance, the MJS offers students a 10-point Likert scale for 48 items taken directly from the

Table 2

Distribution of Students in SL and Non-SL English, Philosophy, and Sociology Classes.

	SL			non-SL			total
	Male	**Female**	**Total**	**Male**	**Female**	**Total**	
English	3	6	9	5	3	8	17
Philosophy	2	6	8	1	3	4	12
Sociology	1	7	8	4	5	9	17
Total	6	19	25	10	11	21	46

text of dilemmas they must rate from "not important" to "very important." The content of the answer selections on this test are assigned point values based on their implementation of different ethics. The measure consists of two subscales—the Care subscale and Justice subscale. Cronbach's alpha is .75 for the Care subscale and .64 for the Justice subscale. Test-retest reliability is reported to be 0.61 for the care subscale and 0.69 for justice (Gump; Gump, Baker, & Roll, 2000).

The third measure administered to participants (at Time 2 only) was the SL Outcomes Scale (Rowe & Chapman, 1999). This self-report questionnaire collected data regarding students' perceptions of how their coursework impacted the richness of their educational experiences, understanding of social problems, problem solving ability, ability as future leaders, efficacy to make the world a better place, compassion, and sensitivity. Items included a statement such as "This class has made me more compassionate." Students were asked to rate their agreement with the statement on a 5-point Likert scale (1 = "strongly disagree" and 5 = "strongly agree"). Psychometrics for the SL Outcome Scale were not reported by the original authors.

Finally, a survey of personal characteristics developed for this study was administered at Time 1 and at Time 2. At Time 1, participants were asked their gender, age, and number of semesters enrolled in colleges and universities. Participants also were asked about their previous SL classes. Participants were asked to list previous service experiences, as well as the site, type, and duration (in hours). A second background measure was included at Time 2; this included items asking students to report their in-class and out-of-class service involvement during the semester in terms of hours, location, and type of service.

Results

To examine for a potential self-selection bias, groups were compared at Time 1 to ensure that groups were not significantly different in terms of age, gender, ethnicity, and previous service experience. There were no significant differences between groups on these variables. However, a trend was found for gender [x^2 (1, 46) = 2.81, p = .094]. The SL group tended to be composed of more females (n = 19) than males (n = 6) whereas the non-SL group had a more equal distribution of females (n = 11) and males (n = 10). Because of this trend, hypotheses regarding moral reasoning and orientation were tested using a 2 x 2 mixed ANCOVA (Group x Time) with gender as the covariate. SL outcomes were tested using a 2 x 2 ANOVA (Group x Gender).

Moral Reasoning Scores

No significant main effects of Group [F(1, 42) = .639 , p =.429] or Time [F(1, 42) = .877, p =.354] were found for moral reasoning scores. No interaction effect of Group x Time was found [F(1, 42) = .066 , p =.788]. See Table 3 for adjusted means for P scores by group and time.

Table 3
Adjusted Principled Reasoning (P) Scores of Students in SL and Non-SL Classes.

	SL	non-SL
N=	25	20
	M (SD)	M (SD)
P score (time 1)	32.13 (10.73)	30.75 (10.07)
P score (time 2)	36.27 (12.08)	32.60 (11.68)

Moral Orientation

For Care scores, no significant main effects of Group [F(1, 40) = .091 , p = .765] or of Time [F(1, 40) = 1.37, p = .249] were noted. No interaction effect of Group x Time was found [F(1, 40) = .138, p = .712]. See Table 4 for adjusted means for Care scores by group and time.

Table 4
Ethic of Care Scores of Students in SL and Non-SL Classes.

	SL	non-SL
N=	24	19
	M (SD)	M (SD)
Ethic of Care (time 1)	7.43 (1.04)	7.43 (0.96)
Ethic of Care (time 2)	7.48 (0.97)	7.46 (0.92)

For Justice scores, no significant main effects of Group [F(1, 39) = .103, p = .750] or of Time [F(1, 40) = 0.72, p = .402] were noted. No interaction effect of Group x Time was found [F(1, 39) = .068, p = .795]. See Table 5 for adjusted means for Justice scores by group and time.

Table 5
Ethic of Justice Scores of Students in SL and Non-SL Classes.

	SL	non-SL
N=	24	18
	M (SD)	M (SD)
Ethic of Justice (time 1)	7.18 (1.37)	6.97 (1.19)
Ethic of Justice (time 2)	7.22 (1.01)	6.99 (0.82)

SL Outcomes Scale

Table 6 presents descriptive statistics for self-reported changes by Group and Gender. Students in SL classes reported a significantly better ability to understand social problems [$F(1, 35) = 6.66$, $p = .014$], and a significantly greater efficacy to make the world a better place [$F(1, 35) = 7.89$, $p = .008$] than did non-SL students. SL students also perceived that their coursework had given them a greater ability to be compassionate than did non-SL students [$F(1, 35) = 4.94$, $p = .033$]. SL students also tended to perceive that their coursework had made them more sensitive than did students in non-SL classes [$F(1, 34) = 3.14$, $p = .085$]. No significant differences were reported by students regarding their courses' impact on the richness of their educational experience, their ability to solve problems, or their ability to become future leaders. Also, no significant main effects of Gender or Group x Gender interactions were found for any variable.

A series of Group x Time x Class ANOVAs also were conducted to detect any main effects of course discipline (English, philosophy or sociology) on moral reasoning, justice orientation, or care orientation. No main effects or interaction effects were significant. One three-way interaction of Group x Time x Class indicated a non-significant trend in which philosophy students' increase the amount of time they reason at principled levels

more than English students (by 8%) [$F(2, 39) = 2.42$ $p = 1.02$]. This interaction may be spurious, based on the small cell sizes listed in Table 2 where no cells contain a population greater than 10 students. Accordingly, a larger sample would need to be gathered and data collected before speculating about these relationships.

Discussion

The purpose of this study was to examine the impact of SL on moral reasoning, moral orientation, and self-perceptions. The results of this study do not support the claim by Boss (1994) and Gorman et al. (1994) that taking SL courses promotes the moral reasoning of students more than analogous courses with no service component. In addition and contrary to our hypothesis, students who completed SL courses also were not found to use an ethic of care to resolve moral dilemmas more frequently than students in comparable non-SL courses. However, students who completed SL courses did perceive more positive changes in themselves than did students in traditional courses.

In contrast to a similar study completed by Boss (1994), but consistent with the findings of another similar study by Fenzel and Leary (1997), this study found no significant differences in moral reasoning P scores on the DIT between students taking SL and traditional courses. Before rejecting the

Table 6
Self-Reported Outcomes of Students in SL and Non-SL Classes.

	SL			non-SL		
	M	**F**	**All**	**M**	**F**	**All**
N=	6	19	25	9	11	20
Better Understanding	4.00	4.53	4.39	3.43	3.56	3.50[a]
of Social Problems	(0.63)	(0.62)	(0.66)	(0.79)	(1.33)	(1.10)
Provide Rich Educational	4.00	4.47	4.35	3.86	4.00	3.94
Experiences	(0.89)	(0.51)	(0.65)	(0.90)	(1.00)	(0.93)
Efficacy to Make World	3.67	4.47	4.26	3.86	4.00	3.25[b]
A Better Place	(1.21)	(0.62)	(0.86)	(0.76)	(0.97)	(0.86)
Better Problem Solvers	3.50	4.06	3.91	3.43	3.22	3.38
	(1.05)	(0.43)	(0.67)	(0.79)	(1.12)	(0.96)
Increase Ability as	3.50	3.94	3.83	3.43	3.22	3.31
Future Leaders	(1.05)	(0.83)	(0.89)	(0.53)	(1.20)	(0.95)
More compassionate	4.17	4.29	4.26	3.57	3.67	3.63[c]
	(0.75)	(0.77)	(0.75)	(0.79)	(0.87)	(0.81)
More sensitive	4.00	4.18	4.13	3.67	3.33	3.47
	(0.63)	(0.95)	(0.87)	(0.82)	(1.12)	(0.99)

Note. [a] Main effect of Group on 2x2 ANOVA (Group x Gender) [$F(1,28) = 6.66$, $p < .05$]
    [b] Main effect of Group on 2x2 ANOVA (Group x Gender) [$F(1,28) = 7.89$, $p < .05$]
    [c] Main effect of Group on 2x2 ANOVA (Group x Gender) [$F(1,28) = 4.94$, $p < .05$]

possibility that a single semester of SL courses could influence moral reasoning, one must consider that methodological differences might account for the different outcomes found between this and Boss' (1994) study. The differences include sample size and participant recruitment, age of the participants, and the SL context assessed in each study.

The samples used by Boss (1994) may have made it more likely that she found effects of SL on moral reasoning. Boss' study used a larger sample size (71 versus 46 participants), allowing for a greater likelihood of finding a significant impact of the SL courses. These differences, however, are unlikely to fully account for the difference in results. Based on an estimate of effect size from Boss' study (0.61), the sample in this study had sufficient power to detect such an effect.

Boss (1994) assigned students to SL and non-SL sections of the same course randomly, increasing the internal validity of her study. Although all selection effects may not have been controlled, she was able to control for the student's motivation to enroll in a SL course. In this study, students self-selected into SL and non-SL courses during registration. A potential difference in motivation to do service was assessed at baseline where no significant difference was found with respect to hours of service completed prior to course enrollment.

Age is an important variable to consider when assessing moral development outcomes (Rest, 1986) and it varied between studies. Boss (1994) used an older sample of college students (mean age = 20.3 years) than used in the present study (mean age = 18.49 years). These age differences were reflected in baseline moral reasoning scores. It might be that the age group used in this study (first semester freshmen) was cognitively different from the age group used in Boss' research. In Boss' study, the average pre-test DIT score across both classes was 40.0 (compared to 31.52 in this study) and the mean pre-test DIT score for the SL group was two points higher than the control group (as was the case in this study; both are non-significant between group differences). It is possible that the older students in Boss' sample were at a developmental stage where they were more ready to reorganize their thinking and elevate their level of moral reasoning to a state of principled thinking than were the incoming freshmen used in this present study. Thus, the timing of SL courses in a student's college career may affect the type of outcomes that can be expected from such experiences.

Boss (1994) served as the instructor for both sections of the course, assuring that the major difference in sections was one attributable to the SL component. However, since Boss served as the instructor for both sections of the courses, it is possible that experimenter effects may have contributed to the results. Finally, Boss (p. 186) included in her curriculum lectures regarding moral development and discussions concerning moral dilemmas, with a special focus on Kohlberg's theories of moral development and Gilligan's theories of sex differences in moral orientations and reasoning styles. Such factors may have accentuated differences in learning between groups and made it more likely for her to find positive effects of the SL courses.

Contrary to our hypotheses, SL students did not increase in their use of an ethic of care orientation over the course of the semester. Scores on the Moral Justification scale remained relatively high (7 out of 10) for students across both SL and non-SL groups at both time points. Thus, students felt that Care issues were important in each moral dilemma even before the semester began, and regardless of SL or non-SL course enrollment, perhaps leaving little room for change over time or differences between groups. It may be that SL courses do not engender an increase in the use of an ethic of care as hypothesized.

Although we continue to believe that examining changes in moral orientation as an outcome of SL is warranted, such study is likely to be hampered by a number of methodological issues. First, it is unclear whether it is more appropriate to measure moral orientation as a single construct in which ethics of care and justice represent mutually exclusive orientations or as two parallel constructs that can be measured along two separate continua. The MJS treats an ethic of care and an ethic of justice as two separate variables and measures each with separate items following each dilemma. In this study, however, these dimensions were highly correlated (r = .87).

The substantial correlation between Justice and Care scores argues against measurement of these variables as orthogonal constructs and is counterintuitive to the theories of Gilligan (1982). Gilligan states that there are two distinct moral orientations and that a person tends to use one orientation across the majority of situations. Walker, DeVries, and Trevethian (1987) refuted this finding, reporting that the tendency of people is to not use one orientation consistently. They also reject Gilligan's claim that women tend to use an ethic of care more often, having found no significant gender differences for usage of either ethic. Before moral orientation can be adequately explored as a domain for developmental change, these constructs need to be refined and a more nuanced measure created on the basis of such theory where Justice and Care scores correlate at a lower rate. While measures of inter-

nal consistency are impressive, a content validation of the MJS (Gump, 1994; p. 99) did not include a factor analysis. Instead, a panel of eight judges was asked to determine whether items were indicative of a justice orientation or a care orientation.

Despite the lack of significant findings on the effects of SL on moral development and moral orientation, students who took SL courses perceived themselves to be more socially conscious, more effective in changing the world, and more compassionate because of their coursework than did students in non-SL courses. Greene (1997) argues that these types of psychosocial outcomes are building blocks that will allow for development to higher levels of moral reasoning. The perceived significant increase in these constructs may represent an attitude change that suggests a restructuring of thinking, even though it may not translate to gains in moral reasoning. These findings support the employment of SL pedagogy.

Perhaps researchers should consider that en route to higher levels of moral reasoning, students first must develop individual constructs upon which principled reasoning is based. If this is the case, the benefit of SL courses might not be captured by the current theoretical models, which gauge level of moral reasoning and moral orientation. These measures may not be sensitive enough in their design to account for changes in individual constructs such as appreciation of human dignity, empathy, and a concern for social justice issues. Where Greene (1997) did not find significant gains in ethical reasoning in his study of the effects of SL courses, he did find that students developed a more mature understanding of dignity, justice, and equality, constructs that are elements of post-conventional moral reasoning (Kohlberg, Higgins & Powers, 1989). Future research into moral development might consider that a continual development of these psychosocial constructs may be a cognitive stepping-stone to developing the ability to reason at principled levels and should incorporate measures sufficiently sensitive to gauge these changes. We must also underscore the importance of conducting qualitative research (such as thematic analysis of reflections) to assess subtle changes that may by overlooked by quantitative measures.

Limitations of the Study

The quality of the SL experience is dependent upon both service experience duration and reflection experience intensity (Eyler & Giles, 1997). Thirty hours of service may have been insufficient for the hypothesized changes in moral development and orientation to occur. Eyler and Giles suggest that "lack of findings in the SL literature may be due to relatively short and low intensity experiences that have been studied" (p. 68) and suggest that a SL experience must endure at least one year to have significant impact. When discussing SL context, it also is relevant that this study was conducted at a smaller, private, Catholic university. It can be argued that students at such a school may differ in moral character from students at institutions without a religious affiliation.

Conclusion and Future Directions

In this study of the impact of SL on moral reasoning and moral orientation, no significant differences were found with respect to students' ability to reason at principled levels or with a greater orientation toward issues of justice or care as a result of one semester of SL participation. Although scores on moral development and orientation did not change significantly, students taking SL courses reported becoming more compassionate and more sensitive, having a greater understanding of and ability to solve social problems, and possessing a greater efficacy to make the world better when compared to non-SL peers. A single-semester exposure to SL and/or SL in the context of courses without an explicit focus on moral development may be too limited to affect moral development. However, the changes participants reported in themselves might be precursors to such developmental changes.

It is clear that some changes do take place in even one semester of SL enrollment, though they were not captured by the moral development measures used in this study. Some might argue that the choice of moral reasoning and moral orientation are not good outcomes by which to assess the efficacy of SL courses. Most SL courses are not explicitly aimed at increasing students' moral reasoning maturity about hypothetical dilemmas; they focus instead on moral practice and aim to influence the way students eventually behave in the world (Dewey, 1938; National SL Clearinghouse, 2008; Westheimer & Kahn, 2004). This raises questions about how we should assess moral development as a part of SL pedagogy. Is moral development germane enough to SL pedagogy to be assessed as a primary outcome? If so, should it be assessed using hypothetical dilemmas or through student reflections and their activities at the service site? What would an appropriate and pragmatic study look like that indicates students have developed morally as a result of their SL experience?

With the exception of some moral philosophy courses, few SL courses spend time addressing Kohlberg, Gilligan, and other content that would change principled moral reasoning and orientation

scores the way Boss (1994) did in her classes. Given the lack of consistent findings using the Defining Issues Test (Boss; Fenzel & Leary, 1997; Gorman, et al. 1994; and this study), it might be beneficial for researchers who aim to assess moral development to incorporate a variety of tools that measure moral thinking and action, in addition to measuring understanding of hypothetical moral principles.

When attempting to tell the story of how students change as a result of SL, measures gauging reasoning ability and depth of orientation are necessary but not sufficient. An improved study of the impact of SL on moral development might include a mixed method design. Researchers should continue to assess hypothetical moral development using the DIT and measures of moral orientation as it is important to track changes in students' moral reasoning ability and orientation. In addition, it would be helpful to include a qualitative element, potentially as part of the course reflection, where researchers can observe *how* reasoning changes over time, and what critical events spur reasoning to become principled, justice-oriented, or care-oriented. Finally, an additional instrument measuring instances of moral practices would confirm that increases in moral reasoning abilities lead to greater frequency of actions guided by principled morality, authenticating the pedagogies of Dewey (1938). While the development of this type of longitudinal inventory may prove difficult, such a tool would enhance the discussion of moral development as an outcome of SL by serving as a concrete measure of service grounded in principled morality linking developmental changes in reasoning with a greater frequency of morally-driven actions.

References

Astin, A., Vogelgesang, L., Ikeda, E., & Yee, J. (2000). *How service-learning affects students.* Los Angeles: University of California, Higher Education Research Institute.

Boss, J.A. (1994). The effect of community service on the moral development of college ethics students. *Journal of Moral Education*, *23*, 183-198.

Dewey, J. (1938). *Experience and education.* New York: Collier Books.

Eyler, J. & Giles, D.E. (1997) The importance of program quality in service-learning. In A. S. Waterman (Ed.), *Service-Learning: Applications from the research* (pp.57-76). Mahwah, NJ: Lawrence Erlbaum Publishers.

Eyler, J. S. & Giles, D. E., Jr. (1999). *Where's the learning in service-learning?* San Francisco, CA: Jossey-Bass, Inc.

Eyler, J., Giles, D.E., Stenson, C.M. & Gray, C.J. (2001). *At a glance: What we know about the effects of service-learning on college students, faculty, institutions and communities, 1993-2000.* Third Edition. Washington, D.C.: Corporation for National Service.

Fenzel, L.M. & Leary, T.P. (1997) *Evaluating outcomes of service-learning courses at a Parochial College.* Paper presented at the Annual meeting of the American Educational Research Association, Chicago, IL.

Gibbs, J.C., Arnold, K.D., Morgan, R.L., Schwartz, E.S. Gavaghan, M.P. & Tappan, M.B. (1984). Construction and validation of a multiple choice measure of moral reasoning. *Child Development*, *55*, 527-536.

Giles, D. E. (1991). Dewey's theory of experience: Implications for service-learning. *Journal of Cooperative Education*, *27*(2), 87-90.

Giles, D. E., & Eyler, J. S. (1994). The impact of a college community service laboratory on students' personal, social and cognitive outcomes. *Journal of Adolescence*, *17*, 327-339.

Gilligan, C. (1982). *In a different voice: Psychological theory and women's development.* Cambridge, MA: Harvard University Press.

Gorman, M., Duffy, J. & Heffernan, H. (1994). Service experience and the moral development of college students. *Religious Education*, *89*(3), 422-31

Greene, D. (1997). The use of service-learning in client environments to enhance ethical reasoning in students. *American Journal of Occupational Therapy*, *51*(10), 844-852.

Gump, L. S. (1994). The relationship of culture and gender to moral decision-making. Unpublished doctoral dissertation, California School of Professional Psychology, San Diego, CA

Gump, L.S., Baker, R.C. & Roll, S. (2000), The Moral Justification Scale: Reliability and validity of a new measure of care and justice orientations. *Adolescence*, *35*, 67-76.

Harris Interactive. (2006). SL and transitioning to adulthood. New York: Harris Interactive. Retrieved January 4, 2008 from http://www.nylc.org/rc_downloaddetail.cfm?emoid=14:664&si=1

Hatcher, J.A. (1997). The moral dimensions of John Dewey's philosophy: Implications for undergraduate education. *Michigan Journal of Community Service Learning*, *4*, 22-29

Kohlberg, L. (1971). *Philosophy of moral education.* New York: Harper & Row.

Kohlberg, L., Higgins, A. & Powers, F.C. (1989). *Lawrence Kohlberg's approach to moral education.* New York: Columbia University Press.

Kolb, D. (1984) *Experiential learning: Experience as the source of learning and development.* Englewood Cliffs, NJ: Prentice Hall.

National Service Learning Clearinghouse (2008). *What is service-learning?* Retrieved January 4, 2008 from http://www.servicelearning.org/what_is_SL/SL_is/index.php

Rest, J.R. (1986) *Manual for the Defining Issues Test, 3rd Edition*. Minneapolis: Center for the Study of Ethical Development, University of Minnesota. 5.1-5.7.

Rest, J.R., Davidson, M.L. & Robbins, S. (1978). Age trends in judging moral issues: A review of cross-sectional, longitudinal, and sequential studies of the Defining Issues Test. *Child Development*, *49*,263-279.

Rest, J.R. & Narvaez, D. (1991). *The college experience and moral development. Handbook of moral behavior and development* (pp. 229-245). Hillsdale, NJ: Erlbaum..

Walker, L.J., de Vries, B., & Trevethian S.D. (1987). Moral stages and moral orientations in real-life and hypothetical dilemmas. *Child Development*, *58*, 842-858.

Waterman, A.S. (1997) The role of student characteristics in service-learning. In A.S. Waterman (Ed.), *Service-learning: Applications from the research* (pp. 95-105). Mahwah, NJ: Lawrence Erlbaum Publishers.

Westheimer, J., & Kahn, J. (2004). What kind of citizen: The politics of educating for democracy, *American Educational Research Journal*, *41*(2), 237-269.

Authors

MATTHEW BERNACKI is Research Coordinator at the Faith-Justice Institute at Saint Joseph's University and has studied SL throughout his academic career. He holds a Master's degree in experimental psychology with a focus in developmental psychology and a Master's degree in Social Work. He is currently in doctoral study in Educational Psychology at Temple University.

ELIZABETH JAEGER is an Associate Professor of Psychology at Saint Joseph's University. Dr. Jaeger's research concerns basic and applied issues as they relate to the impact of early care and education on children's development. From 1995-2001, she served as Co-Principal Investigator on the NICHD Study of Early Child Care and Youth Development. Dr. Jaeger has conducted a number of evaluations of interventions aimed to improve the quality of early care and education settings. Currently, she is Principal Investigator of an evaluation examining the impact of a facilities fund on program quality. In addition to research on early care and education, Jaeger's research focuses on the antecedents and consequences of parent-child attachment relationships across the life cycle.