Service-Learning and Student Diversity Outcomes: 
Existing Evidence and Directions for Future Research

Matthew A. Holsapple
University of Michigan

As today's business world and society become more diverse, it is essential for colleges and universities to prepare students to work and live in that diverse world, and service-learning is one tool for that education. This study presents a critical review of 55 studies of the impact of service-learning participation on students' diversity outcomes, identifying six diversity-related outcomes that emerge from these studies. The paper also identifies five major limitations of the existing body of research, and offers suggestions for researchers to conduct and write about this research in ways that provide an empirical basis for effective service-learning practice.

The increasing diversity of the American population is influencing the make-up of colleges and universities. The proportion of students of color among undergraduates rose from 16% in 1976 to 29% in 2000 (Rankin & Reason, 2005), and Carnevale and Fry (2000) estimated that percentage would grow to 37% by 2015, with states such as Hawaii, California, and New Mexico expected to see students of color outnumber White students. In several states, minority students already make up more than one-third of students (Pike & Kuh, 2006).

The worlds in which students grow up prior to entering college are not seeing a comparable diversification. Elementary and secondary schools are growing more segregated (Orfield, Bachmeier, James, & Eitle, 1997), and even in diverse schools, the Black and White students are almost completely socially self-segregated (Echenique & Fryer, 2007), suggesting that even students from integrated high schools may have little experience with diversity. College is the first time most students experience significant interaction with diverse others (Hurtado, Engberg, Ponjuan, & Landerman, 2002; Rankin & Reason, 2005).

When students arrive on campus and encounter a range of diverse fellow students and new ideas, they enter a living laboratory with great potential for allowing them to learn and grow. This environment can provide students an opportunity to learn about different perspectives and cultures and confront issues of racism, bigotry, and oppression (Chang, Astin, & Kim, 2004; Gurin, Dey, Hurtado, & Gurin, 2002; Pike & Kuh, 2006), but there is also potential for interactions in this diverse environment to harm students' attitudes about diversity (Hurtado, Milem, Clayton-Pedersen, & Allen, 1998), especially if those interactions are not supported and facilitated by the institution (Chang, 1996; Gurin, 1999). When considering interventions and programs designed to improve students' diversity outcomes through structural, instructional, and interactional diversity, most researchers have focused on classes focused on multicultural issues and diversity (Bowman, 2010; Hurtado, 2001; Marin, 2000), intergroup dialogues (Hurtago; Nagda & Zuniga, 2003), and other diversity-focused programs (Hyun, 1994; Milem, 1994; Whitt, Edison, Pascarella, Terenzini, & Nora, 2001). Service-learning programs, however, are largely absent from the larger discussion of improving students' exposure to interactional diversity. Engberg (2004) reviewed 13 studies of the impact of service-learning participation on students' racial bias in his review of diversity-related educational programs; but despite a myriad of studies linking service-learning to diversity outcomes, other researchers studying ways that college improves diversity outcomes have not considered service-learning as a diversity intervention (Chang, 1999; Denson, 2009; Gurin et al., 2002; Milem, Chang, & Antonio, 2005; Pascarella, Edison, Nora, Hagedorn, & Terenzini, 1996).

Despite this absence in much of the literature on diversity outcomes in higher education, service-learning scholars and advocates have long promoted the pedagogy as a way to address diversity-related outcomes. Zlotkowski (1996) suggested that “issues of diversity and multiculturalism” (p. 26) is one of the areas of education for which service-learning is best-suited. This lies primarily in the potential for students to interact across difference and form relationships with the members of the served populations while engaged in their service experiences (Jones & Hill, 2001). Soukop (1996) stated that service-learning was an ideal setting to provide students with
Holsapple

“first-hand experience of diversity and multiculturalism” (p. 9), by providing students with experiences with people different from themselves. In a recent survey of 1,100 colleges and universities, more than 60 percent reported supporting service-learning programs with a focus on diversity issues (Campus Compact, 2009).

Although the extant literature does provide support for service-learning participation encouraging the development of diversity outcomes, there has been little attempt in the literature to develop a cohesive body of research on the topic that systematically builds a body of knowledge or supports mechanisms for the development of the outcome from the experiences. Instead the literature is largely comprised of studies of individual service-learning courses, with little connection to one another (Furco, 2003). This lack of connection among studies has limited the development of a foundation of knowledge that supports improving the design of service-learning programs to enhance diversity outcomes.

Recognizing that lack of foundation, this paper addresses a central research question: What is the impact of participating in service-learning programs on students’ diversity outcomes? I will present a review of 55 published empirical studies assessing diversity outcomes from service-learning, synthesizing results and presenting a classification of these outcomes. I also will discuss the methodological limitations of this body of literature and the restrictions of the knowledge claims based on these limitations, and make suggestions for researchers that will lead to the building of that missing foundation.

Methods and Data

When identifying published papers with diversity outcomes, I relied on a definition that draws from the Openness to Diversity/Challenge assessment tool by Pascarella et al. (1996). The authors were interested in moving the research on the effects of diversity in college beyond outcomes such as reporting a commitment to racial understanding; instead they aimed to move the discourse toward the investigation of how “specific programmatic or policy-relevant college experiences increase the value on experiencing the different dimensions of diversity” (p. 176). Openness to diversity is not only an assessment of an individual’s openness to cultural, racial, and value diversity, it also taps the extent to which an individual enjoys being challenged by different ideas, values, and perspectives as well as an appreciation of racial, cultural, and value diversity (p. 179).

When selecting papers with diversity outcomes, I used several other criteria to identify the studies to include in this review: a publishing date between 1998 and 2010; an undergraduate context; the study of credit-bearing service-learning; a focus on U.S. institutions, and an empirical investigation with explicit reporting of student outcomes. Because the peer review process for conferences is inconsistent and often less rigorous than journal articles, conference proceedings were not considered for this analysis. I did, however, include two studies published outside of journals – Where’s the Learning in Service-Learning? (Eyler & Giles, 1999) and How Service-Learning Affects Students (Astin, Vogelgesang, Ikeda, & Yee, 2000) – because of their prominence in the scholarly discourse surrounding the effects of service-learning participation. I searched for studies through several channels, including searching a range of electronic databases such as ERIC and PsychInfo, broader searching through Google Scholar, conducting an issue-by-issue search of the Michigan Journal, and consulting reference lists in found articles. Due to the wide range of discipline- and field-based journals, I cannot claim that this search has identified every article matching the established criteria. However, analysis of the articles in this review reached theoretical saturation and information redundancy (Strauss & Corbin, 1990), suggesting that the sample of reviewed papers was sufficient for analysis. The studies included in this review comprise a collection much larger than any previous published review of this type.

While conducting qualitative analysis of these studies, each was first coded to catalogue aspects of the service-learning program that was studied and the methods used to study it. I coded each study based on the field or discipline of the course; the theoretical foundation of the study; a description of the study participants, including sample size, student characteristics, and number of institutions; the community service experience; and the methods and data used in the study.

I then used a constant comparative open coding process (Strauss & Corbin, 1990) for the results sections of each study. I considered the results based both on the way the authors described the results and the voices of the students themselves when they were included in the paper. The open-coding process resulted in identifying 197 individual codes relating to diversity outcomes. I then used axial coding (Strauss & Corbin) to group these codes into six main themes of the outcome: tolerance of difference, stereotype confrontation, recognition of universality, interactions across difference, knowledge about the served population, and belief in the value of diversity. The themes are presented in order number of appearances in the sample; no hypothesized developmental order, which is far beyond the scope of this type of secondary analysis, should be interpreted.
These 55 studies represent a wide range of fields of study, programs and interventions, and types of diversity addressed. Most commonly the studies consisted of presenting the effects of a single service-learning course or program. Consequently, 45 of the studies have a single-institution focus; the remaining 10 studies with a multi-institutional focus include students from a range of 2-388 schools. Thirty-three studies were conducted using qualitative methods, 11 used only quantitative methods, and 11 used mixed methods. The most common source of data in the studies were students’ assigned written reflections for the course, with 36 of the studies using some type of assigned course work for data and 22 using only assigned coursework as data. Tables 1 and 2 provide more information about the studies included in the sample for this paper.

Results
Stereotype Confrontation (n=32)

The most common diversity outcome reported in the reviewed studies was students’ confrontation of personal or societal stereotypes about the population with which the students worked. Of the 55 studies reviewed, more than half (32) reported that students confronted previously held stereotypes during the course of their service. Eyler and Giles (1999) interviewed 133 students at 12 institutions who had participated in a range of service-learning programs – the largest qualitative study of its kind examining the impact of service-learning. They reported a reduction of negative stereotypes was the most common outcome that students described from their experiences. Several students worked at social service agencies that provided service to HIV/AIDS patients. These students reported being surprised to learn that fellow volunteers with whom they had been working were gay, and reread their personal stereotypes after getting to know gay people who did not fit those stereotypes.

This confrontation and reduction of stereotypes through relationships with diverse others is a consistent theme in the reviewed studies. Teranishi (2007) found that students who conducted service with local families in Mexico described their relationships with the families as invaluable to reducing stereotypes. Several researchers found that students who completed service-learning programs consisting of socialization and reminiscence with socially isolated elderly adults – focused on developing relationships with them – led to reducing negative stereotypes about seniors (Brown & Roodin, 2001; Dorfmann, Murty, & Ingram, 2004; Dorfmann, Murty, Ingram, Evans, & Power, 2003; Wakefield & Erickson, 2003).

Other researchers found the development of relationships during service helped students to confront a wide variety of stereotypes about a wide range of categories of difference, including sexual orientation (Williams & Reeves, 2004), HIV/AIDS patients (Jones & Hill, 2001), race and ethnicity (Boyle-Baise, 2005; Everett, 1998; Long, 2003); religion (Giles & Eyler, 1994) and disability (Smith, 2003). In some cases, the design of the service-learning experience itself directed that relationship specifically toward the negative stereotypes that students held. Hale (2008) reports on a service-learning program in which preservice teachers tutored Mexican immigrants in a high school equivalency program. One student said of the experience:

Going in there, I had stereotypes. Then actually getting to work with them and talking with them … I saw they wanted a successful future. Seeing these guys sitting down and studying, asking questions, and being very involved in their classroom changed my whole perspective of what I used to think. (p. 59)

This type of stereotype confrontation was common in educational settings with such served populations as incarcerated prisoners (Amtmann, 2004), non-literate, Spanish-speaking adults (Plann, 2002), and the elderly (Williams & Kovac, 2001). Targeting of educational stereotypes also helped students to confront their stereotypes about low-income elementary and secondary students (Etheridge & Branscombe, 2009; Hughes, Welsh, Mayer, Bolay, & Southard, 2009; Simon & Clear, 2006) and their parents and families (Baldwin, Buchanon, & Rudisill, 2007; Boyle-Baise & Lanford, 2004; Childs, Sepples, & Moody, 2003; Davi, 2006).

Most of the support (23 studies) for the confrontation and reduction of stereotypes of service-learning comes from qualitative studies; however, there are also two quantitative studies that also found support for this outcome. Spezio et al. (2007) surveyed more than 1,200 students at four institutions – of whom 524 were involved in service-learning. Students were surveyed in the first and last weeks of the semester; those who participated in the service-learning courses reported a statistically significant increase to the survey item stating that they are better than the average student at “being aware of my own biases and prejudices” (p. 282), while non-service-learning students did not demonstrate a statistically significant change on the same item. Seven studies found similar support in the quantitative portions of mixed methods studies.

Knowledge about the Served Population (n=28)

After stereotype confrontation, the other most common reported outcome is the development of knowledge about the served population, which was found in 28 papers. This outcome applied to different types of
Table 1
Details of the 55 Studies Comprising the Analytic Sample

<table>
<thead>
<tr>
<th>Study</th>
<th>Outcomes</th>
<th>Methodology</th>
<th>Number of Students</th>
<th>Number of Institutions</th>
<th>Data Collecting Method(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amtmann, 2004</td>
<td>Stereotype confrontation, Recognition of universality</td>
<td>Mixed-methods</td>
<td>9</td>
<td>1</td>
<td>Interviews, Surveys</td>
</tr>
<tr>
<td>Astin, Vogelgesang, Ikeda, &amp; Yee, 2000</td>
<td>Tolerance of difference</td>
<td>Mixed-Methods</td>
<td>22,236</td>
<td>19</td>
<td>Surveys, Interviews</td>
</tr>
<tr>
<td>Astin &amp; Sax, 1998</td>
<td>Tolerance of difference, Interactions across difference, Knowledge about the served population</td>
<td>Quantitative</td>
<td>3,450</td>
<td>42</td>
<td>Surveys</td>
</tr>
<tr>
<td>Astin, Sax, &amp; Avalos, 1999</td>
<td>Tolerance of difference, Interactions across difference, Knowledge about the served population</td>
<td>Quantitative</td>
<td>27,064</td>
<td>388</td>
<td>Surveys</td>
</tr>
<tr>
<td>Baldwin, Buchanon, &amp; Radisill, 2007</td>
<td>Stereotype confrontation, Belief in the value of diversity</td>
<td>Qualitative</td>
<td>41</td>
<td>2</td>
<td>Interviews, Assigned written reflections, Researcher observations</td>
</tr>
<tr>
<td>Borden, 2007</td>
<td>Tolerance of difference</td>
<td>Quantitative</td>
<td>40</td>
<td>1</td>
<td>Surveys</td>
</tr>
<tr>
<td>Boyle-Baise &amp; Sleeter, 1998</td>
<td>Stereotype confrontation, Recognition of universality, Belief in the value of diversity</td>
<td>Qualitative</td>
<td>117</td>
<td>2</td>
<td>Interviews, Assigned written reflections, Researcher observations</td>
</tr>
<tr>
<td>Boyle-Baise &amp; Langford, 2004</td>
<td>Stereotype confrontation</td>
<td>Qualitative</td>
<td>8</td>
<td>1</td>
<td>Interviews, Assigned written reflections, Researcher observations</td>
</tr>
<tr>
<td>Boyle-Baise, 1998</td>
<td>Stereotype confrontation, Interactions across difference, Knowledge about the served population</td>
<td>Qualitative</td>
<td>65</td>
<td>1</td>
<td>Surveys, Group Interviews, Course projects, Assigned written reflections</td>
</tr>
<tr>
<td>Boyle-Baise, 2005</td>
<td>Stereotype confrontation, Belief in the value of diversity</td>
<td>Qualitative</td>
<td>24</td>
<td>1</td>
<td>Researcher observations, Interviews, Assigned written reflections</td>
</tr>
<tr>
<td>Brown &amp; Roodin, 2001</td>
<td>Stereotype confrontation, Recognition of universality, Knowledge about the served population</td>
<td>Qualitative</td>
<td>104</td>
<td>1</td>
<td>Class discussions, Course evaluations</td>
</tr>
<tr>
<td>Childs, Sepples, &amp; Moody, 2003</td>
<td>Stereotype confrontation, Recognition of universality</td>
<td>Qualitative</td>
<td>Not reported</td>
<td>1</td>
<td>Assigned written reflections</td>
</tr>
<tr>
<td>Davi, 2006</td>
<td>Stereotype confrontation, Knowledge about the served population, Belief in the value of diversity</td>
<td>Qualitative</td>
<td>Not reported</td>
<td>1</td>
<td>Assigned written reflections</td>
</tr>
<tr>
<td>Dorfmann, Murty, &amp; Ingram, 2004</td>
<td>Stereotype confrontation, Belief in the value of diversity</td>
<td>Quantitative</td>
<td>13</td>
<td>1</td>
<td>Surveys</td>
</tr>
<tr>
<td>Dorfmann, Murty, Ingram, Evans, &amp; Power, 2004</td>
<td>Stereotype confrontation, Knowledge about the served population</td>
<td>Mixed-Methods</td>
<td>59</td>
<td>1</td>
<td>Surveys</td>
</tr>
<tr>
<td>Elwell &amp; Bean, 2001</td>
<td>Recognition of universality, Knowledge about the served population</td>
<td>Qualitative</td>
<td>28</td>
<td>1</td>
<td>Class discussions, Assigned written reflections</td>
</tr>
<tr>
<td>Essson, Stevens-Truss &amp; Thomas, 2005</td>
<td>Interactions across difference</td>
<td>Mixed-Methods</td>
<td>58</td>
<td>1</td>
<td>Surveys, Assigned written reflections</td>
</tr>
<tr>
<td>Ethridge &amp; Branscomb, 2009</td>
<td>Stereotype confrontation</td>
<td>Qualitative</td>
<td>19</td>
<td>1</td>
<td>Assigned written reflections</td>
</tr>
<tr>
<td>Evertt, 1998</td>
<td>Stereotype confrontation, Knowledge about the served population</td>
<td>Mixed-Methods</td>
<td>105</td>
<td>1</td>
<td>Surveys, Assigned written reflections</td>
</tr>
<tr>
<td>Eyler &amp; Giles, 1999</td>
<td>Tolerance of difference, Stereotype confrontation</td>
<td>Mixed-Methods</td>
<td>1,544</td>
<td>45</td>
<td>Surveys, Interviews</td>
</tr>
<tr>
<td>Flannery &amp; Ward, 1999</td>
<td>Tolerance of difference, Knowledge about the served population</td>
<td>Qualitative</td>
<td>103</td>
<td>1</td>
<td>Assigned written reflections</td>
</tr>
<tr>
<td>Greene, 1998</td>
<td>Knowledge about the served population</td>
<td>Quantitative</td>
<td>36</td>
<td>2</td>
<td>Surveys</td>
</tr>
<tr>
<td>Hale, 2008</td>
<td>Stereotype confrontation, Knowledge about the served population</td>
<td>Qualitative</td>
<td>8</td>
<td>1</td>
<td>Interviews, Assigned Written Reflections, Student presentations</td>
</tr>
<tr>
<td>Handa et al., 2008</td>
<td>Tolerance of difference, Knowledge about the served population</td>
<td>Qualitative</td>
<td>Not reported</td>
<td>1</td>
<td>Assigned written reflections</td>
</tr>
<tr>
<td>Hollis, 2004</td>
<td>Stereotype confrontation</td>
<td>Qualitative</td>
<td>105</td>
<td>1</td>
<td>Assigned written reflections</td>
</tr>
<tr>
<td>Hughes, Boyd, &amp; Dykstra, 2010</td>
<td>Knowledge about the served population</td>
<td>Qualitative</td>
<td>49</td>
<td>1</td>
<td>Assigned written reflections</td>
</tr>
<tr>
<td>Hughes, Welsh, Mayer, Bolay, &amp; Southard, 2009</td>
<td>Stereotype confrontation, Knowledge about the served population</td>
<td>Qualitative</td>
<td>32</td>
<td>1</td>
<td>Assigned written reflections</td>
</tr>
<tr>
<td>Jakubowski, 2003</td>
<td>Knowledge about the served population, Belief in the value of diversity</td>
<td>Qualitative</td>
<td>4</td>
<td>1</td>
<td>Assigned written reflections</td>
</tr>
</tbody>
</table>
Table 1 continued
Details of the 55 Studies Comprising the Analytic Sample

<table>
<thead>
<tr>
<th>Study</th>
<th>Outcomes</th>
<th>Methodology</th>
<th>Number of Students</th>
<th>Number of Institutions</th>
<th>Data Collecting Method(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jones &amp; Hill, 2001</td>
<td>Stereotype confrontation, Recognition of universality, Interactions across difference, Knowledge about the served population, Belief in the value of diversity</td>
<td>Qualitative</td>
<td>14</td>
<td>1</td>
<td>Interviews</td>
</tr>
<tr>
<td>Keen &amp; Hall, 2009</td>
<td>Interactions across difference</td>
<td>Quantitative</td>
<td>700</td>
<td>25</td>
<td>Surveys</td>
</tr>
<tr>
<td>Keselyak, Simmer-Beck, Bray, &amp; Gadbury-Am Hoy, 2007</td>
<td>Interactions across difference, Knowledge about the served population</td>
<td>Qualitative</td>
<td>23</td>
<td>1</td>
<td>Assigned written reflections</td>
</tr>
<tr>
<td>King, 2004</td>
<td>Interactions across difference, Knowledge about the served population</td>
<td>Qualitative</td>
<td>4</td>
<td>1</td>
<td>Interviews, Program application essays</td>
</tr>
<tr>
<td>Long, 2003</td>
<td>Stereotype confrontation, Knowledge about the served population, Belief in the value of diversity</td>
<td>Qualitative</td>
<td>11</td>
<td>1</td>
<td>Assigned written reflections, Class discussions</td>
</tr>
<tr>
<td>Malone, Jones, &amp; Stallings, 2002</td>
<td>Tolerance of difference, Knowledge about the served population</td>
<td>Mixed-Methods</td>
<td>108</td>
<td>1</td>
<td>Survey, Assigned written reflections</td>
</tr>
<tr>
<td>Marullo, 1998</td>
<td>Tolerance of difference, Knowledge about the served population</td>
<td>Quantitative</td>
<td>41</td>
<td>1</td>
<td>Surveys</td>
</tr>
<tr>
<td>Miciano, 2006</td>
<td>Knowledge about the served population</td>
<td>Qualitative</td>
<td>15</td>
<td>1</td>
<td>Assigned written reflections, Student focus groups</td>
</tr>
<tr>
<td>Moely, Furco, &amp; Reed, 2008</td>
<td>Belief in the value of diversity</td>
<td>Quantitative</td>
<td>2,233</td>
<td>7</td>
<td>Surveys</td>
</tr>
<tr>
<td>Moely, McFarland, Miron, Mercer, &amp; Illustre, 2002</td>
<td>Belief in the value of diversity</td>
<td>Quantitative</td>
<td>741</td>
<td>1</td>
<td>Surveys</td>
</tr>
<tr>
<td>Morris, 2001</td>
<td>Belief in the value of diversity</td>
<td>Mixed-Methods</td>
<td>95</td>
<td>1</td>
<td>Surveys</td>
</tr>
<tr>
<td>Muradha-Watts, 1998</td>
<td>Stereotype confrontation</td>
<td>Qualitative</td>
<td>22</td>
<td>1</td>
<td>Interviews, Assigned written reflections</td>
</tr>
<tr>
<td>Paolotti, Segal, &amp; Totino, 2007</td>
<td>Stereotype confrontation, Recognition of universality, Belief in the value of diversity</td>
<td>Qualitative</td>
<td>40</td>
<td>1</td>
<td>Assigned written reflections</td>
</tr>
<tr>
<td>Pasricha, 2008</td>
<td>Tolerance of difference, Knowledge about the served population</td>
<td>Qualitative</td>
<td>20</td>
<td>1</td>
<td>Assigned written reflections</td>
</tr>
<tr>
<td>Plan, 2002</td>
<td>Stereotype confrontation, Recognition of universality, Interactions across difference</td>
<td>Qualitative</td>
<td>Not reported</td>
<td>1</td>
<td>Assigned written reflection, Course evaluations</td>
</tr>
<tr>
<td>Reed, Jernstedt, Hawley, Reber, &amp; DuBois, 2005</td>
<td>Interactions across difference</td>
<td>Quantitative</td>
<td>33</td>
<td>1</td>
<td>Surveys</td>
</tr>
<tr>
<td>Race &amp; Brown, 1998</td>
<td>Stereotype confrontation, Knowledge about the served population, Belief in the value of diversity</td>
<td>Mixed-Methods</td>
<td>230</td>
<td>1</td>
<td>Survey, Course evaluations</td>
</tr>
<tr>
<td>Sedlak, Donehy, Panthofer, &amp; Anaya, 2003</td>
<td>Stereotype confrontation, Belief in the value of diversity</td>
<td>Qualitative</td>
<td>94</td>
<td>1</td>
<td>Assigned written reflections</td>
</tr>
<tr>
<td>Shaw &amp; Jolley, 2007</td>
<td>Knowledge about the served population</td>
<td>Qualitative</td>
<td>9</td>
<td>1</td>
<td>Interviews, Assigned written reflections</td>
</tr>
<tr>
<td>Simons &amp; Cleary, 2006</td>
<td>Stereotype confrontation, Knowledge about the served population, Belief in the value of diversity</td>
<td>Mixed-Methods</td>
<td>140</td>
<td>1</td>
<td>Surveys</td>
</tr>
<tr>
<td>Smith, 2003</td>
<td>Stereotype confrontation, Recognition of universality, Interactions across difference</td>
<td>Qualitative</td>
<td>7</td>
<td>1</td>
<td>Interviews, Assigned written reflections, Researcher observations</td>
</tr>
<tr>
<td>Spezio, Baker, &amp; Boland, 2005</td>
<td>Tolerance of difference, Stereotype confrontation, Knowledge about the served population</td>
<td>Quantitative</td>
<td>1,243</td>
<td>4</td>
<td>Surveys</td>
</tr>
<tr>
<td>Stachowski, Bodle, &amp; Morrin, 2008</td>
<td>Interactions across difference, Knowledge about the served population</td>
<td>Qualitative</td>
<td>88</td>
<td>1</td>
<td>Assigned written reflections</td>
</tr>
<tr>
<td>Teranishi, 2007</td>
<td>Tolerance of difference, Stereotype confrontation, Recognition of universality, Knowledge about the served population, Belief in the value of diversity</td>
<td>Mixed-Methods</td>
<td>11</td>
<td>1</td>
<td>Surveys, Assigned written reflections</td>
</tr>
<tr>
<td>Wakefield &amp; Erickson, 2003</td>
<td>Stereotype confrontation, Knowledge about the served population</td>
<td>Qualitative</td>
<td>19</td>
<td>1</td>
<td>Assigned written reflections</td>
</tr>
<tr>
<td>Williams &amp; Reeves, 2004</td>
<td>Stereotype confrontation, Recognition of universality, Belief in the value of diversity</td>
<td>Qualitative</td>
<td>21</td>
<td>1</td>
<td>Assigned written reflections, Focus groups, Course evaluations</td>
</tr>
<tr>
<td>Williams &amp; Kovacs, 2001</td>
<td>Stereotype confrontation</td>
<td>Qualitative</td>
<td>Not reported</td>
<td>1</td>
<td>Assigned written reflections</td>
</tr>
</tbody>
</table>
knowledge, including factual knowledge about population traditions; knowledge of marginalization; and an understanding of diversity within the population.

Gaining factual knowledge about traditions of the served population was most likely to occur when working with international or immigrant populations. Long (2003) describes this in a program in which students perform at least 100 hours of service with the local Spanish-speaking community, primarily recent immigrants and their families. One student described a conversation he had about differences between Mexican food served in the U.S. and in Mexico; the student said,

Luis told me that the food there isn’t very authentic. He told me that in Mexico, they eat less meat. He likes the food in [the restaurant], but he told me that they don’t eat such large portions like they gave us. He couldn’t finish the whole burrito. (p. 228)

Other studies reported students learning about cultural practices – such as meals, dress, and household activities – from members of the served population (Jakubowski, 2003; King, 2004; Pasricha, 2008). Astin and Sax (1998) found that students who had participated in service activities – including service-learning – were more likely than those who hadn’t to say that they had gained “Knowledge of people of different races and cultures” (p. 258) during college.

In other studies, outcomes moved beyond factual knowledge to a more thorough understanding of the ways disadvantage and marginalization affect the served population. Hale (2008) reported that students working with Spanish-speaking children gained an understanding for the difficulty they face in school when they and their parents speak no or little English. One student said,

The parents are mainly Spanish speaking with little or no English, so the children are forced to translate, but who knows what is lost in translation. In some cases the parents want to be more involved in their child’s education, but the language barrier, and even the level of education they have, created many problems for parents. (p. 63)

These types of realizations were common for students in several other studies (Boyle-Baise, 1998; Elwell & Bean, 2001; Everett, 1998; Gorlick, 2002; Handa et al., 2008; Hughes, Boyd, & Dykstra, 2010; Long, 2003; Miciano, 2006; Teranashi, 2007).

The third type of knowledge about diverse others students gained during service-learning programs is the recognition of diversity within the served population. Prior to service-learning experiences, students thought of the served population as a homogenous group, with the same experiences, backgrounds, and perspectives; this changed during the program. Jones and Hill (2001) report on a study of a program in which students spend a semester working at one of two social service agencies, one of which provides services to HIV/AIDS patients; students involved with that agency gained an understanding of the wide range of races, sexual orientations, and backgrounds of those patients. Greene (1998) found the same effect from a program in which occupational therapy students conducted social visits with socially-isolated adults. After performing content analysis on students’ responses to a set of open-ended questions, he found that 58.7% of students’ responses on these questions reflected gaining a greater understanding of diversity among seniors and adults with disabilities. Other studies reported the same outcome from programs that emphasized prolonged contact with people with disabilities (Keselyak et al., 2007; Shaw & Jolley, 2007), seniors (Brown & Roodin, 2001), children (Wakefield & Erickson, 2003), and residents of developing countries (Stuchowski, Bodle, & Morrin, 2008).

### Table 2

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Total</th>
<th>Qualitative (n=33)</th>
<th>Quantitative (n=11)</th>
<th>Mixed-Methods (n=11)</th>
<th>Some Data Derived from Assigned Coursework (n=36)</th>
<th>All Data Derived From Assigned Coursework (n=22)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tolerance of difference</td>
<td>12</td>
<td>3</td>
<td>5</td>
<td>4</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>Stereotype confrontation</td>
<td>32</td>
<td>23</td>
<td>2</td>
<td>7</td>
<td>24</td>
<td>14</td>
</tr>
<tr>
<td>Recognition of universality</td>
<td>11</td>
<td>9</td>
<td>0</td>
<td>2</td>
<td>9</td>
<td>5</td>
</tr>
<tr>
<td>Knowledge about the served population</td>
<td>28</td>
<td>18</td>
<td>4</td>
<td>6</td>
<td>20</td>
<td>12</td>
</tr>
<tr>
<td>Interactions across difference</td>
<td>12</td>
<td>7</td>
<td>4</td>
<td>1</td>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td>Belief in the value of diversity</td>
<td>18</td>
<td>11</td>
<td>3</td>
<td>4</td>
<td>11</td>
<td>7</td>
</tr>
</tbody>
</table>

Belief in the Value of Diversity (n=18)

The conclusions of 18 studies provided support for...
belief in the value of diversity as an outcome of service-learning. Morris (2001) reported a study of students in a Spanish service-learning course in which students were required to work each week with social agencies supporting the Spanish-speaking community. Of the 152 students who took the course, 95 were selected for the study based on responses to a pre-class interview and survey that suggested they had “low motivation towards learning Spanish and indifference towards the cultures and/or the speakers of Spanish” (Morris, p. 247). Morris conducted pre- and post-test surveys which suggested that students had gained a belief in the value of diversity during their experiences. On the survey’s quantitative measures, students were more likely to agree to statements supporting the importance of both learning Spanish and gaining a better understanding of “Spanish-speaking people and their cultures” (p. 250), as well as motivation to continue learning.

In the survey’s open-ended questions, students reported gaining a belief in the importance of learning about the lives and culture of Spanish-speaking people. A student wrote,

I never wanted or cared to learn Spanish. I did it because it was the thing to do. But now I believe it because Spanish is a rich language tied to great cultures and traditions. I want to learn more about the language and the cultures. I am fascinated by the different people I have met and I look forward to meeting more. (p. 251)

Morris’s findings provide support for the potential to significantly improve students’ belief in the value of diversity, particularly when students demonstrated a prior lack of interest.

Simons and Cleary (2006) conducted a mixed-methods study of a program in which pre-service teachers tutored low-income students of color. The college students’ responses to open-ended questions suggested they had gained an appreciation for the value of different experiences and cultures, and quantitative results showed that students demonstrated improved diversity attitudes during the semester. Other studies found similar results, demonstrating service-learning experiences can be associated with a belief in the value of diversity, particularly in K-12 school-based experiences (Boyle-Baise & Langford, 2004; Davi, 2006; Paolletti et al., 2007) and experiences with international or immigrant populations (Jakubowski, 2003; Long, 2003; Sedlak, Doheny, Panthofer, & Anaya, 2003; Stachowski et al., 2008; Teranishi, 2007).

Tolerance of Difference (n=12)

Pascarella et al. (1996) described openness to diversity as an extension beyond simply the lack of negative feelings toward diverse others, a state termed “tolerance” by Astin (1993), Chickering and Reisser (1993), and others. While this tolerance is not a satisfactory end goal of education, it is a necessary first step for some students and was presented as an outcome of service-learning in 12 of the studies analyzed for this paper. Eyler and Giles (1999) surveyed 1,544 students at 47 institutions, and found students who participated in service-learning were more likely to report growth in the development of tolerance of difference than their non-service-learning counterparts, while controlling for students’ characteristics, relationships, and other activities. Other quantitative studies found that service-learning participation positively predicts students’ increasing “global understanding” and “respecting the views of others” (Marullo, 1998, p. 268), and “being tolerant of other people’s differences” (Spezio, Baker, & Boland, 2006, p. 282), and decreases in ethnocentrism (Borden, 2007). Astin and colleagues found similar results in a series of studies of outcomes for students participating in service, including service-learning (Astin & Sax, 1998, Astin, Sax, & Avalos, 1999; Astin, Vogelgesang, Ikeda, & Yee, 2000).

Using qualitative methods, Pasricha (2008) reported that students participating in a service-learning course requiring 15 hours of service providing tutoring for refugee communities learned tolerance and respect of the communities with which they had worked; one student wrote, “Overall, I have to say I have a new respect for people trying to move to a new country and learn a different language” (p. 50). Other researchers found similar results in other studies in which they analyzed students’ reflection journals and other reflective writings (Flannery & Ward, 1999; Malone, Jones, & Stallings, 2002; Teranishi, 2007).

Interactions across Difference (n=12)

In service-learning, learning is often hypothesized to occur through regular interaction between the student and the served population, but in this sample 12 studies present the development of these relationships as its own outcome. Authors report students becoming more comfortable across difference and more likely to interact with diverse others outside of the service-learning environment.

Two large quantitative studies support the idea that service-learning can help students to interact across difference. Astin and Sax (1998) examined longitudinal responses for 3,450 students at 42 institutions, including more than 2,000 who had participated in service activities, including service-learning. Those students were more likely to report than they had improved their “ability to get along with people of different races and cultures” (p. 258). In another large study, Astin et al. (1999) found those who had partic-
ipated in service activities were more likely to report “socializing with persons from other racial/ethnic groups.” Other studies lent further support to interactions across difference as an outcome. Esson, Stevens-Truss, and Thomas (2005) report on a chemistry course in which students worked in groups to teach chemistry to elementary students. In a post-test survey, students agreed that “I am now more comfortable working with people of cultures other than my own” and “This service-learning project increased my ability to get along with different kinds of people with different lifestyles” (p. 1171).

Keselyak, Simmer-Beck, Bray, and Gadbury-Amyot (2007) described a study in which dentistry students provided dental services to children with severe disabilities, and students in the program reported that they became more comfortable interacting with the disabled children and their families as well as more comfortable providing dental services to people with disabilities. Similarly, Reed, Jernstedt, Hawley, Reber, and DuBois (2005) reported on a service-learning course on end-of-life issues in which students spent time interacting with and socializing with end-of-life patients. After the program, students reported that they were more likely to report being comfortable “talking with dying patients” (p. 363) than a randomly selected control group.

Recognition of Universality (n=11)

The final outcome in the reviewed studies is the recognition of similarities and common ground with people who at first seem to be different. Williams and Reeves (2004) report changes in their students that are typical of this outcome; after spending a week volunteering at a camp for severely burned students, one student wrote in her reflection journal,

I remembered someone asked me, ‘Are they pitiful?’ I responded, ‘No, they’re kids! You know, they have joy just like other kids and they’re running around having fun.’ And I found myself not pitying (them). I just see them as kids, and I wasn’t feeling sorry for them. I was glad about that. (p. 393)

The student expressed happiness that she was de-emphasizing the difference between the students at the camp and other children, and instead was emphasizing the ways in which these children were like all others. Other studies found similar outcomes for service-learning programs involving children and adolescents (Boyle-Baise & Langford, 2004; Boyle-Baise & Sleeter, 1998; Childs et al., 2003; King, 2004; Paoletti, Segal, & Totino, 2007).

This universality outcome was also reported in service-learning programs in which students worked with adult populations. Plann (2002) and Teranashi (2007) reported that students recognized universality through providing tutoring to Spanish-speaking immigrants. Jones and Hill (2001) reported students recognizing universality in regards to HIV/AIDS patients; the students recognized that these patients were like them in ways they did not expect. One student said in an interview,

Stigma is placed on people with AIDS and so it was nice to see, no, they’re not different … It was very eye-opening in that it made me realize just how very alike everybody is in one way or another. (Jones & Hill, 2001, p. 209)

Through working closely across difference, the student came to see the served population as more similar to her than she has first assumed.

Literature in this review represents research on the effect of service-learning on student’s diversity outcomes. The review provided consistent support for the potential of service-learning to encourage several diversity outcomes – tolerance of difference, stereotype confrontation, recognition of universality, interactions across difference, knowledge about the served population, and belief in the value of diversity. The studies most commonly addressed the confrontation of stereotypes and gaining knowledge about diverse others.

Methodological Limitations of the Research

The 55 studies reviewed in this paper demonstrate generally consistent results that service-learning can encourage the development of students’ diversity outcomes. However, significant methodological concerns in the body of the literature limit the confidence one can have in that support and how that knowledge can be used in the development of future service-learning programs. In this section, I outline five such limitations: lack of theoretical models, sample selection, external validity to other programs, lack of detail, and trustworthiness of data.

Lack of Theoretical Models

While the pedagogy and use of service-learning have grown and become more sophisticated over the past two decades, the research on service-learning largely has not followed the same path. Furco (2003) raises concerns that most of the published studies of service-learning “have been unable to make definitive statements about the impact of service-learning on students” (p. 13) because service-learning programs are so diverse that the study of one program can often tell researchers little about other programs. The research agenda of the field has done little to draw connections between these diverse programs:

Unlike most other fields in which one research
The studies reviewed for this paper continue to fit the model of research that Furco described. To move beyond this model of research, studies – and service-learning programs themselves – must be designed with a foundation in theoretical models and relationships that provide a framework for understanding the outcomes from the service-learning experience. Scholars such as Bringle (2003), Bringle and Hatcher (2000), Butin, (2003) and Engberg (2004) point out that this theoretical foundation is missing or limited in the majority of service-learning research, which severely limits claims researchers can make about the efficacy of different types of service-learning initiatives on specific outcomes.

Few authors in the sample explicitly use a theory or model to identify mechanisms for learning within the programs they studied. Studies consistently present results that support the potential of service-learning to enhance diversity outcomes and many authors point to the importance of interactions and relationships with the served community as being important, but there is little exploration of mechanisms that create learning from these service experiences and relationships.

There are exceptions to this lack of explicit theory informing the research. Reed et al. (2005) rely in part on the theory of self-perception. Miciano (2006) explicitly drew upon four theories of learning which he integrates into one conceptual model of how a tutoring service-learning program will develop outcomes in both the students and the served population. Studies by Astin and colleagues (Astin et al., 2000; Astin & Sax, 1998; Astin et al., 1999) rely on Astin’s own Inputs-Environments-Outcomes model (1993) which identifies student outcomes as a product of students’ characteristics and the whole of their college experiences. They used this framework primarily to identify covariates in analysis, but they do not suggest explanations as to why inputs and environments interact in specific ways.

Without these explicit frameworks, a large amount of the literature is performed without explicit reasons for the sampling methods, choice of data sources, and specificity or breadth of research questions. So, instead of research that builds upon itself creating theory and a more thorough understanding, the field is left with the idiosyncratic studies in research vacuums.

**Sample Selection**

One significant limitation of the body of literature is the sampling methods employed. In the majority of the studies, little is done to select a sample in a way that is either exogenous to the outcomes or purposefully based on criteria important to the study. Most often, the sample is selected simply because of students’ decisions to participate in the studied course, severely limiting the generalizability of the studies because of the error that selection bias introduces into results by not accounting for ways that students who take the course may be different – in observable and unobservable ways – from students who do not choose to take the particular course. College service and character education activities, like service-learning, “thrive on nonrandomness” (Russell, 2004, p. 106).

In quantitative studies, researchers can include a large set of covariates in regression models to account for observable differences between participating and non-participating students (when a control group was used); these cannot address unobserved differences in treatment and control groups. Some studies used other methods to limit the effect of unobservable differences. Hollis (2004) taught the same service-learning course with different methods in two consecutive semesters and compared student outcomes between the semesters. Marullo (1998) taught two sections of the same class, and only on the first day of class told students that one of the sections would have a service-learning component, and he compared outcomes between the sections. More studies like these would allow researchers to begin understanding effects of service-learning in a way that better controls the effect of selection bias.

A common sampling method employed was to simply sample all students who voluntarily chose to take the service-learning course, with little attention paid to their other characteristics. These studies investigate outcomes of a given program, but even this approach allows for more purposeful sampling. Morris (2001) included in his sample only some of the students who were participating in the program over multiple semesters, choosing students who had the least interest in learning about Hispanic cultures. This selection method enabled him to examine aspects of the service-learning program that led to these students changing their attitudes about the course content. More exogenous and purposeful sampling is vital to advancing service-learning research so as to understand the way student characteristics affect service-learning experiences, and they are largely missing from the studies reviewed for this paper.

**External Validity to Other Programs**

As Furco (2003) described, service-learning research is limited by the focus on individual, idiosyncratic programs rather than a systematic examination of the effects of different aspects of service-learning on the different types of students; most studies...
reviewed here follow this pattern. Of the 55 reviewed studies, 47 examine an individual service-learning course. These studies approach service-learning programs as monolithic entities without differentiating specific aspects of the programs, such as the types of reflection or differences in relationships with the served population, limiting the way results can be applied to other programs, whereas understanding component parts would produce results that speak more to programs that are similar in those specific ways. Currently, though, research is presented more in terms of individual program assessment, with little to no attention paid to the specific aspects of the program that may be contributing to the outcomes.

Hollis (2004) takes a more systematic approach, comparing outcomes from levels of support on reflection activities and amounts of preparation before service experiences. Eyler and Giles (1999) and Astin et al. (2000) include questions about aspects of the service-learning experience, including the amount and type of reflection, levels of faculty support, and amount of service conducted. Results from these studies can point to ways to make the programs more beneficial rather than simply making claims about the effectiveness of an entire course or program that would be impractical or impossible to implement wholesale on another campus.

Lack of Detail

Exacerbating the limitations on external validity in the body of research on service-learning as an intervention to encourage diversity outcomes is a lack of detail provided about the studied service-learning programs themselves. Even with a body of research made up mostly of the examinations of individual programs as described above, conclusions could begin to be drawn about what types of programs promoted each of the six diversity-related outcomes if enough detail was provided about those programs. In the sample of studies analyzed for this paper, however, insufficient detail about the specific experiences of students enrolled in the service-learning courses was provided.

An original goal of the current study was to identify the practices and conditions present in programs that led to each of the six diversity-related outcomes detailed above. This was impossible, however, because only six of the studies in the sample provided the level of detail about the programs studied to allow for these types of claims. The authors of these six studies provided the reader with information about the types of interactions students had with the served population, the ways that service was incorporated with course material, and other details to allow the reader a fuller understanding of the experiences of students in the studied program. For example, Baldwin, Buchanon, and Rudisill (2007) examined two service-learning programs in which students worked with children at minority-serving social service agencies. They went into extensive detail about both programs, including the ways students prepared for their service, the context for the service and the served population, the role of in-class time, and the way sessions between the students and served population progressed. This is the level of description of the program and students’ experiences necessary to draw conclusions about the most effective practices and how those practices encourage any of the six outcomes. Other studies that provided extensive detail about the service-learning experiences were Borden, 2007; Childs, Sepples, & Moody, 2003; Jakubowski, 2003; Smith, 2003; and Stachowski, Bodle, & Morrin, 2008. Unfortunately the vast majority of studies do not include this type of information, making it impossible to draw these conclusions from the existing research.

Trustworthiness of the Data

The reviewed studies exhibit a major limitation based on the type of data often used. Of the 44 studies in the sample with qualitative or mixed methods, 36 relied on data collected from students’ reflection journals, final papers, and other course assignments. These data were not created by the students to explicitly provide truthful accounts of their experiences, responses, and outcomes from the program; rather, results were derived from data created by study participants to receive grades. Trustworthiness of the data requires the assumption that students were less concerned about providing responses that they believed would result in high grades than in providing honest feedback about their responses and outcomes.

This is a particularly concerning assumption for studies that address diversity-focused programs, as college students are well-practiced in parroting responses about racism and inequality that represent what they believe instructors want to hear (Paoletti et al., 2007). Stanton (1990) expressed concern about this trustworthiness, and also suggested that without the right kinds of questions and supports in reflection this data likely only scratches the surface of the response to and outcomes from the experience: “Their learning outcomes are likely to be described vaguely with phrases such as, ‘I learned a lot;’ or ‘I got so much out of my experience’” (p. 185). The assumption that these course assignment data represent the true effect of the program is a difficult one to make.

Discussion

As students matriculate to colleges and universities, most are encountering diverse living and learning environments for the first time, and the world
they will enter and work in after graduation is becoming more diverse every year. Institutions employ significant resources in helping students achieve diversity outcomes, and researchers have shown that a wide range of these interventions are effective. Despite a long line of service-learning practitioners and scholars positioning service-learning as an intervention that, by design, can be particularly effective in encouraging diversity outcomes, service-learning has largely been missing in larger conversations in the literature about these interventions. In this paper, I reviewed studies that present diversity outcomes as a result of service-learning participation. Taken as a whole, these studies provide significant support for the claim that service-learning can promote the outcome, with six specific diversity-related outcomes found consistently across studies: tolerance of difference, stereotype confrontation, recognition of universality, interactions across difference, knowledge about the served population, and belief in the value of diversity.

Based on these findings, we can conclude that diversity outcomes are common from service-learning participation. This re-emphasizes the place that service-learning should have in the discourse on diversity outcomes in higher education. To continue ignoring the potential of service-learning is to ignore a practice that could make a real difference for students not currently being served or affected by other practices and present an incomplete picture of the efforts by institutions to educate students. Diversity researchers and higher education leaders should give serious consideration to service-learning when they are considering ways to help prepare students to work and live in a diverse world.

This conclusion, however, also illuminates the need to take the research being done in this vein further than it has gone to date. If we can demonstrate such strong evidence of the effect of service-learning participation on students’ diversity outcomes, to continue to demonstrate this connection is not sufficient for the future direction of research. Addressing limitations of the current body of research – a lack of theoretical models, sample selection, external validity to other programs, lack of detail, and the trustworthiness of data – will allow researchers to begin answering more fine-grained questions, investigating what aspects of service-learning experiences are most likely to lead to what types of diversity outcomes. This will provide an empirical basis to make claims about best practices and suggestions for educators, something very difficult to do from the studies analyzed for this paper.

As Eyler & Giles (1997) have pointed out, service-learning practitioners and researchers have different immediate goals and, as such, ask different types of questions. Most studies in this review were written by service-learning practitioners writing about their own service-learning classes and programs; the questions of practitioners dominate the research on service-learning and diversity outcomes. It is imperative for more general researchers to add to the research to service-learning as a diversity intervention and balance that equation, the way that they have to other types of interventions. This will help bring service-learning into discourse of ways to enhance diversity outcomes among students, providing educators and institutions with an additional tool in their “diversity toolbox” to reach students. Balancing published studies between a practitioner perspective and a researcher perspective would also allow for diversity outcomes to be examined across service-learning programs and perspectives. This also would allow the research to move beyond the body of disconnected studies to which Furco (2003) alluded and to empirically identify effective practices that practitioners could use when developing and refining service-learning programs to encourage students’ diversity outcomes.

References

(Asterisked citations comprised the analytic sample)


Holsapple


Author

MATTHEW A. HOLSAPPLE (mapple@umich.edu) is a doctoral candidate at the Center for the Study of Higher and Postsecondary Education at the University of Michigan. His research focuses on students’ moral and ethical development during college and the educational experiences and institutional characteristics that encourage that development. He holds master’s degrees from the University of Michigan and Purdue University in education research methods and college student affairs.