COMMUNITY SERVICE-LEARNING:

THE EFFECTS ON ADOLESCENTS' CIVIC ENGAGEMENT, ACADEMIC ACHIEVEMENT, AND PERSONAL DEVELOPMENT

by

Robert Lawrence Gandy

A Project Submitted to the Faculty of

The Evergreen State College

In Partial Fulfillment of the Requirements

for the degree

Master in Teaching

2009

This Project for the Master in Teaching Degree

by

Robert Lawrence Gandy

has been approved for

The Evergreen State College

by

_____

George Freeman, PhD.

Member of the Faculty

_____

June, 2009

ABSTRACT

This paper examines the potential of community service-learning to better engage adolescent students with their local communities, increase their academic achievement, and enhance their personal development. A historical overview illustrates how service and education have become intertwined, and the theoretical foundation of the community service-learning model is credited to Dewey, who advocated the values of experiential learning. A critical review of the literature suggests consistently increased levels of civic participation and social responsibility for youth participants in community service-learning. The quantitative results on the effects of community service-learning on academic achievement are less conclusive, but some qualitative research suggests cognitive growth in higher-order thinking skills. Personal development, exhibited through increased self-esteem and personal responsibility, and decreased delinquent behavior, is consistent in several studies' conclusions. Finally, attention to service design suggests that the teacher, duration and type of service, and level of student ownership significantly influence measured outcomes. Suggestions for further research are included, including stronger quantitative methodology with a quasi-experimental design, greater attention to the structure of the students' service experience, and longitudinal data on larger, more representative samples.

TABLE OF CONTENTS

TITLE PAGE.....i

APPROVAL PAGE.....ii

ABSTRACT.....iii

ACKNOWLEDGEMENTS.....vi

CHAPTER ONE: INTRODUCTION..... 1

Introduction..... 1

Rationale..... 1

Statement of Purpose..... 6

Definition of Terms..... 6

Limitations..... 7

Summary..... 8

CHAPTER TWO: HISTORICAL BACKGROUND..... 10

Introduction..... 10

The Expanding Role of Practical Education in Higher Learning..... 10

Industrial Education vs. Traditional Liberal Arts: The Education of the African-American Community..... 12

Dewey and the Foundations of Community Service-Learning..... 15

Social Workers and Government Lend Rise to the Movement..... 15

Shifting Epistemologies in 20th and 21st Century..... 18

Strengthening the Focus On K-12 Community Service-Learning..... 22

Summary..... 23

CHAPTER THREE: CRITICAL REVIEW OF THE LITERATURE..... 25

Introduction 25
Civic Engagement 26
Special Populations 55
Community Involvement and Social Responsibility 65
Academic Achievement 76
Personal Development 97
Service Design 107
Summary 121
CHAPTER FOUR: CONCLUSION 122
Introduction 122
Summary of Findings 124
Classroom Implications 134
Suggestions for Further Research 137
Conclusion 142
REFERENCES 145

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my dedicated faculty team for guiding me through this research process. Terry Ford lent my cohort her expertise in analyzing qualitative and quantitative research. Masao Sugiyama instructed me in statistics, and the tedious but important task of formatting. Finally, my thesis adviser, George Freeman, provided me with valuable resources, suggestions, and encouragement. Most importantly, he supplied a much needed dose of humor to stay sane when knee-deep in articles.

I also received much appreciated assistance from my colleagues, most notably Lori DeFoor. She and I spent a lovely summer's day reading and revising each other's work, and it was a pivotal moment in taking my work to closer to completion. I would like to thank Chalen Kelley for the valuable sharing of research materials on our topic. I cannot fail to also mention my silent advisor, James Brownley Horne, an MIT alum whose thesis served me time and time again as a tangible reminder of what a completed masters paper should look like.

I would like to thank my mother and father for allowing me the time and space to complete my most valuable work on this project during our precious summer visit—which quickly became a research retreat. Also, thank you to Emily Lardner, and my summer writing class, for reading several drafts of my first two chapters. Emily's teaching inspired me to become more experimental with style and punctuation in the final stages of my revisions. And lastly, thank you to Justin Hellier, for being attentive through all my many pages in the last round of editing.

CHAPTER ONE: INTRODUCTION

Introduction

One of the fundamental reasons children are educated through formal schooling in the U.S. is to create a strong, democratic society. The role of education to promote citizens with a disposition to act on knowledge in the pursuit of good can be traced back in Western thought to the work of Plato and Aristotle (Rocheleau, 2004). Modern philosophers such as Locke, Kant, Mill, and Rousseau continued to assert that education plays an essential role in forming positive values and citizenship; they saw education as necessary for "capable and sensible" civic participation (Rocheleau, p.3). Rousseau (1968, trans.) deemed the educational prerequisite for citizenship "the social contract," and it signifies the transition humans make from "a state of nature" to "civil society" (p. 20). Education—for Rousseau—was what allowed humans to evolve "from a stupid, limited animal" and be lifted to a "creature of intelligence" (p. 20). Du Bois (1903) also understood the power of education to develop the character necessary to access civil society, and he called upon it for his fellow African Americans. At the turn of the 20th Century, he warned a white culture of power, "If you do not lift them up, they will pull you down" (Du Bois, p. 156).

Rationale

Today society is confronted with the dilemma of how to educate and empower citizens that are engaged in democracy. The U.S. has seen the participation of its citizens in government decline; politics have become a "niche activity" particularly exclusive to youth (Campbell, 2000). Social philosophers

identify a “crisis of democracy” reflected in the widely held belief that politics are governed by money and powerful lobbies rather than the public good. Americans have fewer social connections with opportunities to engage in political and social debate than prior in American history (Putnam, 2000). This is a decline in what sociologists define as social capital—the “features of social organization such as networks, norms, and social trust that facilitate coordination and cooperation for mutual benefit” (Campbell, p. 21). The loss of a sense of community in American culture and the resulting alienation directly impact youth. Over the past 50 years the dominant U.S values have become materialism, individualism, and competitiveness (Speck & Hoppe, 2004), and the traditional classroom perpetuates these values, contributing to the rise of egoistic individualism in youth (Rocheleau, 2004).

Socialization scholars and educators are concerned about young people’s lack of knowledge, interest and engagement in political activities such as voting (Owen, 2000). Youth are perceived as being apathetic toward political institutions and parties; their feelings of national pride and patriotism are weaker than those of other age brackets. Statistics by the National Association of Secretaries of State cite that fewer than 20% of youth between the ages of 18 to 25 voted in the 1998 election (Van Benschoten, 2000). Political participation is seen as prerequisite for successful, democratic societies, and citizenship is a primary mission of public schools (Flanagan & Sherrod, 1998). As a result, local and national leaders are confronted with the task of providing youth with opportunities for success as students in the classroom, and citizens in society.

From a different perspective, some scholars prefer a more developmental approach to changing civic behavior, as opposed to top-down, structural, macrosocial studies of the problem. The developmental background of youth can be pivotal in forming their adult capacity to take part in the renewal and reform of civic society, and thus it becomes important to examine how civic identity is constructed at the individual level (Youniss, McLellan & Yates, 1997).

Scholars have employed Erikson's theoretical framework in attempting to explain how generational identity formation is a critical component in creating civically engaged, and socially-minded adolescents (Nakkula, 2003; Youniss & Yates, 1997). Erikson's work examined the ego identity—he believed this identity developed through one' s relationship with others. Adolescents need to learn how to perform and achieve cooperatively if they are to successfully enter a democratic culture as adults. Youth are searching for identifying ideologies, which is a means for them to simplify, organize, and make meaning of their experience of the world and their connection to it. Ideologies allow youth to identify commonly at a higher level, transcending their personal identity and sharing ideals with others. As youth become parts of a larger group their ego identity gains strength, and they develop the agency needed to collaboratively create a better world.

Bringing youth to reflect on political and moral ideologies during the developmental process allows them to construct identities that integrate ideological stances with political and moral outlooks connected to society's history and concern for its civic future (Youniss & Yates, 1997). Van Benschoten

(2000) argued that youth are forming their civic identity in new ways, volunteering, serving, and finding new ways of affecting social change besides participating in political structures.

In an attempt at creating civic reform and greater social responsibility in youth, education research and practice has looked to constructivist methods of teaching that require the social interaction of the individual with their environment. Dewey (1938) claimed that for youth to have an “educative” experience, learning must be characterized by interaction and continuity. Knowledge becomes relevant to children when they participate in their local community and their learning is connected to their interests and future. One example of a teaching and learning approach founded on constructivist theory is community service-learning.

Advocates hail community service-learning as a way to improve youth’s participation and education within society. Teens confront social issues such as pregnancy, violent crime, drug abuse, and dropping out; community service-learning is a way to combat these ills and empower youth to be positive, contributing members of society (Wade, 1997).

Community service-learning is also argued to enhance the cognitive development of the child. Rooted in the cognitive constructivist theory of Piaget (Singer & Revenson, 1996), Vygotsky (Wertsch, 1985), and Rogoff (2003), community service-learning is collaborative problem solving which strengthens mental structures. It requires that students interact with the outside world, and engages them in social dialogue that is crucial to their intellectual maturation

(Pritchard & Whitehead, 2004). As Rogoff asserted, “People develop as participants in cultural communities” (p. 3). Educational reformers see community service-learning as a way to help their students retain more knowledge and become life-long learners. Based on theories put forth by Dewey (1938) and Kolb advocates assert that direct experience and reflection are essential to effective learning (Wade, 1997). Community service-learning is thus seen as experiential education, with criteria that ensure it is an educative experience for the student.

While those in favor of community service-learning argue for its potential to improve students’ academic achievement and civic engagement, there is still doubt and a need for further research to assess its effects and viability. Without empirical evidence demonstrating a positive, lasting impact on youth, there is no solid ground to support the use of valuable class time on service work. In U.S. schooling, standards-based assessment is the dominant approach; teachers and schools are held accountable with high-stakes tests to demonstrate students’ mastery of content knowledge requirements. Community service-learning demands significant time and energy from students and educators, and if it cannot demonstrate that it improves student performance on state assessments, then conventional methods will be favored.

The prominence of community service-learning as a pedagogical tool lies in the hands of educators and researchers who can demonstrate whether it can be sustained and institutionalized in the U.S. educational system. Billig (2001) feared that the approach might become just a fad unless a research agenda is

implemented, and methods are improved to demonstrate community service-learning's efficacy in a K-12 setting. As of 2001, there was no significant body of research literature demonstrating enduring changes in student knowledge, attitudes, and behavior (Pritchard, 2001). However, in the last seven years there has been a concerted effort to refine and develop research on community service-learning in a K-12 setting to complement the extensive work that has been done in higher education. The focus of the most recent research in the field is on the effect of community service-learning on academic achievement, civic engagement, and social and personal development.

Statement of Purpose

This paper intends to analyze the current research in the field of community service-learning and determine its effects on adolescent students' civic engagement, academic achievement, and personal development. Additionally, the elements of service design will be examined to better understand how the aforementioned outcomes may be enhanced. The results of this research can then be used to help educators determine how community service-learning could best be prioritized and structured in their curricula.

Definition of Terms

The National Commission on Service Learning defines service-learning as "a learning teaching approach that integrates community-service with academic studies to enrich learning, teach civic responsibility, and strengthen communities" (Pritchard & Whitehead, 2004, p.4). The research literature varies in its use of community-service, service-learning, and community service-learning; in this

paper, the term community service-learning will be consistently used when not limited to a particular study, as it is the most inclusive label. This paper defines community service-learning as:

> "curriculum-based community service that integrates classroom instruction with community service activities. The service must: Be organized in relation to an academic course or curriculum; Have clearly stated learning objectives; Address real community needs in a sustained manner over a period of time; and assist students in drawing lessons from the service through regularly scheduled, organized reflection of critical analysis activities, such as classroom discussions, presentations, or directed writing." (Pritchard, 2001, p. 20)

The concept of civic engagement presented in this paper includes but extends beyond voting and volunteering. Civic engagement is about "helping citizens recognize their authority, learn[ing] the skills to create change, and organiz[ing] a base with others who share a common vision" (Van Benschoten, 2000, p. 3). It implies a sense of agency, and belief that one can make a difference in the world and has the social responsibility to do so (Youniss & Yates, 1997). This term will overlap with the terms community involvement and social responsibility.

Limitations

In attempting to make this work as relevant as possible to K-12 educators, the research presented here on community service-learning has been

limited to that conducted on samples of K-12 students. All but one seminal piece of research has been drawn from the past decade to ensure the analysis concentrates on the most recent and relevant findings. The greatest body of research done on the effects of community service-learning, however, is actually on sample populations at the college undergraduate level; the research methodology at the college level is more refined, and the academic dialogue on the subject more robust. Since the findings of this research are not necessarily transferable to K-12 students, this work has been omitted from the literature review.

One of the primary difficulties in establishing the efficacy of community service-learning in a K-12 environment is the shortage of research at this level with strong methodology. Many researchers were not consistent in their terminology and definitions of what constitutes community service-learning. Much of the literature that existed prior to 2001 is conceptual, theoretical, or opinion-oriented. There is an abundance of how-to guides and examples of service-projects, but there is a lack of empirical research done on the effects of community-service learning on K-12 student samples.

Summary

This chapter has provided a brief overview of how youth today are perceived in terms of their relationship and involvement with their communities and greater civic society. Schooling as an institution is idealized to empower citizens capable of continuing and strengthening a democratic society, but the behaviors of youth today suggest that education is not serving these ends. The

problem of declining involvement and disengagement of youth can be perceived either structurally or developmentally, and schools have begun seeking to develop new ways of helping students become more civically engaged and socially responsible, as well as increasing their academic achievement and promoting positive personal development. Community service-learning is one possible way to impact students in these areas, and the research documenting its effects in a K-12 setting is growing.

The integration of community service with academic curriculum is still a relatively new method in the history of education in the U.S. A historical overview of how service and education have come together for American youth will now be presented to contextualize the contemporary research literature presented in this paper.

CHAPTER TWO: HISTORICAL BACKGROUND

Introduction

While community service-learning is a new field in American education, service to one's community has a long-standing history beginning with the philosophical foundations of the young republic. De Tocqueville commented that voluntary associations to serve the public were "the principle of interest rightly understood," and that they were crucial to American vitality (Pritchard, 2001, p.4). Thomas Jefferson believed in basic, universal education for males to prepare them for the necessary conditions of citizenship. While his view of who should be educated was aristocratic, he assumed those who sought education did so to ultimately serve the public interest as civil servants and leaders (Rocheleau, 2004). Many Western philosophies of education saw the aim to be the "knowledge of fundamental moral principles and the development of a character with the disposition to act on these principles" (Rocheleau, p. 4).

The Expanding Role of Practical Education in Higher Learning

In the 19th century, American higher education began moving toward a more practical and accessible view of education for the farmers and laborers of the land. Until then, universities had been known as "conservators of knowledge and transmitters of mostly foreign culture" as a legacy of European ways (Zieren, 2004). The year 1857 marked the creation of the first agricultural colleges in the U.S., beginning in Michigan and spreading to Iowa, Pennsylvania, Maryland, and New York. These colleges blended book knowledge with mandates of manual labor, and were an effort to improve the farmer to be a citizen of democracy. The

Morrill Land Grant Movement was a federal act that donated the proceeds of land sales to public education for farmers and mechanics. It was the first national attempt to combine liberal arts with practical and vocational learning, and it defined a break with the past of traditional schooling. Instead of liberal arts colleges offering only the “discipline and furniture of the mind,” they now served the interests of the corporate state (Spring, 2005, p. 316). The Morrill Land Grant was a sign of a greater orientation to service in higher education, and this would directly benefit the industrial and agricultural sectors.

By 1860, Westpoint and the Rensselaer Institute were known for their practical training and applied instruction. International expositions such as Philadelphia Centennial Celebration and the Chicago World’s Fair of 1896 demonstrated the benefits of practical training and technical education through the innovations displayed by Germany and Russia. This had a sweeping effect on higher education in the U.S., and classroom training was combined with practical internships. Professional organizations of teachers, doctors, lawyers, social workers, and engineers began to lobby the legislature for more educational standards and on-the-job training in their professions (Zieren, 2004).

The view of education as gaining knowledge in preparation to serve or work began to change with the Progressive Era and hands-on education became more prevalent. Colleges and universities became much more prominent in community life, and curriculum offerings became more diverse. At the University of Wisconsin, President Charles Van Hise developed the “Wisconsin Idea,” which promoted the use of university expertise for the aims of reform and civic

improvement (Zieren, 2004, p. 30). Students and faculty learned through application of their knowledge to modern problems, and farmers could use extension programs to improve their practice. By the 1900s, education had become more democratic in its content, with courses expanding to include home economics, civics education, and citizenship courses. Vocational education and community colleges were created and "designed to connect more closely the worlds of work and education, and education with the community" (Zieren, p. 31).

Industrial Education vs. Traditional Liberal Arts: The Education of the African-American Community

While the purpose of education diversified in the U.S. to become more inclusive of Euro-Americans of varying socioeconomic backgrounds, the educational landscape became increasingly segregated for African-Americans. The shift in education toward vocational and practical experience was intended to unify and integrate White communities, but it purposefully excluded and divided communities of color. The Supreme Court ruled in 1896 that public schools could be segregated under the "separate but equal doctrine." Funds to segregated African-American schools became minimal, and the resistance by white society to offer any form of education to people of color was high (Spring, 2005). The effort to provide education to the African-American community rested chiefly on its own shoulders, financially and emotionally.

As the idea of practical education geared toward service in the industrial and agricultural sectors spread throughout higher education, the distinction between vocational education and traditional liberal arts curricula became

particularly divisive within the African-American community. The education many African-Americans pursued was practical: a commonly segregated, industrial education, based in skills that would help them negotiate contracts and measurements and become better workers (Spring, 2005). Their primary motivation was to improve their economic and political positions in Southern society. In the 1890s, this stirred a great debate over what type of education best served African-Americans, and the controversy was represented by the clashing philosophies of Booker T. Washington and W.E.B. Du Bois.

Washington was born into slavery and in his freedom had received an industrial education that emphasized occupational training, manual labor, better work habits, and moral development (Spring, 2005). As he became a leader of the education crusade for African-Americans, he compromised with white demands and advocated for segregated, industrial education. Washington felt this practical education would best help his community prove themselves economically, and this would eventually give them grounds to stand on for equality.

Du Bois came from a very different background, which explains his opposition to Washington, and his stance for traditional, liberal arts curricula. Born in the free state of Massachusetts, he received a PhD at Harvard, studied in Europe, and became a professor of sociology (Spring, 2005). Du Bois envisioned the purpose of education as the shaping of new leaders that would protect the social and political rights of the African-American community (Du Bois, 1903). This view of education as the breeding ground for social and political leadership

was similar to what Thomas Jefferson had promoted, only Du Bois was not looking to staff the white aristocracy that Jefferson had in mind. He hoped to instill a legacy of struggle for his people, and help them hold on to their discontent until true equality was granted to all African-American citizens.

The two leader's views on education could not have contrasted more, and they created a lasting dichotomy between practical, hands-on skills and liberal arts that community service-learning works to dismantle. Washington scorned the teachers of the North, with their liberal arts education in which he saw only false promises for his community (Spring, 2005). Industrial education was practical, occupational training that would benefit all immediately, and classical studies only served to disappoint and develop useless vanity. Du Bois, on the other hand, thought all African-American educators should be trained in the traditional schools of the North, to be empowered and think critically, and then take their knowledge to lift those of the South (Du Bois, 1903). The tension between the two types of education they each fought for remained until schools were finally desegregated beginning in the 1960s. The movement for community service-learning today recognizes the value of both forms of education that Washington and Du Bois believed in, and does not seek to favor one over the other, but rather it works to integrate them. Advocates of community service-learning posit that the critical thinking skills and civic concern Du Bois sought to cultivate could be enhanced by working more directly with communities, through hands-on, experiential learning.

Dewey and the Foundations of Community Service-Learning

Dewey was one of the most influential voices on education in the Progressive Era, and his philosophy formed the theoretical foundation for modern-day community service-learning. His work emphasized the participatory nature of education, and he was an early advocate for "participatory democracy" (Rocheleau, 2004, p. 7). Dewey wanted to see active citizen participation in identifying social problems and devising solutions. In his eyes, the importance of learning was chiefly its benefit to the community as a whole, not simply the advancement of the individual. While Dewey never used the term community service-learning, he established the basis for the model, and helped shape criteria to facilitate cognitive development through problem-solving. His work influenced a new generation of theorists and educators, and led to the founding of the Dewey School at the University of Chicago. William Kilpatrick was a follower of Dewey's work, and he proposed the "project method" of education, an early predecessor of the community service-learning model. Kilpatrick argued that learning should take place outside the school and involve efforts to meet real community needs (Conrad & Hedin, 1991). This work was expanded upon in 1936 with Hanna's *Youth Serves the Community*, which applied progressive, project-based education to social problems (Rocheleau).

Social Workers and Government Lend Rise to the Movement

The field of social work in higher education also contributed dramatically to the foundations of the community service-learning model, and since the Progressive Era it has kept service-learning alive. Edward Lindeman was a key

figure in this movement between the 1920s and 1930s as a social work educator and reformer. Known as the “democratic man,” he called for an effort by society to ensure that adults possessed the requisite civic skills to create an ideal society (Zieren, 2004, p.32). He adopted Dewey’s philosophy and proposed community service-learning as one of bases of the social work profession. The social work field came to use a “field practicum” that was a guided learning experience with service to the community. A strong civic ideology developed to support community service-learning, and the social work profession has provided a model to society that has developed over time.

Between the 1930s and the 1970s, the U.S. federal and state governments began creating opportunities for youth to serve their communities. An influential group of social workers in the 1930s led by Harry Hopkins and Aubrey Williams were an integral part of the New Deal. Their influence on Eleanor Roosevelt led her to convince President Roosevelt to create a massive employment program for youth (Zieren, 2004). President Roosevelt initiated the largest community service-learning project at that time, creating the National Youth Administration (NYA) by presidential executive order in 1935. The NYA established jobs and educational opportunities for youth ages 16 to 25 in vocational and academic settings. The program offered youth two options, the Works Project Program and the Student Aid Program. In the Works Project youth were given full-time jobs in industrial and community settings, with goal of skill attainment and meeting the needs of the community. Student Aid operated similarly to what is known as work-study today, and its focus was working on

community problems. The purpose of NYA was “to teach by example ‘the practice, responsibilities and rewards of citizenship’” (Zieren, p. 35). It exposed a new generation of young people to social work’s concept of service and community betterment.

This work continued when the Civilian Conservation Corps (CCC) was established by Harold Ickes in the Department of the Interior, which operated from 1933 to 1942 (Wade, 1997; Zieren, 2004). With the urging of President Roosevelt, the organization was built around the ideals of service to community and modeled closely after the NYA. It was an attempt to place youth in an environment where they could learn and serve simultaneously. The NYA and CCC enrolled over 1.4 million youth between 1935 and 1939 and had an even greater impact than higher education at that time on creating models of service and linking education with the practical application of skills (Zieren). President Kennedy established the Peace Corps as an international service program in 1961, and President Johnson’s “War on Poverty” led to the creation of Volunteers in Service to America (VISTA) in 1964. The Youth Conservation Corps began in 1970, and from that program many states developed programs of their own, such as Washington, California and Vermont. The initiatives of the Federal and State Governments, and the strong influence of social work’s aims of public advocacy and community organizing were central forces in shaping what community service-learning is today.

Shifting Epistemologies in 20th and 21st Century

A cultural shift in views on education that occurred from 1940 to 1985 challenged the potential of schools to be part of this growing service movement, and Dewey's progressive ideas went out of vogue. Theories of logical positivism and objectivist epistemologies developed, and the value of knowledge for foreign competition became great (Rocheleau, 2004). A defining moment was the launch of Sputnik in 1957 that set the U.S. into a competitive frenzy. Hands-on, progressive approaches to education were abandoned in favor of more scientifically rigorous education (Spring, 2005). History repeated itself in the 1980s after the rise of Japan as an economic power. U.S. education was scrutinized for weaknesses, and U.S. Secretary of Education William Bennett declared unproven teaching methods caused lower test scores. The 1983 Nation at Risk Report condemned the state of education, and more conservative, mechanistic approaches to education were taken.

Academia in the late 20th century witnessed an epistemological and philosophical turn through the revival of Progressivism and Dewey's work. As a result of the tide of poststructuralism in academic thought, the objective authority of facts was questioned, and there was a renewed interest in the active interpretation of knowledge, practical learning goals, and the social context of learning. Academic research in education was influenced by phenomenological and pragmatic conceptions of the role of the individual and society (Rocheleau, 2004). Dewey's instrumentalist conception of inquiry was researched and expanded upon, most notably by Kolb in 1984.

Kolb established a model of experiential learning that has provided one of the strongest theoretical foundations for community service-learning. His model integrated the work of Piaget, Lewin, and Dewey with the core concept that "reflection transforms experience into new and usable understanding" (Pritchard & Whitehead, 2004, p. 11). In summary, the model operates as a cycle moving the learner between quadrants defined by concrete experience, reflective observation, abstract conceptualization, and active experimentation. The model is congruent with the cognitive goals of community service-learning, and has been used in planning and setting criteria for successful and educative service experiences.

Political philosophy at the end of the 20th century became more concerned with the status of democracy and citizenship. There was a call for the rejuvenation of participatory democracy, and a desire to move away from the egoism of the 1980s that had been career-centered and fostered political and social alienation (Rocheleau, 2004). Economic growth had led to a decline in concern for others, and many social and political philosophers believed social solidarity and commitment to the common good was necessary for the civic health of society. Community service-learning emerged as a way to combat the individualism and competitiveness present in schooling, and it was seen as a way to promote cooperative learning with a communal goal of benefit for one's community. Around the mid-80s, networking among various service areas began to increase dramatically and community service-learning spread in K-12 schools and higher education. By the late 80s, the governors of 20 states had started

initiatives for statewide youth service programs (Wade, 1997). As initiatives were launched, private corporations such as Ford, Mott, and W.K. Kellogg began contributing millions and their support helped build the service movement.

The federal government resumed its initial role as promoter of service through several important legislative acts in service history. In 1989, President Bush founded the Office of National Service and Points of Light Foundation to promote volunteerism (Wade, 1997). Congress followed this with the passing of the National and Community Service Act of 1990 that created the Commission on National and Community Service. The commission provides program funds, training, and technical assistance to states and communities for service opportunities. President Bush and Congress approved $73 million dollars for program activities in 1992, and again in 1993. Over a quarter of these funds went to community service-learning in the K-12 and higher education setting. In 1993, with strong bipartisan support, President Clinton and Congress passed the National and Community Service Trust Act. This legislation established the AmeriCorps program, which was a domestic service program for youth that created opportunities for further education and hands-on development in areas of strong community need. The act also increased funds allotted to states for K-12 community service-learning, and incorporated VISTA and other federal volunteer programs under the Corporation for National Service.

With the amount of federal, state, and private support for community service-learning, academic interest and research began to spark in the 1980s. Within education, the field of community service-learning was a “new pedagogy”

and by the mid-90s the research literature had exploded (Speck & Hoppe, 2004, p. vii). The work of Conrad and Hedin was a pioneer study in the field, and now stands as a seminal piece of the research literature (Shumer, 2005). Their work examined 27 experiential youth programs, including community service, and documented some of the earliest findings on community service-learning's impact on academic and social outcomes. Conrad and Hedin concluded that experiential learning had the potential to improve students' grades, and help students become more socially responsible, with greater moral reasoning (Conrad and Hedin, 1982). On service design, they found a reflective seminar, student choice, positive adult interaction, intense service hours, and a program duration of many months facilitated the greatest cognitive effects. The *Michigan Journal of Community Service Learning* was first published in 1994, and this marked the beginning of systematic research specifically on community service-learning. As service and learning became more connected, research literature began hyphenating the two, with the hyphen serving as a "symbolic nexus that linked inextricably praxis and theory so that they are no longer two separate activities but symbiotic" (Speck & Hoppe, p. viii). The social perception that the American community had been fractured fueled the belief and interest in community service-learning as a remedy. The academic community generated intense interest and some began to see community service-learning as capable of producing a generation of citizens to restore community. However, the body of community service-learning research is not monolithic, and it has become quite controversial with multiple theoretical strains (Speck & Hoppe). The varying

theories of community service-learning include the philanthropic, civic, and communitarian models and these are held in tension by researchers. This paper will focus its analysis on the civic model of community service-learning, as it is specifically structured to develop participants' civic involvement.

Strengthening the Focus On K-12 Community Service-Learning

With the beginning of the 21st century, community service-learning researchers have taken on the task of refining and strengthening the data on K-12 schools. There has been an increasingly available amount of literature on community service-learning in higher education, with a well developed research agenda. The academic community is working to complement the body of work in higher education, and fill the void for K-12 schools. Most of the literature up until 2000 consisted of program evaluations, descriptions, or anecdotes, and no clear research agenda had been defined (Billig & Furco, 2001). The questions raised by policy makers and educators are just beginning to be addressed: How does community service-learning impact students' academic achievement as measured by state accountability systems? What is the impact on students' civic engagement? In July 2001, a group of researchers convened to begin addressing these questions systematically. They discussed the state of research in K-12 community service-learning and generated a research agenda for the decade modeled after the literature on higher education. The direction of their research is focused on the impact of community service-learning on academic achievement, civic engagement, and personal and social development.

Researchers believe that there is a rich body of social constructivist learning theory to inform their work (Billig & Furco, 2001). The challenge is developing hypotheses and seeking funding to continue the work. The research documenting academic achievement and civic engagement does not cohesively point to one conclusion. There is ambiguity in the use of terms and theoretical frameworks employed, and therefore room for community service-learning research methodology to be improved.

Summary

This chapter has offered an overview of how ideas of service and different philosophies of education have evolved in U.S. history. Beginning with the philosophical foundations of service in early America, service to others was presented as an integral component of U.S. education. The expanding role of practical education in higher learning signified the need for Americans of multiple social classes and occupations to have access to schooling. The educational history of the African-American community demonstrates that need, however, was not universally met. The contrasting philosophies of industrial education and traditional liberal arts curricula adopted by leaders of the African-American community led to conflict over how to best serve itself, and signified the greater question of how all would be better served and empowered by schooling.

Dewey's theoretical work on experiential education contributed to the foundations of community service-learning today. His work helped establish criteria by which the movement could build upon educative experiences, and it inspired those after him to construct models of learning based on experience and

reflection. Social workers lent rise to the community service-learning movement by utilizing Dewey's work, establishing the field practicum, and influencing the federal government to prioritize youth service programs. National legislation brought greater attention the service movement, and also attracted the support of private donors and organizations to help it gain momentum. Shifting epistemologies and waning political participation by young U.S. citizens led to the adoption of community service-learning in schools as a way to help strengthen civic society. As this work grew in the 1990s, it captured the attention of academia. Institutional literature and research initially focused on the effects of community service-learning within higher education, and currently there is work being done to strengthen the focus on K-12 community service-learning. The following chapter examines the most recent directions research has taken, and then illustrates how K-12 research in community service-learning is a growing and vital field of inquiry.

CHAPTER THREE: CRITICAL REVIEW OF THE LITERATURE

Introduction

The first chapter discussed the problem of today's youth, alienated from civic society, with declining interest, engagement, and participation in the democratic process. Erikson's work around adolescents' identify formation suggested that adolescence is a crucial window of development for youth to establish a sociopolitical identity and make a lasting connection to the community around them (Nakkula, 2003; Youniss & Yates, 1997). The educational theory of Dewey (1938) and Kolb (Pritchard & Whitehead, 2004) emphasized the importance of experiential education to create meaningful, educative experiences for youth in schooling. Community service-learning is argued by scholars and educators to be an experiential, constructivist pedagogy that encourages youth to become more civically engaged, while enhancing their cognitive development.

Chapter Two explored how a culture of community involvement and social responsibility has been cultivated through education over time. It tracked the rise of practical education in higher learning, and the tension that this evolution created between liberal arts curriculum and more experiential forms of learning, such as industrial and agricultural education. Over the last two decades, legislation by federal and state governments has established and promoted national and community service programs in U.S. schools, creating a swell of academic interest and research.

Chapter Three reviews the most recent research surrounding community service-learning in secondary schools, and is divided into four

sections: civic engagement, academic achievement, personal development, and service design. In each section, the research is summarized and analyzed based on the methodology and conclusions of each article.

Civic Engagement

The initial 12 studies of this chapter analyze the leading edge of community service-learning research: civic engagement. Scholars, educators, and politicians have heavily promoted community service-learning in response to the fledgling political participation of young adults, extolling it as a pedagogy that can alleviate the problem. The following research evaluates whether claims of improved civic outcomes from adolescent experiences of service can be supported empirically. Rutter and Newmann (1989) begin this section where the seminal authors Conrad and Hedin (1981) left off in Chapter Two; they questioned the power of service learning, a form of experiential education, to increase youth's commitment to civic participation. This section includes studies that measure a variety of civic outcomes, often referred to as civic engagement, commitment, or participation, and the research attends to both attitudinal and behavioral outcomes. Although some researchers also published findings regarding academic and personal outcomes, they will reappear with more detail in later sections.

As the research linking service with civic engagement grew more refined, Perry and Katula (2001), Riedel (2002), and Billig, Root, and Jesse (2005) increasingly discussed and controlled for interaction variables such as the background characteristics of the participant, the type of service program, and

the school, home, and neighborhood context of service. The section continues with research that focuses on the needs of special student populations, such as gifted adolescents studied by Terry (2000) and Lee, Olszewski-Kubilius, Donahue, and Weimholt (2007), as well as specific cultural communities, discussed by Yamauchi, Billig, Meyer, and Hofschire (2006). Finally, this section concludes with research documenting service's effect on the civic aspects of community involvement and social responsibility with research by Metz and Youniss (2003), McGuire and Gamble (2005), and Halfacre, Chessin, and Chambless (2006).

Rutter and Newman's (1989) seminal research led the inquiry into community service-learning's impact on the civic engagement of adolescent youth. The authors evaluated eight community-service programs they designated as exemplary and in a mixed-methods, nonequivalent control group design examined the programs' impact on students' civic responsibility (N = 314), personal development, and developmental opportunities.

The authors used a questionnaire to measure changes in students' sense of civic responsibility. Data was collected from treatment groups (n = 155) and comparison groups (n = 159) at the beginning and end of service, and the questions examined social and political items such as sense of school and non-school responsibility, social competence, political efficacy, future social affiliation, and anticipated political participation. None of the changes in mean scores were statistically significant.

To assess the effects of service in the area of personal development, the authors selected four students randomly in each program and interviewed this sample (n = 32) with open-ended questions on why they decided to enroll in voluntary service and what they expected to learn, before and after service. Students' responses were tape-recorded and coded into five categories: acquisition and pursuit of social relationships; personal growth and development; acquisition of useful skills and knowledge; community awareness and involvement; and vocational experience. Social relationships and personal growth were the most frequently noted effects of service; while only 11 students had anticipated growth in these areas, all but 5 students indicated productive social relationships being the most important thing gained from their experience.

Rutter and Newman (1989) measured the developmental opportunities perceived by the youth by asking them how often they experienced activities and roles considered to promote adolescent development. These experiences ranged from making difficult judgments and accepting blame to examining personal values and being exposed to new ideas. Students responded with the frequency they encountered these opportunities in school classes, family life, extracurricular activities, work, and community-service programs. The mean score on the scale used was highest for community service programs (M = 3.88; SD = .63). Mean scores of the treatment group (M = 3.58; SD = .60) were higher than the control group (M = 3.50; SD = .59) in each of the other settings, with the greatest mean difference found in school class opportunities.

Finally, eight students were selected from each school program to compare their fieldwork experiences with students of other programs. Newman and Rutter (1989) discovered commonalities in student discussion, such as "being treated more like adults…being given more responsibility, having more freedom, receiving more respect, and being regarded as responsible persons" (p. 372). Students were more enthusiastic about their service programs than their conventional class work, and attributed this to their enjoyment of active rather than passive learning.

The study's history and maturation was strengthened with the use of pre- and post-data, on both a treatment and control group. There also appears to be triangulation through a mix of quantitative and qualitative data. The external validity of this study was greatly reduced however, by the authors' omission of any sociodemographic details on the samples and schools studied. The criteria for choosing "exemplary" programs were not provided, and the authors made little attempt at randomization. All participants had elected to do community service, and thus the effects on students without a predisposition towards service is unexamined. The authors did not detail the sources of their instrumentation, nor did they report its internal consistency. There were no details on how the data for the qualitative research was collected, or mention of member-checking and acknowledgement of self as instrument. Rutter and Newmann (1989) acknowledged none of these limitations, and as a result, the credibility of the findings is weakened. Despite the critique of this particular study, the authors'

research is frequently cited in the research literature, and they are considered one of the most influential voices in the field (Shumer, 2005).

Another pair of prominent service-learning researchers are Yates and Youniss (1998), whose findings concerning sociopolitical identify formation were influential for later research to come. In a case study, they examined how the civic and extracurricular activities of high school juniors from a Catholic school (*N* = 160) connected to a sense of political identity in adulthood. They found that service-learning encouraged youth to think about their political role in society and become more civically engaged with government and policy. The participants were predominantly African-American, from middle and lower middle-class family backgrounds, and the balance of female to male was approximately equal.

An investigation of a year-long service learning program through a mandatory junior-year religion course on social justice provided the data for this study. As part of the course, students served four times at the same downtown soup kitchen for the homeless. The participants completed questionnaires during the first week of class and at final exams, and they wrote essays as part of course curriculum, which the authors collected after each quarterly visit to the soup kitchen. The authors ran quarterly discussion sessions with small groups (*n* = 10-15), in which students exchanged ideas about their experiences, and they observed participants in the classroom (40 hours) and at the soup kitchen (104 hours).

Analysis of the essays and discussion groups suggested that service encouraged students to reflect on societal problems and their own sense of

responsibility, and that these reflections became increasingly broader and complex over a one-year period. When students felt sadness, good about helping, and anger, they were more likely to express a commitment to continued service after the program. The authors found the effects of the program were an essential interaction of the service activity, class time, and peer discussions. Writing and discussion initiated reflection about service activities in terms of civic and political issues, which in turn, prompted them to consider their own civic participation. Yates and Youniss's (1998) content analysis of essays and discussions exhibited students' reflection "…on their role in enacting social change and touched upon politically related questions regarding the impact of personal initiative, government spending, and government responsibility" (p. 500). They found that 22% of the students (*n* = 36) reflected on civic responsibility and government policies, and three out of four quarterly group discussion sessions, each with different students, included these themes. Individual service experiences stimulated students to also consider the role of government policy and the individual in enacting change, and students discussed the relationship between critical thought and action. They also discussed the limitations of individual initiative and political action, and students demonstrated an awareness of the differences between individual and government action.

A strength of this study was the authors' clear acknowledgement of their theoretical position, which framed their research on political socialization with Erikson's identity development theory. They also worked with a large sample size of African-American youth of varying socioeconomic statuses, which is a valuable

contribution to the research literature that often overlooks large, low-income communities of color. The sample is from a private Catholic school however, which makes the research less transferable to the public school environment. The authors expressed that the inclusion of a control group could have strengthened the credibility of their findings. Interestingly, they were concerned that the mandatory nature of the program and its strong connection to the course curriculum limited the findings; however, this simply eliminated the confounding variables associated with student's self-selection and ensured the program met the community service-learning criteria established by this paper. The authors would have been better off to discuss the lack of an outside evaluator for the qualitative data, which is a greater limitation on the confirmability of their findings. Other unmentioned weaknesses include no discussion of member-checking or acknowledgement of self as instruments in their analysis, and the authors did not explain their procedure for selecting the research sample.

Kahne and Sporte's (2007) research agenda closely aligns with the Yates and Youniss's (1998) work. With the final publication and peer-review of their research still in progress, they examined civic curriculum's effect on low-income, students of color's commitment to civic participation. The authors pulled data from 2003 and 2005 surveys for the Consortium on Chicago School Research for a correlational study, and found a substantial impact in their findings. Their sample included 4,057 students from 52 schools in Chicago Public Schools (CPS), for which data was collected during their freshmen and junior years. The authors only included students who had taken the survey at both points, and had

remained in the same school, to establish longitudinal data with school-level effects.

To analyze the survey data, the authors used single-item measures on a four-point scale, and they constructed multi-item measures with Rasch analysis and modeling. Instead of summing individual item responses, Rasch modeling puts all items on a hierarchical scale and places suites of items on the same scale according to the degree and likelihood that they were “endorsed” by respondents. This model allowed the authors to measure the latent outcome variable of “commitment to civic participation” with a five-item measure (α = .73). From the survey data and CPS records, they extracted predictor variables such as classroom and school characteristics (α = .80 - .84), extracurricular activities, family and parent contexts (α = .73), and demographic and achievement values. An individual item assessed students’ service learning experience, outside of the broader measure of classroom civic learning opportunities (α = .74). Finally, the authors considered the impact of their prior commitment to civic participation on later levels by comparing the 2005 data with the 2003 reports. Data was analyzed using hierarchical modeling to test the significance of individual and group characteristics on the outcome variable in four models: the first stage controlled only for individual demographic characteristics; the second added indicators of family and parental context; the third added educational contexts and practices, including extracurricular activities; and the fourth included all variables with the added control for prior commitment to civic participation. The

authors provided the standardized coefficients for each model, and calculated the effect sizes by dividing the coefficient by the standard deviation of the outcome.

The results of the fourth model, controlling for all predictor variables, indicated that classroom civic learning opportunities as a group measure had a significant impact on commitment to civic participation with the highest effect size (.41; $p < .001$), and service learning opportunities had the second highest effect size (.26; $p < .001$). A prior commitment to civic participation (.19; $p < .001$), parental discussion of current events and politics at home (.12; $p < .001$), and neighborhood social capital (.14; $p < .001$) were equally significant, but with smaller effect sizes. Differences in demographic characteristics had no significant effect.

The authors claimed the large sample size and their control for students' prior commitment to civic participation as leading strengths of their work. While there was an overrepresentation of low income students of color in their sample, they felt it was beneficial to focus on this population, whose political voice is often marginalized, for generalization of the findings to other urban contexts. They acknowledge that beyond urban locales, the research does not have high external validity. The authors clearly stated their conceptual framework and personal bias in pursuing effects on civic commitment, influenced by Dewey's (1916) philosophy of experiential education and the importance of sociopolitical identity formation in late adolescence as discussed by Erikson. They employed a similar four-stage hierarchical analysis as Hoffman and Xu (2002), which thoroughly accounted for predictor variables, and they improved upon this work

by accounting for neighborhood characteristics such as social capital, which proved to have a significant effect. Key factors limiting the internal validity of Kahne and Sporte's (2007) work were the reliability and design of the data measures. Most had fair reliability scores, although some measures were in the .73 - .74 range. More importantly the authors note that many items, including one as important as the indicator of service-learning experience, are single items and therefore not desirable. No valuable details are offered about the service type, duration, frequency, and so forth, to allow more precise comparison with other research findings. Additionally, the authors lamented that they were unable to nest data at the classroom level, due to the limits of the database and the confounding variable of civic learning in multiple subjects.

The following study collected data at the classroom level, which is the context of most civic learning that uses a service-learning methodology. In a nonequivalent control group design with mixed methods, Lakin and Mahoney (2006) researched the outcomes of civic responsibility and intention to be involved in community action for students in an urban, school-based community service program. The participants (N = 60) were ages ten to thirteen, and approximately 40% were Caucasian, 30% biracial, 20% African-American, and 10% Hispanic. Ninety percent of the school received free or subsidized lunch.

The authors randomly assigned three sixth-grade classes on a classroom-wide basis to either the treatment (n = 40) or control (n = 20) condition during the first week of classes. They implemented the program separately, but concurrently, for the two intervention classes chosen. The process experience of

these two classes was assessed with a quantitative measure administered to program participants as a posttest, as well as qualitative feedback obtained from the participants in a classroom discussion following the completion of the youth service project. All three classes completed assessment surveys prior to the program implementation and following program completion to provide preliminary data on program effects. Quantitative assessment of the process experience used the Learner Empowerment Measure (α = .90), which was adapted for the age group. Ongoing feedback was collected as well as surveys, and discussions and reflections were held and recorded for qualitative data. The authors constructed a five-item measure (α = .83) to assess the construct of civic responsibility, and used questions to determine how important it was to help others, understand others, and contribute to society. Additionally, they added two questions to assess how important respondents felt that their actions and those of others are to the future of their neighborhood and the world. The author measured the participants' intent to be involved in future community action with a three questions they created for the study (α = .72). They measured empathy toward others with eighteen items from the Index of Empathy for Children and Adolescents (α = .72). Lakin and Mahoney (2006) used *t*-tests and Chi-squared tests to evaluate both the demographic compatibility of the intervention and control groups, and the representativeness of the sample. They used three mixed design ANOVAs to control for changes in target variables within subjects across time, and differences in the treatment versus control group.

The qualitative findings indicated that the program was successfully experienced as empowering and promoting a sense of community for its participants. It helped students feel a sense of accomplishment and be more connected to one another. Participation in the program significantly affected the youth's intent to be involved in future community action, F (42, 1) = 4.3, p = .046, n^2 = .09, as well as increased empathy toward others, F (42, 1) = 10.1, p = .00, n^2 = .20. Paired sample t-tests also indicated service programs increased empathy, t (28, 1) = 4.2, p = .00, and intention to be involved in future community action, t (28, 1) = 1.5, p = .16. However, involvement did not significantly increase participants' sense of civic responsibility.

While much of Lakin and Mahoney's (2006) work demonstrated strong research methods, the generalizability of the findings is difficult with a small sample size. The instruments only measured self-reported attitudinal data, and did not measure behavioral changes. The scope of the research was also very limited, due to a treatment period of only ten weeks. The testing measures for "intent to be involved with future community" and "empathy" had low reliability scores of .72, and therefore are not strong indicators of the treatment's effects on those variables.

Moving from the classroom to the larger school system, the impact of service learning is seen as impacting civic engagement as well as other variables such as academic achievement, discussed in the next section. Kielsmeier, Scales, Roehikepartain, and Neal (2004) surveyed K-12 school officials (N = 1,799) with a pre-experimental, static group comparison. Their national study

focused on the scope and nature of service-learning in U.S public schools. Using mail and follow-up telephone interviews based on a 1999 US Dept. of Education survey, with an initial mailing in January, 2004 and a follow-up phone interview in February, 2004, the researchers' final sample represented a 91% participation rate.

Sixty-one percent responded that service-learning had a very positive impact on their students' citizenships/civic engagement and 31% responded it was somewhat positive. On academic achievement, 32% responded that there was a very positive impact, and 51% responded a somewhat positive impact. Schools with higher poverty levels reported a higher percentage of very positive impact on the areas of academic achievement and school engagement.

These results may reflect wishful thinking on the part of the principals. This study failed to provide any statistical data used in calculating percentages. The authors provided no reliability data on the survey instrument, no information on how they selected participants, and there was no indication that the sampled population represents a diversity of ethnic and socioeconomic groups. The principals' impressions may not reflect the actual responses by students or their lived experience, and the authors provided no data on how the principals made the distinction of "very positive," and "somewhat positive" in their survey responses. The large sample size suggests that the sample may be representative despite the absence of data to substantiate this possibility.

Perry and Katula (2001) also worked at the broader school system level, and reviewed 37 empirical studies to determine whether there was a connection

between service and citizenship, and, if so, by what set of causal processes. Thirteen of the studies worked with participants in a K-12 setting, and the authors focused their examination on three mechanisms of how service affects the individual to bring about change: intellectual stimulation, socialization, and practice.

The authors developed a set of search terms, and used a limited set of databases to come up with their records including: Academic Search Elite, Book Where, Dissertation Abstracts Intl., Eric, Govt. Documents, PAIS, IPSA, Socio File, and Psyc Info. The majority of the records were from ERIC. They merged the databases to create a master database, and within this, a subset of research was collected relating directly to the relationship of service and citizenship. Perry and Katula (2001) culled 219 records in this process, and they examined the research according to a theory of change or logic model that consists of five sets of variables: antecedents, attributes of service, attributes of server, individual changes, and institutions. Finally, the authors summarized the 37 research studies in table form, and identified six categories of outcomes: citizenship-related cognitive understanding, citizenship attitudes, citizenship skills, institutional change, philanthropic and civic behaviors, and political behavior.

Perry and Katula's (2001) review of the literature found that most of the research focused on the citizenship-related attitudinal outcomes of service, usually measuring civic or social responsibility. There was an even division between the positive and null findings correlating service-learning with the development of higher civic responsibility. The authors commented that the small

number of empirical studies on the effects of service-learning prevent substantial generalizations from being made, but they offered the following summary: service appears to influence favorably citizenship-related cognitive understanding; service and volunteering appear to positively influence later giving and volunteering; the type of service that produces the most consistent positive results is service-learning. Researchers in the field of service-learning are divided over how important the attributes of service and the server are in determining its effects on citizen development.

Perry and Katula (2001) suggested that these factors surrounding service might be important moderators of attitudinal outcomes, and warrant more research. Additionally, research about citizenship skills and behaviors, particularly political behavior, has largely been neglected in studies of service; this makes it difficult to assess the outcome of service. Attitudinal outcomes—such as self-esteem and civic responsibility—are most easily measured, and thus the psychological effects of service have dominated the literature. But with the need for more active citizenship to combat political apathy, more attention to political behavior as an outcome of service would be worthwhile.

The following five research studies are more attentive to the characteristics of server and service that Perry and Katula (2001) discussed as potential moderators of civic outcomes. Morgan and Streb (2001) sought to elucidate how student voice in the planning and design of service projects might enhance their civic outcomes. They conducted a one group pretest/posttest study to determine the impact of community service-learning on adolescents' (N = 210)

self-concept and political engagement. They reported that in all cases there was significant support for the positive effects of service-learning on youth leadership in each of the three domains of civic attitudes and orientations—when students had voice in the project. The authors selected participants evenly from each high school grade level, and also included a small number of eighth grade students. Three-quarters of the sample was female, and 80% were Caucasian, 17% African-American, and the remaining 3% Asian or Pacific-Islander.

The Indiana Dept. of Education administered surveys to students in ten classrooms in ten different schools during the 1997-98 school year, before and after the students' service learning project. Each site received funding to support their projects, and the authors selected schools that submitted an application and described their plan to implement a service-learning project. Each school individually designed a plan, and their goal was to learn more about service-learning. The authors pooled the responses found that student voice and ownership, and authentic, significant learning as identified by Dewey (1938) and Beyer (1996) emerged as major themes.

Self-reported student voice was the key independent variable in the study, which Morgan and Streb (2001) based on four indicators (Conrad & Hedin, 1985; Melchior, 1998). Students rated statements based on a four-point scale and statements included: "I had real responsibilities; I had challenging tasks; I helped plan the project; I made important decisions" (p. 160). The authors examined the changes in students' attitudes in relation to level of student voice by using the average level of voice reported by the whole class. This put the focus on the

amount of voice the teacher allowed the class, as opposed to the feelings of one individual.

Responses about political engagement provided data on students' political attentiveness using a series of four questions based on questions in the National Election Studies as well as the Abt Report. Morgan and Streb (2001) also studied students' interest in social action, and whether service broke down negative attitudes about the elderly and disabled using a modified Abt Report. To control for the effects of student involvement in a variety of classes, the authors used a robust variance estimator that corrected the standard errors in the presence of heteroskedasticity—or differing variances—and allowed them to relax the assumption of independence within classes. Other variables commonly examined in community service-learning literature are the amount of time spent in service, and the amount of reflection structured around the service. The study controlled for these variables. To ensure that the positive effects of community service-learning were not limited by race, class, and gender, the authors included interaction terms between each of these individual characteristic variables and the level of student voice.

The results demonstrated that students' levels of efficacy and belief in their own competence increased with high levels of voice in service-learning programs. For each one-unit increase in student voice in the project, student efficacy increased .34 points on a 5-point Likert scale, and competence increased .49, as indicated by a t^2 coefficient, ($p < .001$). For each additional unit of student voice, the average score on the political attentiveness scale increased

by .37 points ($p < .001$). Students commented that they had learned to “help others,” and “volunteering just a few hours can make a big difference for other people” (Morgan and Streb, 2001, p. 163). Finally, student attitudes toward the elderly and disabled improved as a result of service, as measured by students’ beliefs that they could work with the elderly, and believed that those with disabilities can hold jobs competently. The t^2 coefficients for this change were .17 and .25 ($p < .01$). There was no evidence that the interaction variables significantly affected these results.

The sample size is large, with demographics that allow for external validity. While the demographics of the study are disproportionately Caucasian, the authors claimed that the racial composition of the sample is closely aligned with the national percentage of school-aged youth that are African American. The authors controlled for confounding and interacting variables in a method consistent with peer research. The high coefficients of the findings have strong statistical significance, with *p*-values of .01 and .001. The internal validity of the results is questionable, however, since no alpha scores are reported on the survey instruments to indicate their reliability. A few other factors also stand out as problematic; with the absence of a control group, there is no data to compare the results from a similar sample group not participating in service-learning and control for maturation. Additionally, the 31% mortality rate is high, and the gender ratios are heavily disproportional.

Riedel (2002) drew attention to how different types of service programs affect students’ feeling of civic obligation, a dimension he argued had yet to be

studied in the domain of political engagement. His mixed-methods, nonequivalent control group design focused on four Minnesota high schools that he selected based on the types of service programs and social studies courses offered, as well as the schools' desire to participate. Three of the sites were public, suburban schools outside Minneapolis, and the fourth was an urban, private Catholic high school. The programs included an elective service course, two mandatory service programs connected to social studies courses, and two social studies courses, one of which had a service option. All participants (N = 294) were in the 11th and 12th grade, and the social studies course with no service option served as the comparison group (n = 33).

Students completed questionnaires before and after service to measure the independent variable of feelings of civic obligation. Participants rated the importance of eight possible citizen obligations on a scale (α = .66), including items such as voting, keeping informed, and volunteering in the community. Riedel (2002) conducted semi-structured interviews with 14 teachers and service coordinators involved in service programs, and two social studies teachers not involved with service. He spent 42 hours observing at students' service sites and casually talking with students and site supervisors, systematically sampling sites that were representative of each program. Riedel dichotomized the programs into participatory and private types of service: "Participatory community service programs envision good citizenship as mainly as public action to address problems of wide concern…By contrast, private community service programs envision good citizenship mainly in terms of personal virtue" (p. 503). Information

he gained through interviews and observations were used to designate programs as participatory or private. He checked his judgments against participants' reflective essays, searching for participatory or private themes, and found his descriptions to be accurate. In courses with mandatory service, the author divided participants into two groups—those who met the service requirement, and those who went beyond 40 hours, whom he dubbed "intense service." Finally, Riedel used a multiple regression model to predict civic obligation alongside other controls on student characteristics, such as influential adult education and extracurricular activities. This static-score method allowed the author to control for participants' differences in initial levels of the dependent variable. The coefficients he calculated are unstandardized least squares regression values.

Three of the participant groups showed small gains in their mean civic obligation scores: students in the mandatory service program at the private Catholic school completing regular service, those completing intense service, and students from the social studies course with a service option. Riedel (2002) had designated both these programs with civic gains as participatory programs. The multiple regression model then predicted posttest scores on the civic obligation scale controlling for individual characteristics, and indicated that the two strongest predictors of gains were involvement in extracurricular activities (.256, $SE = .052$, $p < .01$, two-tailed), and participatory service programs (.759, $SE = .350$; .514, $SE = .049$, $p \leq .05$, two-tailed). In the mandatory private school program, only students with intense service showed significant gains.

Riedel (2002) compensated for history and maturation in the study with use of pre- and post-data and a comparison group, and his instrumentation is consistent throughout. His findings were statistically significant, and the least square coefficients of his findings are substantial in supporting his conclusion of participatory service programs heightened efficacy. He created triangulation of data between survey scores, interview data, and his observations. The limitations of his research begin with the low alpha score of his civic obligation scale (α = .66), creating doubt around its reliability as the primary instrument. Riedel provided no information on the gender and ethnicity of his participants, weakening the external validity of the work. His sample size was adequate, but suffered a high mortality rate, losing 35.3% of the control group (n = 33), and 26.9% of the social studies class subsample which he claimed made significant civic gains. The attrition in the study was not explained, and it resulted in a small control group of only 20 students for a sample size of roughly 275 students.

In terms of his qualitative research, he makes a distinguished contribution to the field in his method of coding of programs as participatory or private, but an outside perspective could have enhanced the confirmability of his analysis, as could more transparency in his coding methods. No member-checking is conducted, and he relied exclusively on himself as instrument without acknowledgement. Finally, criteria distinguishing community service—be it participatory or private—from service-learning is not discussed, thus the contribution to service-learning research as a specific pedagogy is not as strong.

With a mixed-methods, nonequivalent control group design, Billig, Root, and Jesse (2005) questioned the impact of service-learning on civic engagement, and civic knowledge and skill acquisition. They also examined to what extent program quality, and teacher characteristics and practices moderate civic outcomes. The study sample included 1,052 students, grades 9-12, with both service-learning participants (*n* = 645), and a comparison group (*n* = 407). The sample was 55% Hispanic, 31% Caucasian, and 13% African-American. Sixty percent of the participants were female. The authors selected high school sites for the studies that were first nominated by service-learning experts, including active researchers and state Learn and Serve directors. Teachers from the nominated sites were interviewed, and then the first pool of nominations was analyzed to find comparison groups. Finally, the authors identified the sample sites based on program quality, availability of a comparison site, school district approval, and the diversity of the school's sociodemographic profile.

Billig, Root, and Jesse (2005) administered student (*n* = 1052) and teacher (*n* = 26) surveys in the fall and spring of the 2003-04 school year that measured civic and academic outcomes. The authors analyzed the data with both MANOVA and ANOVA models, and a hierarchical MANOVA with repeated measures using nested pairs controlled for variation in student grade level and subject area. Additionally, they ran multivariate *t*-tests to compare student groups, regression analyses to determine the contribution of specific moderating variables on outcomes, and canonical correlation in the case of multiple independent and dependent measures. The National Assessment of Academic

Progress (NAEP) measured civic knowledge, and the authors created subscales to measure for the following constructs (α = .462 - .884): academic engagement, valuing school, school attachment, civic knowledge, self-reported civic knowledge, civic skills, community attachment, civic dispositions, efficacy, and civic engagement.

The authors assessed moderators of outcomes, including variables such as characteristics of the service experiences—duration, nature and type of service, and quality—as well as characteristics of the teachers and classroom practices. Qualitative information from focus groups and interviews supplemented questions of service quality from students and teachers on survey items from the Student Service Learning Subscales (α = .689- .844).

The measured impact of service-learning on student outcomes was not significantly higher than the outcome measures of the comparison group. In some cases the measured outcomes were equal, or in the case of civic knowledge, the control group (M = 2.709) actually scored slightly higher than the treatment group (M = 2.7). The results on the moderators of outcomes showed that student-selected service activities yielded significantly higher gains in civic knowledge (M = 3.05; SD = 1.886), than teacher-selected civic issues (M = 2.83; SD = 1.596; $p < .02$) or environmental problems (M = 1.86; SD = 1.167; $p < .025$).

A MANOVA analysis on the relationship between type of service—direct, indirect, or political or civic action—was significant, F (24, 1,028) = 5.326, $p < .001$, Eta-squared = .111. Indirect service (M = 4.11; SD = .611) was linked with

higher levels of academic engagement than civic or political action (*M* = 3.86; *SD* = .722). Civic knowledge, however, benefited significantly from direct civic or political action (*M* = 3.01; *SD* = 1.626) over direct (*M* = 1.67; *SD* = 1.15) and indirect service (*M* = 2.21; *SD* = 1.353). The authors found the duration of service was statistically significant on outcomes *F* (36, 1,1,866) = 6.326, *p* < .001, Eta-squared = .109. Civic knowledge measures were highest with service projects of one-year (*M* = 2.81; *SD* = 1.696), and civic skills were at their highest with service for one semester (*M* = 3.04; *SD* = .480). The final moderator relevant here that the authors mentioned were the teacher characteristics of experience and active versus passive instructional strategies. Canonical correlation analysis demonstrated a significant relationship between teacher experience, ranging from 2 to 24 years, and student outcomes *F* (12, 244 = 5.320), *p* < .001, Eta-squared = .113. Teacher's experience, specifically with service learning, was positive correlated with students' posttest scores on two subscales: civic knowledge *r* (540) = .174, *p* < .001, and civic dispositions *r* (538) = .139, *p* < .01.

The research conducted by Billig, Root, and Jesse (2005) had the advantage of a large sample with control groups, with a unique focus on a primarily Hispanic population that was heavily bilingual; however, because of this, the external validity of the study was not high. History and maturation were accounted for by the research design. The authors created triangulation of the data between teacher and student surveys, and interview data. The instrumentation is consistent in pre- and post-surveys, but the subscales constructed by the authors to measure civic outcomes did not have high

internally consistency, with some reporting low alpha levels ($.46 < \alpha < .80$). The authors had the opportunity to add some dimension to the credibility of their research with qualitative data, but there is very little analysis of the interview and focus group material collected. Also, there was no discussion of member-checking, author bias, or mention of who collected the qualitative information and how it was coded.

The next two studies also explored service-learning's impact on participants' civic behaviors, as well as attitudes. Melchior and Bailis (2002) used a quantitative, nonequivalent control group design to evaluate two service programs, the first of which was the federally-sponsored Learn and Serve (LS) program. The LS evaluation took place in seven middle schools and ten high school-based programs across nine states with both a treatment (n = 608) and a control group (n = 444). It used pre- and post-program surveys and an analysis of school record data for program participants. At the one-year follow-up, attrition in both groups resulted in a total of 764 participants with 460 completing the LS program and 304 in the control group. The 17 programs evaluated were fully implemented service-learning programs, selected according to the following criteria: program standing of longer than one year; higher than average hours of service; regular, written and oral reflection; and linkage to a formal, academic curriculum.

To assess the efficacy of the programs, the authors used the following measures: The Social Welfare Subscale drawn from The Personal and Social Responsibility Scale of the Search Institute Social Issues Survey, The

Community Involvement Subscale drawn from the same survey, and The Service-Learning Leaderships Scale taken from Sieber's Service Learning Leadership scale, as well as self-reported data on involvement in community service and hours of service performed. The authors' adapted measure of Service Leadership assessed students' sense of awareness of community needs, and their capacity and commitment to address those needs both in the present and future.

In their statistical analysis, Melchior and Bailis (2002) used an ANCOVA adjusting for differences in baseline scores by including them as separate variables, and used as covariates, similar to age and gender, in the regression formula. The intention here was to allow an estimated outcome adjusted for differences in both baseline scores as well as participant characteristics. During the study, the authors added a second method of analysis known as the change score or "difference-in-differences" approach that adjusted for differences at the baseline by subtracting the baseline score from the post-program score. The authors believed that the ANCOVA model tended to under-adjust for differences at the baseline, which resulted in an overstatement of program impacts. They relied on the change score method to provide a less biased, and more conservative estimate in their results and included both sets of analyses in their final report.

The results of the initial LS evaluation using the ANCOVA demonstrated significant effects on all of the participants' outcome measures with average differences in the scale scores ranging from 2% to 5% (p = .01). Using the

change score method, three variables demonstrated change (p = .01): an increase in Service Leadership for high school students (5-8% average score difference); Involvement in Community Service; and Hours of Service. The one-year follow-up revealed no significant results in any of the outcome measures.

Melchior and Bailis (2002) evaluated a second program, the foundation-sponsored initiative Active Citizenship Today (ACT). In their research on the ACT program (N = 588), the authors used the same instruments as the LS evaluation, but also included pilot testing of an open-ended "problem solving exercise" designed by ACT teachers and program coordinators to explicitly measure the application of civic skills taught through case study scenarios. The study drew participants from 12 middle and high schools in Omaha, NE and Jefferson County, CO. In addition to the instruments described above, the authors also evaluated participants' self-assessed knowledge of their community, sense of belonging, and attitudes toward teamwork.

Melchior and Bailis (2002) again employed the change score method, in conjunction with an independent sample t-test to assess differences between the treatment (n = 298) and control groups (n = 290). There were no differences in demographic characteristics of the sample groups and the only significant impact found for the ACT program was a change in students' perception of their communication skills (p = .05) as measured by an adapted version of the Communication Subscale of Greenberger's Psychosocial Maturity Scale.

The evaluation of the LS and ACT programs by Melchior and Bailis (2002) was strengthened by the authors' acknowledgement of their positive bias toward

service-learning. They did not assert implications for K-12 schools that were not based on their data. The authors could have strengthened the internal validity of their research by reporting the alpha levels of the instruments they used to measure the outcome variables, and reporting the mean differences in pre- and post-surveys for each measure. Also, the reliance on self-reported data in the significant outcomes measured is a potential weakness in the reliability of the findings. The external validity of the study is lacking due to no inclusion of the demographic data of the sample, and no mention of what states were selected to gather the sample from, or a rationale of how those states were selected.

In a correlational study using data from the 1999 National Household Education Survey (NHES), Schmidt, Shumow, and Kackar (2007) investigated if participation in service activities contributed to high school students civic efficacy and knowledge, as well as academic and behavioral adjustment. They also explored if the type of service performed was related to a specific outcome. The survey used random digit dialing to contact adolescents and their parents over a four-month period, with a focus on service participation. The authors concentrated on a sample of high school respondents (N = 4,306) attending either public or private schools, evenly distributed through grades 9 through 12. The sample was evenly split between males and females, and was 63% Caucasian, 17% Latino, 15% African American, and the remainder Asian and Other.

Schmidt et al. (2007) measured civic efficacy with a six-item composite variable, on a scale from 6, signifying less civic efficacy, to 12, signifying more

civic efficacy (α = .41). They measured civic knowledge with a five-item composite variable, on a scale of 1, meaning less civic knowledge, to 5, meaning more civic knowledge (α = .66). To measure behavioral adjustment, the authors constructed a five-item composite behavior variable with a scale of 1, indicating fewer problems, to 4, indicating more problems (α = .51). Finally, parents reported students' grade point averages, which the authors translated to a 5-point scale.

The authors coded service activities into three types: direct contact with individuals in need, assisting organizations, and benefiting the environment and/or animals. They used several control variables in their analysis, including gender, race/ethnicity, type of school, and socioeconomic risk status. Multiple regression models estimated the impact of service on the four outcome measures controlling for background variables, and Schmidt et al. calculated an effect size for each *b* coefficient, which appears in parentheses below.

The authors indicated a statistically significant, positive correlation between participation in any service and increase in grades, b = .34 (.12), SE= .03, β = .18; $p < .001$, civic efficacy, b = .26 (.03), SE = .04, β = .11, $p < .001$, and civic knowledge, b = .34 (.16), SE = .04, β = .12, $p < .001$. The study suggested a significant negative correlation between service and behavior problems, b = -.18 (.15), SE = .03, β = -.10, $p < .001$.

Independent *t*-tests examined the relationship between the three types of service the authors coded, and the outcome measures, on a subsample of students that participated in a longer version of the NHES survey (n = 857). The

results indicated that students in direct service earned higher grades ($t = 2.92$, $p < .01$, $d = .35$). Adolescents assisting organizations reported greater civic efficacy ($t = 2.86$, $p < .01$, $d = .21$) and civic knowledge ($t = 1.94$, $p < .05$, $d = .15$) than students of other service types.

Schmidt et al. (2007) made a substantial contribution to the literature by studying a large, nationally representative sample with high external validity. They did not attempt to infer causality or directionality in their findings, holding true to the limitations of their correlational design. The authors admitted that data from national surveys has serious limitations on the precision and detail of research, which prevented them from gathering more background on the nature of service conducted and the predispositions of the participants. They disclosed that this is why a distinction was not made between community service and service-learning, with clear criteria established for both. The authors made it clear that the internal consistency of their instruments was very low, with reliability levels ranging from .41 - .66, and revealed that while their findings were significant, instrumentation with greater validity might have increased effect sizes. The authors acknowledged that the effect sizes indicated by Cohen's *d* were relatively small in comparing the relationships between service types and increased outcome measures. The measured effects may have been a result of selection effects—meaning predispositions—rather than the effects of service.

Special Populations

In recent research, scholars have addressed the lack of empirical research demonstrating the civic outcomes of community service-learning on

special student populations, such as adolescents in gifted or culturally relevant educational programs. The next two articles focused exclusively on how service-learning affected gifted students, followed by a third article that examined service-learning's effectiveness in meeting the cultural needs of a community.

Terry (2000) was a leader in service-learning research on gifted students, to which she contributed a qualitative case study examining her students' attitudes and perspectives on the effects of their participation in service-learning. She selected middle school participants (*N* = 4) with purposive sampling techniques, who had been part of two Community Action service projects she facilitated through her class teaching—The Backstage CREW and SWAMP—both of which received national media attention. Students assessed their community's needs, and selected and designed the projects themselves. Backstage CREW was a five month program in which students planned the restoration of a historic theater in rural Georgia, and SWAMP was a three-year program in which students created a waste management plan for their county.

Terry (2000) led focus group interviews with her students between one and five years after the service projects were completed, allowing a longitudinal perspective. The sample was primarily female, with three Caucasian students, and one African-American student. Terry chose students to represent all the task groups from the projects, and to create a mix of male and female perspectives. She recorded and transcribed student interviews, and used Glaser and Strauss's constant comparative method of analysis to gradually identify a core of emerging concepts in the data. Segmenting the information to code the transcribed

interviews also assisted Terry in identifying additional categories and themes. The next step of her analysis was axial coding, in which the author looked for categories across all the data sets. The five themes she identified from the data were: methodology, attitudes, personal and social development, commitment, and empowerment.

The findings indicated that students found the methodology of the service-learning course to be important. Students in both projects followed the Osborn-Parnes model of creative problem solving, which allowed students to develop higher-order thinking skills, and reflect on their work. They enjoyed taking on unique tasks that were self-disciplined, and working at their own pace. Student attitudes were consistently positive throughout the projects, and students completed their work with strong feelings of accomplishment, support, respect and having left “a mark” (Terry, 2000, p.11). In terms of personal and social development, students developed new understandings and perspectives, academic skills, self-development, the ability to overcome problems, and better cooperation. Students demonstrated a strong sense of caring and commitment from their service work, as a result of the ownership they felt about their learning, and the visible impact they could see they had on the real world. The student leader of the SWAMP project remained committed to the environmental work she began in high school, and pursued an environmental consulting degree at the college level. The final attitudinal outcome measured through the interviews was empowerment, and as result of their service, participants felt as if their ideas were listened to “…as adults instead of kids” (Terry, p.14). Empowered by the

success and attention their work received by the community, another female student responded, “It makes me want to be President!” (Terry, p. 14.) Students reflected that with more time devoted to solving community problems, they could accomplish even more.

Terry (2000) was very transparent in her coding of the qualitative data, and she employed multiple, independent investigators throughout the process to reduce personal bias and increase the confirmability of her research. Triangulation of data and member-checking with one of the interviewed participants of the study greatly added to the data’s credibility. She acknowledged that a limitation of her work is its transferability to other student population and types of service. In explaining her rationale for selecting students for the focus groups, Terry did not mention the demographics of the school’s gifted populations or indicate how representative of the larger school population her sample was. Certainly the focus on gifted students is a factor that limits how representative her findings are, and in discussing why service-learning best suits the need of gifted students, she made no concession to how the same benefits she propounded for the gifted could benefit all students. When Terry claimed, “Gifted students need more from their schooling than a simple regurgitation of facts from a teacher or textbook” (p. 17), she failed to point out at that all students—labeled gifted or not—deserve more than an inadequate education.

Despite the hopeful outlook offered in the implications of Terry’s (2000) findings, and her call to action for more empirical studies relating the benefits of service-learning to gifted populations, Lee, Olszewski-Kubilius, Donahue, and

Weimholt (2007) still described a paucity of research seven years later. Their work examined the effects of service-learning on gifted adolescents' civic attitudes, behaviors, and leadership, in the short- and long-term, through a quantitative, nonequivalent control group design. The study drew its service-learning participants (n = 118) from the Civic Leadership Institute (CLI), a three-week residential summer program intended to develop civic responsibility and leadership skills through traditional academic work, service, and field experience. The author selected a control group (n = 112) from the Equinox Program, also a three-week summer program, offering high-level and AP core courses for high school credit. Both programs were intended for gifted adolescents who qualified with their SAT and ACT scores. The balance of male to female participants was roughly equal. Caucasians comprised 53% of the sample, Asians 32.4%, and African-American, Hispanic, and Native American students made up the remainder.

The authors used three instruments in this study to measure the outcome variables: the Civic Responsibility Survey: English Version (CRS; α = .93), the Civic Behavior Survey (CBS), and the RSL (α = .85). The CLI institute designed the CBS to examine the civic behaviors of its participants. For the treatment students, the authors administered the surveys multiple times between Summer 2003 and Spring 2004, with initial testing before the three-week session, a second test after the three-week session, and a third, six months after the program's end. The control group was also tested with all three survey items, but only before the program and six months after. The authors collected data for the

first and second tests on-site, during class, and the final tests were mailed to all students who had completed the previous surveys. The response rate for the final survey was 45.2%.

Lee et al. (2007) created participants scores on the CRS for the three clusters of items: connection to the community, civic awareness, and civic efficacy. The authors computed participants total leadership scores on the RSL using all 26 of the survey's items. In investigating the thirteen CBS items, they adjusted alpha levels to .004, by dividing .05 by the number of items. They analyzed all data using SPSS 11.0 software, and compared students' performances on the surveys by type of program, and time of administration. Using a mixed-model ANOVA, the authors examined whether the differences in their performance was significant as affected by these variables. Independent sample *t*-tests served to explore differences in performance, specifically as an effect of program type, between the control and treatment group students on the first and final surveys (CLI, $n = 57$; Equinox, $n = 47$).

On the CRS, treatment students demonstrated significant differences in their performance on two of the three cluster areas, with higher mean scores for the CLI students. For connection to the community, CLI students earned a mean score of 4.57 ($d = .29$; $p < .05$) on test one, which rose to 4.96 ($d = .57$; $p < .05$), while the comparison group rose from a mean score of 4.34 to 4.55. The difference was significant, $F(1,102) = 5.59$, $p = .02$, partial eta = .05. On civic efficacy, the difference was also significant, $F(1,87) = 4.32$, $p = .04$, partial eta = .05. Mean scores rose for the CLI students from 4.23 to 4.51 ($d = .54$; $p < .05$),

and comparison group scores rose from 3.98 to 4.18. There was no significant difference between the groups in the cluster of civic awareness.

The authors determined there were significant pre-program mean differences between the students of the control and treatment groups, favoring CLI students on the measures of connection to the community, t (224) = 2.19, p = .03, and civic efficacy, t (222) = 2.58, p = .01. On the final survey, differences between the CLI and Equinox students in civic responsibility were again significant for connection to community, t (102) = 2.91, p = .00; civic efficacy, t (87) = 2.50, p = .01; and now, civic awareness, t (87) = 2.72, p = .01. Mean scores for the CLI students were higher in all three areas of civic responsibility, with medium effects sizes for the mean differences ($.5 < d < .8$). There were no differences in civic behavior of leadership skills between or within the control and treatment groups.

Lee et al.'s (2007) findings coincided with Terry's (2000) report of heightened civic awareness for gifted adolescents as a result of service learning, raising the research's reliability. The effect sizes for the treatment group, indicated by Cohen's d, are in the .5 range and are therefore moderate. The instrumentation was consistent for both groups, and the internal validity of the CRS and RSL surveys were high. The CBS, created by the CLI however, did not have a reported alpha score to demonstrate its validity. Additionally, it was designed specifically for CLI participants, and so the survey's internal validity for other student groups is questionable. While the authors included a control group to account for maturation, they acknowledged their choice was problematic, first,

because the content of the two programs was quite different; the CLI had a focus on leadership and civic responsibility, which matched the outcome variable, while Equinox had no intention to affect students' civic responsibility. The authors also admitted significant differences indicating the treatment group began with a heightened sense of civic responsibility. The authors acknowledged this was attributable to participants' self-selection for the CLI institute, and predisposition toward civic engagement. Lee et al. were able to capture more longitudinal data by conducting the final survey six months after the program's end, but the time lapse, combined with their reliance on mail to collect the data, resulted in a high mortality rate of 54.8%. This also led to a small sample size of 57 participants, with a demographic profile that does not have high external validity for other school settings, because of students' ethnic and academic backgrounds.

Finally, Yamauchi, Billig, Meyer, and Hofschire (2006) had the unique opportunity to examine the effects of culturally relevant service-learning for high school participants (N = 84) in a case study of the Hawaiian Studies Program (HSP). The HSP integrates the learning of traditional Hawaiian values and culture with traditional Western curriculum, and is open to all students, grades 10-12. The program was evaluated at a rural public high school on the island of O' ahu that serves predominantly Native Hawaiian students. The school serves high numbers of low-income and special education students, and it has a higher than average drop-out rate for Hawaiian schools. For the period of this study—the 2001-02 and 2002-03 school years—the program was a collaboration of three to four teachers and over 12 community based organizations (CBO).

Students completed service-learning field work one full day each week, and rotated through four placements over the course of a semester. The mixed-methods study focused on the HSP's effects on participants' (*n* = 55) civic attitudes, connectedness to their community, and career development, and compared this with a control group (*n* = 29).

An external evaluator administered a student survey at the end of the 2003 school year that assessed for: "(a) student experiences with HSP curriculum and instruction; (b) student motivation and school engagement; (c) career-related outcomes; (d) civic development and participation outcomes; (f) personal and social development outcomes; and (g) connectedness to cultural heritage and community" (Yamauchi et al., 2006, p.154). An external researcher from the Center for Research on Education, Diversity, and Excellence (CREDE) interviewed three teachers and fifteen community members involved with the HSP for the 2001-02 academic year. In these 60 to 90 minute interviews that were recorded as part of a larger study on the HSP's development, the researcher asked participants (*n* = 18) about the HSP's initiation, outcomes, and future directions. An external evaluator conducted and recorded a second set of 60 to 90 minute recorded interviews for the 2002-03 academic year that questioned four teachers, nine community members, and three student focus-groups (*n* = 19). This time, researchers explored the focus group students' perceptions of the program, including the perceived influence their participation had on them and others. The outside researchers analyzed the transcribed interview data using thematic data analysis methods, and they identified excerpts

relating to the outcome variables, examining them for patterns across the participant groups.

Student survey results indicated that HSP participants scored significantly higher mean scores on their perceptions of their connection to the community. HSP students felt more strongly than control group students that their community viewed them as valuable members (HSP, $M = 2.65$, $SD = .87$, $p < .01$; Non-HSP, $M = 2.07$, $SD = .75$), and that they took action and made changes in their community (HSP, $M = 2.69$, $SD = .94$, $p < .01$; Non-HSP, $M = 1.83$, $SD = .81$). Data from the interviews suggested that students felt less alienated from their community when they served, and that service-learning expanded the network of mentors and role models students had. Students had a heightened sense of community activism after experiencing how peer mentors can help pass down generational knowledge, and there was more integration between home and school. There were no statistically significant differences between the quantitative measures of students' civic attitudes. For the measure of career development, HSP had the most significant higher mean scores in the areas of time management (HSP, $M = 2.35$, $SD = .52$, $p < .01$; Non-HSP, $M = 1.93$, $SD = .53$) and resume writing (HSP, $M = 2.43$, $SD = .60$, $p < .01$; Non-HSP, $M = 2.00$, $SD = .76$). Interview data indicated students had been exposed to a variety of careers, learned details about them and how to prepare for entry into different fields. Community members felt that HSP participants had greater access and ability to navigate openings in the job market, with enhanced preparation from portfolio and resume development, as well as letters of recommendation.

In discussing how the authors selected the HSP for evaluation, the authors of this study readily admit their bias toward sociocultural learning theory that posited that all higher cognitive functions are rooted in social interaction. Their criteria for choosing a quality service learning program were strong, ensuring the program had preplanned curriculum integration with service, collaboration between school and community, and intensive service over a full-year period with exposure to a variety of service types and adult support. The methodology of the qualitative analysis was clearly described and was highly confirmable through the use of external researchers and evaluators in the data collection process. Triangulation of student surveys, adult interviews and student focus group interviews reduced bias, and also strengthens the credibility of the work. There is, however, no mention of member-checking with participants. While Yamauchi et al. (2006) provided a control group, they conceded a strong limitation of their quantitative analysis is that it does not account for the history of the participants or allow longitudinal data; the survey was only administered once, post-service. The authors reported no alpha score on the survey instrument, leaving to question its internal consistency. The student sample size was small, and participants were self-selected. Additionally, this was a case study working with a special ethnic population in culturally specific program, and therefore the external validity of the study is low.

Community Involvement and Social Responsibility

Closely related to the civic outcome measures above, the following research evaluates what the authors refer to as a sense of community

involvement and social responsibility. These measures overlap with the previous themes of civic engagement, but signify a slightly broader variable that is not as focused on the skills, behaviors, and knowledge particular to American government. Community involvement is more about the attention and importance an adolescent attributes to the general welfare of those around them. Again, some researchers also bundle these items with academic success, but a concentrated analysis of those findings will appear in the next section.

One way to determine youth's community involvement is to assess their volunteer experience, and intentions to volunteer in the future. Metz and Youniss (2003) conducted a longitudinal, nonequivalent control group design study to determine if after required service, high school students served voluntarily, and if their intentions to serve as volunteers in the future were raised or lowered. Working with a middle-class suburban town outside Boston, the authors pulled participants (N = 484) from three consecutive senior classes, during which time the school had implemented a mandatory service policy. The Class of 2000, exempt from service, formed the control group (n = 172), while the Classes of 2001 and 2002 were required to perform 40 hours of service and were combined as a treatment group for the study (n = 312). The majority of the school population was Caucasian, and the remaining quarter was composed of Asian, African American, and Hispanic students. Almost half of the student body was Catholic, and 93% of the school's graduates attended college.

Two groups were formed from the control and treatment groups, based on their inclination to serve. In the control group, the authors defined more-inclined

to serve students as having volunteered in two of the three measured years, and less-inclined to serve students as having completed service of only one year or not at all. For the treatment group, more-inclined students were those who completed their 40-hour service requirement in either 10th or 11th grade, and less-inclined students did not complete the requirement until 12th grade. Metz and Youniss (2003) determined that the more- and less-inclined to serve groups were dissimilar on background measures including gender, parents' volunteerism, mother's education, religiousness, and GPA. They also differed in personality dispositions, measured by the "Big 5" personality assessment for adolescents (α = .74 and .88). The authors used a Pearson Chi-square statistic with the Haberman's adjusted residuals to test for differences in categorical variables among the more-inclined groups from both cohorts, and less-inclined groups from both cohorts, and an independent *t*-test to test for differences in scaleable variables. They determined that the cross-cohort groups were similar according to their inclination to serve.

Participants designed and implemented their own service experiences, with the only stipulations being that their service benefited at least one unrelated person, and that their work must be documented with written reflection and an adult or supervisor's approval. Students were assessed at the beginning and end of 11th grade, and at the end of 12th grade. Participants self-reported data on whether they had performed voluntary or required service during the previous year. The authors examined patterns of volunteering before and after the required service to determine the students' likelihood to volunteer service in the

future. At each assessment point, the authors asked participants to rate from one to five the probability of their performing service after high school graduation.

The results indicated that there was a 67% rate of voluntary service participation for both the treatment and control groups in their senior year. Both groups' participation rates increased over time, although the control group decreased slightly from 69% in its junior year. For more-inclined to serve participants who completed their required service in 10th grade (*n* = 68) and 11th grade (*n* = 96), rates of volunteerism over time increased from roughly 40% before service to 90% after service. The less-inclined to serve participants (*n* = 148) increased their rate of volunteerism from approximately 30% to just above 40%, but the authors did not mention if this was a significant increase. The control groups' (*n* = 172) volunteerism rates increased over time from approximately 55% to close to 70% from grade 10 to grade 12.

Finally, the more-inclined to serve groups of both cohorts demonstrated no increase in their intentions to serve over time, nor did the less-inclined to serve control group students. However, after the 40-hour service requirement, less-inclined to serve students' mean score, reporting their future intentions to serve, rose significantly from approximately 3.1 to 3.4 on a 5-point scale. This score was distinctly higher than the score of the less-inclined to serve students in the control group, whose mean score was slightly below three points. The authors claimed that their regression analysis demonstrated this difference was attributable to no other factor than completion of service.

The authors acknowledged that the data suggests a student's inclination to serve is positively associated with—but not attributed to—mandated service, and this is only one of several background factors. The authors strengthened the precision of their findings by accounting for differences within the cohorts on background and personality variables. They also used regression analysis to isolate the service requirement as a likely factor in raising a high school senior's intention to volunteer in the future, given they were unlikely to volunteer before their service experience. History and maturation were both accounted for in this study by pre- and post-assessments and the use of a similar control group.

The external validity of this study is lessened by the high percentage of Catholic students in the school population; religious faith is a confounding variable that could strongly influence students' interaction with their community. In collecting data from students, the authors do not describe their instrumentation, or give any sense of its reliability. The data is presented solely in tables that list rates of participation and volunteerism, but the authors do not provide details on how those rates were calculated. Precise percentages and mean scores are not reported, and the tables provide only rough estimates of the findings. Additionally, there are no *p*-values reported in two of the three tables, and the authors failed to report the significance of their approximated findings. Most importantly, the qualities of community service-learning are not controlled for in this study, and the high variation of students' service experience with no direct connection to classroom curriculum highly restricts the generalization of these findings to a more structured, academically integrated service requirement.

McGuire and Gamble (2005) explored how adolescents' psychological engagement with service activity accounts for a change in their sense of community belonging and social responsibility, through a quantitative, one group pretest/posttest design with variable findings. The high school participants in the study (*N* = 68) were 68% female, and 32% male, with a mean age of 16.4. Hispanic participants made up 43% of the sample, while Caucasians were 28%, and Native Americans were 23%. Fifteen percent of the sample were English language learners; of the teens with a second language, 50% spoke Spanish, and 30% spoke a Native American language.

The authors derived the data for their study from a larger evaluation of sexuality education programs delivered by high school students to 5th through 8th graders. In collaboration with four high schools in the rural Southwest, staff from the Office of Adolescent Pregnancy Programs presented the Postponing Sexual Involvement (PSI) program, a five hour abstinence based sexuality program focusing primarily on sexual refusal skills. Students trained for eight hours, volunteered for ten hours during the program, and committed to four to five hours of meetings. They completed a pre-survey of 198 multiple choice items prior to their training during the fall semester. At the end of the school year, they completed a similar post-survey, on which the authors excluded demographic items, and added items measuring the degree of engagement with the service.

The post-survey measured time spent leading PSI, psychological engagement, social responsibility, and community belonging. McGuire and Gamble (2005) measured psychological engagement with a scale including

seven items on varying-point Likert scales (α = .84). Community belonging was measured in open-ended questions during the pre-survey to determine the participants' description of the term "community." The scale of community belonging was adapted from Goodenow's Psychological Sense of School Membership Scale; it included 16 items that were measured with the pretest and posttest (α = .90). Social responsibility measurements used a 12-item sub-scale of Greenberger's Measure of Psychosocial Maturity (α = .78 - .84). With the results, the authors completed ANOVAs to compare groups on the outcome measures, by ethnicity, gender, and change in sexual behavior.

In the findings, females reported higher levels of psychological engagement, t (65) = 3.40, $p < .001$, and social responsibility, t (65) = 2.04, $p < .05$, than males; however, they were not different in their reported levels of community belonging. Interactions with gender yielded no other significant observations. Two-tailed paired t-tests compared the raw pretest and posttest scores for community belonging and social responsibility. There was not a significant change for community belonging, t (66) = 1.39, but there was significant increase in social responsibility, t (66) = 2.28, $p < .05$.

Residual change scores were created for community belonging and social responsibility and used as unique variables. A zero-order correlation was made of time spent and psychological engagement with community belonging and social responsibility. Psychological engagement with service accounted for a significant portion of change, r = .37 for community belonging, and r = .39 for social responsibility ($p < .01$), but time spent did not affect either significantly.

Compared with other independent predictors, such as sexual history and gender, psychological engagement was the most significant predictor of change for the individual in their feelings of community belonging and social responsibility ($p < .01$).

McGuire and Gamble's (2005) work benefited from a very multicultural sample, with significant data from Hispanic and Native American participants. The author's acknowledged the study's impact was limited however by a small sample size ($N = 68$), in which the gender balance was weighted toward females. Therefore, the findings do not create high external validity. The sample was not randomly selected, and the authors admitted participation required significant volunteer work, meaning the sample's composition could be greatly influenced by it participants' predispositions toward service. Additionally, they realized the study did not control for maturation with a non-service control group. Confounding variables such as other classes and life experience could have also contributed to the measured change in the outcome variables, beyond the effects of the participants' ten hours of service, especially due to the duration of time between pretest and posttest.

Halfacre, Chessin, and Chambless (2006) conducted a one-group pretest/posttest design with middle school participants ($N = 42$) to investigate the effects of service-learning on their attitudes and actions related to social and academic responsibility. The sample was 59% male, 41% female, and ages ranged from 10 to 14, with the majority in either grades 5 or 6.

In partnership with the Corporation for National and Community Service and the University of Mississippi School of Education, the Learn and Serve After-school Service-Learning Lighthouse Partnership program was created and implemented in the Lafayette County Public School District. The goal for sample selection was to identify students who were struggling in more than one academic area for reasons other than special limitations or behavioral disorders. The criteria were two or more subjects in need of academic intervention, with limited behavioral problems and no special education rulings. Students who met these criteria were invited to participate, and allowed to leave at any time. The service-learning program operated with goals of fostering civic responsibility through participation in service that meets community needs. Service also had to be integrated into curriculum, and provide structured time for reflection. Program coordinators followed and monitored these goals throughout the program.

The administrators of the program worked with students for 90 minutes each Tuesday and Thursday after school. Pre-service teachers tutored the middle-school participants, and collaborated with classroom teachers to choose and plan the service-learning project. Students carried out their projects during the on-campus tutoring program.

The authors used a survey instrument that was a hybrid of four previously validated instruments to determine the effects of service-learning on social and academic responsibility. The survey consisted of forty questions on a Likert scale with 11 questions from the Helping Dispositions Scale, 6 from the Scale of Service Involvement, 13 from the Community Service Attitude Scale, and 10 from

the Lighthouse Partnership After-School Service Learning Program Survey. The authors divided the questions in the survey into two categories—social and academic responsibility—and they aggregated and averaged scores from each to construct the dependent variables Academic Responsibility Score and Social Responsibility Score.

They determined the effects of the treatment by comparing pretest and posttest scores, with measured raw scores from the same instrument. To determine statistical significance, they used a *t*-test for relational means. Halfacre et al. (2005) assumed equal intervals for data in the *t*-test, and calculated a confidence interval from the mean scores.

The mean score for the pretest measuring student attitudes and actions related to social responsibility was $M = 3.11$, $SD = .37$, $SE = .0658$, and the mean score for posttest was $M = 3.36$, $SD = .38$. There was a 95% confidence interval for the posttest social responsibility data with $CI = 2.9784, 3.2466$. The difference between the pretest and posttest means were statistically significant ($p < .001$), with a posttest score that was greater, falling above the 95% *CI*. The mean score for the pretest measuring student attitudes and actions related to academic responsibility was $M = 2.83$, $SD = .32$, and the mean score for the posttest was $M = 3.10$, $SD = .36$. The 95% confidence interval for was $CI = 2.7140, 2.9422$. Again, the difference between pretest and posttest mean was statistically significant, $p < .001$.

Halfacre et al. (2005) reported sizeable mean score differences with high significance levels, and used consistent instrumentation in their pretest/posttest

data. However, they offered no reliability scores on the instrument they constructed making the internal validity of the findings uncertain. The authors also worked with a very small sample size in which 78% of their students were at the 5th or 6th grade level, and a small remainder were at the 7th and 8th grade level. The difference in maturity between the ages of 11 and 14 is great and could have significantly affected the measure of the outcome variables. The study did not include a control group, and the authors make no mention of the effect of maturation on the findings. The sample suffered a high mortality rate of 26%, also with no explanation. In terms of the external validity of their research, the authors provided no socioeconomic data or ethnic demographics to suggest how representative their results are for other populations.

The findings of Perry and Katula's (2001) literature review suggested there was a balance of positive and null hypotheses, while this critical review of the most recent literature suggests the balance is weighted toward positive findings. Rutter and Newmann (1989) and Lakin and Mahoney (2006) found no difference in service participants' civic outcomes, and Lee et al. (2007) and Yamauchi et al. (2006) found no difference in their civic awareness and attitudes, but these null findings are small compared with the positive claims made in the majority of the research. Yates and Youniss (1998), Kahne and Sporte (2007), Kielsmeier et al. (2004), Schmidt, et al. (2007), and Lee et al. (2007) found that for service students there were increased levels of civic engagement, participation, knowledge and efficacy. Melchior and Bailis (2002) found an increase in participants' service leadership, and with Metz and Youniss (2003) the findings

suggested an increase in youth's propensity to volunteer, especially for those who were initially less-inclined to serve. McGuire and Gamble (2005) and Halfacre et al. (2006) both indicated increases in social responsibility, particularly when participants' psychological engagement with service was high. Finally, authors such as Morgan and Streb (2001), Billig et al. (2005), and Terry (2000) suggested there are program characteristics that enhance the civic outcomes above, to include student voice, program design and planning, and student ownership of their learning.

Academic Achievement

With the pressure on schools today to raise students' academic performance, it is not surprising that greater academic achievement is the other primary outcome service learning advocates claim for participants. Research from the next six articles provides data on the significance of the relationship between service and academic measures. The authors' assumptions of what constitutes academic achievement can be limited to solely grade point averages, attendance, and reading and writing levels, as in the first two studies by Kraft and Wheeler (2003) and Magarrey and Francis (2005). However, some authors expand the definition of academic success, to include student perspectives of their own scholastic competence, as in the third study by Scales, Blyth, Berkas, and Kielsmeier (2000). The following article by Scales, Roehikepartain, Neal, Keilsmeier, and Benson (2006) examined the perspective of school principals on the importance of service to their students' academic performance, as well as student data on the achievement gap between low- and high-SES students.

Research by Roberts and Moon (1997) and Jensen and Burr (2006) are the final studies featured in this section, and they focused specifically on service-learning's potential to engage students with content area knowledge.

In a mixed-methods case-study of a charter school, the Youth Learning Cooperative (YLC), Kraft and Wheeler (2003) researched whether service-learning increased students' presence, participation, and academic performance at school. The authors defined presence as students' annual average daily attendance and participation as students' willingness and active engagement in project-based service-learning, their positive contribution to the school climate and their demonstration of good citizenship, attitudes, and skills. The authors defined academic achievement as growth in the core academic areas that was either greater than or equal to one grade equivalent, overall GPA, and advancement to the next grade level. The case study sample included 40 disaffected, behavior disordered, and at-risk high school youth from six rural school districts. The male to female ratio was three to one, and the majority of the participants were Caucasian, and between the ages of 17 and 18. Fifty percent of the sample came from families living in poverty. The authors selected YLC because of its highly integrated service program based on Dewey's philosophy, and the chance to observe its cognitive effects on a traditionally difficult student population. They stated a particular interest in studying resilience within youth, as put forth by meta-learning theory and Marzano's conceptual framework of the self-system.

The authors used the Six-Trait Writing Scale, the Reading Level Indicator created by American Guidance Group Assessments, and comparisons of GPA and course grades in core academic areas. They determined student participation through a Service-Learning and Locus of Control Survey, Student Climate Survey, student interviews and reflections, and documentation of service-learning experiences. Finally the authors measured student presence through attendance records, and the Quality of Service-Learning (QSL) Experience Rubric. Kraft and Wheeler ranked students into three levels on this rubric, with QSL I being the highest quality of service learning and QSL III the lowest. Using the rubric scores for these rankings, the data was analyzed separately for each group to determine the correlation between the quality of their service-learning experience and the participants' presence, participation and performance.

The post-assessment results of the study demonstrated that the QSL I group had a mean GPA of 3.30, compared to the QSL II group's mean GPA of 2.83, and the QSL III group's mean GPA of 1.36. For the QSL I group, attendance was 88.58%, compared with the QSL II attendance rate of 82.69%, and the QSL III rate of 73.5%. The writing assessment revealed a statistically significant increase in mean scores from 16.55 to 20.17 on the Six-Trait Writing Scale. Students performance on the Reading Level Indicator was also statistically significant, with a mean score increase from 28.46 to 30.88. Students' independent reading level also increased from 5th to 6th grade, and instructional reading levels jumped from the 9.5 to 12.2. The authors found a high correlation

between service-learning and resilience for the majority of the students, with 81% reporting that service-learning had enabled them to form caring relationships with adults, and 72% reporting they felt they had something valuable to offer others.

The authors used triangulation to develop credibility in their qualitative research, and did not claim an objective reality. History is properly accounted for with the use of pre- and post-data; however, the sample size they work with is not consistent on each outcome measured. The authors acknowledged that there were 24 students enrolled at the beginning of the school year, but the school population almost doubled, and some students were only enrolled on a part-time basis. Therefore the actual sample size for many of the measurements was quite small (n = 24). QSL I, which demonstrated the most positive results, was only a small subset of this reduced sample (n = 12).

While Kraft and Wheeler (2003) do not directly claim a causal relationship between service and academic performance, they strongly suggest it in their conclusion through phrases such as “seems to be attributable to” and “one can surmise” (p. 234), even though establishing causality in not possible with the data they provide. The internal validity of the quantitative analysis in this study was weak due to the lack of statistical information. Only percentages and raw data scores were reported, and though they were claimed as statistically significant, there was no mention of what type of analysis was conducted, nor were any *p*-values provided. Additionally, no alpha scores were reported for any of the instrumentation creating doubt around their reliability as effective measurement tools. There was no control group, so maturation is not accounted for, and host of

other confounding variables could interfere in identifying service as the primary independent variable in the measured outcomes. The length of participants' enrollment in the school is an important interaction variable with the positive outcomes, and without controlling for its effects it is difficult to isolate the impact of service. Since the participants in the study were enrolled at a charter school, and were primarily rural, Caucasian students, the external validity of this study is also negligible for many diverse, urban, public schools. Interestingly, the authors failed to explore the significance of the 50% poverty level of the sample, and they did not discuss how this figure was calculated or what the correlation of the students' socioeconomic background was with their quality of service and academic background.

Attempting to solidify the link between students' grades and service learning, Magarrey and Francis (2005) led a national survey to explore this relationship and found a positive correlation between the two. Using a Hotbot search engine, they reviewed 222 school websites for wording indicating student participation in service-learning. Over a yearlong identification process, the authors contacted schools in 37 states to be part of the study; 24 schools responded, representing 15 states and a student population of 2,925. The profiles of the schools in the sample were an even distribution of community service-learning programs, community service programs with no academic component, and schools with no service program.

The authors defined and measured academic success by student GPA, with data collected specifically for students with grades from the top 10%, median

10%, and lowest 10% in grades 6-12. They ran the data through a single ANOVA to determine the significance of the interaction between students' GPA and their participation in service, and used the data from the non-service schools as a baseline.

Magarrey and Francis (2005) established there was a mean increase in GPA for students in all categories between grades 6-12 across a period of six years. Non-service students' mean GPA rose .126, while students participating in community service increased their mean GPA by .159, and service-learning students increased .352. For students with GPAs in the top 10%, there were no significant differences between service participants and the control group. In the median 10%, only service-learning participants saw a significant increase in mean GPA, from 2.5 to 2.75. For students in the lowest 10%, service learning participants mean GPA rose from 1.35 in grade 6, to 2.10 in grade 12, compared with little change for community service or non-service students.

A strength of this work is the clarity with which the authors establish criteria for community service-learning, emphasizing the connection to academic content in the classroom. This creates more precision in the findings, and the study offers a useful comparison of community service learning with straight community service, an area lacking in the empirical research. Their collection of data over a six-year period also allows a valuable longitudinal perspective, with a very high sample size. The authors fail however to acknowledge any of the limitations of their research. They assume that each student experiences the same history over their six-year study, which is highly unlikely given geographic

variations in school populations. The external validity of the study is low due to the lack of demographic data on the participants. The authors created their sample from an internet-only search to qualify service programs, which includes only school sites with the staff and resources to create and update a website. This does not create a random sample, and could easily lead to an overrepresentation of wealthier schools with greater technological resources; however, this is hard to conclude without any socioeconomic profile data. In Magarrey and Francis's (2005) reporting of the quantitative data, they did not include the statistical significance of the findings, simply claiming their results either are or are not significant, which limited the internal reliability of the study. In the discussion of their findings, there is no accounting for individual student characteristics, or if program participants were self-selected, what type of service they performed, and if it was mandatory or elective. These are all variables frequently accounted for by other researchers in regression analysis, and thus Magarrey and Francis's methods seemed overly simplistic in the face of confounding variables that are widely recognized. Finally, the authors provided a very limited definition of academic success, as reported solely by GPA, without discussion of the varying cognitive abilities and skills that are not captured by grading.

The next study examined a much broader set of data to determine academic achievement. Scales, Blyth, Berkas, and Kielsmeier (2000) employed a quantitative, nonequivalent control group design to examine whether service-learning had a positive impact on middle school students' (N = 1,153) academic

success and social responsibility. The mean age was 12 years, with an equal distribution from each grade and gender. Seventy percent were white, 15% African-American, 4% American Indian, and 11% biracial. Seventy-five percent of the participants had completed service-learning with a reflection component before.

The authors conducted a national survey to find schools that had service learning programs of a certain quality, defined by number of participants, years of experience, enthusiastic teachers and administrators, integration with a required class, and ability to provide a control group. Forty-one percent of the recommended schools completed a pre-screening and interview. In all schools, the authors assembled participants into teams to achieve a balance of gender, academic performance, and ethnicity. The schools then determined which teams would be service-learning teams and which would not. Control teams had to agree to not use service-learning in the classroom, and those who were more comfortable forgoing the treatment most often became part of the control group. However, some teachers who were strong service-learning supporters also became part of the control group simply to support the research.

A single survey entitled the Survey of Middle School Student life tested for the following measures: social responsibility, personal development opportunities, parent involvement, commitment to class work, engagement with school, perceived scholastic competence, intellectual achievement responsibility, evaluation and mastery goals, academic success, and conduct.

Measurements of social responsibility used three subscales from Conrad and Hedin's Social and Personal Responsibility scale, with a reliability of .83 at the 7th grade reading level. Commitment to class work was measured with a subscale from Epstein and Mac Partland's Quality of School Life Scale, with a reliability of .80. Engagement with school was measured on a four-item academic engagement scale developed by Lee and Smith, using National Educational Longitudinal Data, with a reliability of .64. Perceived scholastic competence was measured using ten items from the Harter Scholastic Competence Scale from the Self-Perception Profile for Children, with reliability of .80 through .85. Finally, academic success was measured with computed GPAs for all subjects for each student during each marking period in the 1996-97 school year, and recalculated to make different schools grading systems comparable on a 13-point scale.

Scales et al. (2000) administered the survey in a single class period as a pretest and posttest, and they did so with student identification numbers that were not linked to student names by teachers. A series of ANCOVAs were conducted to compare treatment and control groups on all dependent variables. Data analysis revealed a serious compromise in which some control students experienced service, and 20% of the treatment students had not experienced service. To correct for this error, the authors reconstructed an uncontaminated sample group (n = 561) to accurately account for service-learning and control groups in comparisons.

Two significant posttest differences were found between the uncontaminated treatment and control groups: service-learning students were

more concerned with the welfare of others than control groups ($p < .01$), and they maintained their concern for others' welfare, whereas the control group's concern declined. Service-learning students also talked more frequently with parents about school than control groups ($p < .01$).

In terms of differences based on exposure and reflection, service-learning students who had done more than thirty-one hours of service had significantly higher posttest scores than control students in their perceived efficacy of helping others ($p < .0001$). Service-learning students who did "a lot" of reflection were more likely to perceive school as a place that offered personal development opportunities ($p < .0001$), and were more committed to doing their class work ($p < .0001$). No significant effects of service-learning were observed on school engagement, perceived scholastic competence, intellectual achievement responsibility, GPA, or conduct at school.

Scales et al. (2000) acknowledged that the sample is not representative. The percentage of parents with college degrees—56%—is more than twice the national proportion of 24%. Also, the 75% of participants with prior service-learning experience is a figure much higher than the 35% of students covered by service-learning requirements in the nation's largest 130 school districts. A chi-square test, however, indicated that there were no significant differences between service-learning and control groups based on socioeconomic or other background variables, and therefore there was minimal confounding of the results with those variables. Scales et al. also used an ANCOVA to compare the treatment group to the control group on dependent variables, and account for

some of the variance. Student surveys were confidential, and an outside party processed and analyzed them, suggesting that students had greater comfort in responding truthfully with anonymity.

Scales et al. (2000) acknowledged that participant schools did not truly meet all the criteria desired for service-learning quality. Despite screening, teachers and students suggested that the programs were not always graded or required parts of the curriculum and were “extras,” and their duration and scope was not as extensive and academically rigorous as desired. There was also relatively little reflection or preparation offered. The authors stated that this diminished the chances of service-learning having more significant effects. The differing types of service experiences and academic activities used to link them to classroom content were not consistent, so they acknowledged there are still confounding variables to control for. The greatest detriment to the quality of this study was the severely compromised sample, which led to a mortality of almost half. Additionally, the author’s did not assign student teams randomly, and teachers with preference or aversion to the treatment could influence their students’ participation.

In a pre-experimental survey, Scales, Roehikepartain, Neal, Keilsmeier, and Benson (2006) pursued the perceived impacts of service-learning programs, the relation of service to achievement gaps, and the effects of longer-term service on participants, with variable findings. The authors worked with three groups of participants, the first of which was the same sample of school principals (n = 1,799) examined in the Kielsmeier et al. (2004) study from the

previous section. In this work, the authors offered a more in-depth analysis of that data, and stratified the participants by their school setting, instructional level, average class size, and minority status. The sample split respondents almost evenly between the elementary and secondary levels. The second group of participants was an aggregate of middle and high school students (*n* = 217,000) from 300 U.S. communities that administered the Search Institute Profiles of Student Life: Attitudes and Behaviors Survey (PSL-AB), and the authors defined 21,883 students as coming from low-SES families. The last group (*n* = 5,136), 6th through 12th grade students from a racially and economically diverse community in Colorado Springs, also participated through their PSL-AB results.

For the data collected from school principals, 2,002 schools were selected from 2001-2002 Common Core of Data public school universe file. Principals were asked whether they thought the impact of service-learning on ten different academic, social development, school climate, and community relations outcomes was “very positive,” “somewhat positive,” or “little or no positive impact.” For each analysis, they calculated the mean “positive impact” score for each of the individual outcomes. In analysis of the principal responses, the authors created three school-level poverty groups at 33% cut points based on percentage eligibility for free or reduced-price lunch.

Scales et al. (2006) measured the data provided by students from the PSL-AB for forty developmental assets, ten risk behavior assets, five developmental deficits, and eight thriving behaviors. They weighted the student sample by race, ethnicity and urbanicity proportions of the 2000 Census to make

it nationally representative. They measured service to others with a single item asking students how much formal volunteering they do in a week, and choices were numerical. The Colorado Springs sample (n = 5,136) was asked how long they had participated in community service or service-learning in the last school year. Students with a few weeks or more were considered to have high levels of exposure. Students were asked to self-report their grades, and how many days of school they had skipped during the past month.

The PSL-AB measured five assets in categories designated as Commitment to Learning. Three items measured achievement motivation (α = 0.64) and four items measured school engagement (α = 0.63); the items include: achievement motivation, school engagement, bonding to school, homework, reading for pleasure, attendance, and grades. The researchers used a series of ANOVAs and Tukey post-hoc tests were performed when overall F tests revealed significance above 95%. Student groups were compared on grades, attendance, Five Commitment to Learning assets according to SES, school poverty, and service-learning exposure.

Principals of urban, high-poverty, or majority non-white schools were significantly more likely than other principals to judge service-learning's impact on attendance, school engagement, and academic achievement as "very positive." Principals of low-SES schools with a majority of African-American students reported the most positive impact of service-learning on academic achievement, compared to schools with predominant populations of other ethnicities ($p < .0001$).

The findings suggested that community service seemed to significantly lessen the achievement gap between high and low SES students ($p < .0001$). High-SES students had the highest means on the five Commitment to Learning assets; however, the low-SES students who had completed service had more of these assets than low-SES students without service, and higher or equivalent levels of academic assets than high-SES students without service. Low-SES students who had completed service reported fewer days missed of school, and higher grades. Students who reported a “few weeks” or more of service learning demonstrated higher results on the Commitment to Learning assets than all other students; however, not significantly higher than those with a few-hours of service-learning. In comparing the gap between high-SES and low-SES students in each of the assets, service-learning narrowed the differential for attendance, achievement motivation, school engagement, reading for pleasure, and most notably, bonding to school ($p < .0001$). The only gap not significantly lessened in comparing the two groups was the measurement of homework and self-reported grades.

The large sample sizes and low *p*-values ($p < .0001$) are strengths of the analysis by Scales, et al. (2006). They conducted ANOVAs, which help account for variance in groups that crossed high and low SES levels with high and low levels of service-experience. The alphas scores the authors reported on their instrumentation do not provide sufficient reliability ($\alpha = .63 - .64$). No comparison or context of the quality of students’ service projects is considered, so the data cannot distinguish between community service and service-learning.

Confounding variables remain, such as the propensity of high-achieving, academically motivated students to complete service. No causal relationship is possible, and more longitudinal data, as well as a stronger quasi-experimental design is necessary.

The next two studies examine academic achievement through the classroom lens of content area knowledge. Roberts and Moon (1997) focused on the relationship between community service-learning and content area knowledge acquisition with a mixed-methods action research project. A secondary goal was to identify if there were certain demographic characteristics of participants that made them more receptive to the cognitive benefits of service-learning. The authors worked with participants (N = 78) from a Midwestern suburban high school, drawing from three sections of the same teacher-researcher's intermediate Spanish course. Given two choices for independent study, control group students chose to complete a traditional book report (n = 12), while the treatment group students (n = 66) elected to create bilingual reading materials as a community service-learning project.

Over the course of a semester both groups were expected to learn how to narrate in the past-tense, and received similar assignments relating grammatical concepts to the application expected in their summative assessment. All students were required to keep reflective journals, and they all received the same classroom instruction, homework, and amount of class time to work on their projects. When the projects were completed, students were given a discrete point grammar posttest to assess their progress toward the learning objective. The test

included a subjective section in which students applied their narrative skills, and both groups were evaluated with identical criteria. Students' test scores were compared with a baseline score, which the authors calculated from the mean of three previous test scores. The authors conducted an analysis of variance between the mean scores of the control and treatment group. The authors also collected journal entries and anecdotal data for evidence of critical thinking and problem-solving demonstrating links between the dependent and independent variables.

The treatment group's mean posttest score was two points higher than the control group, but this was not a statistically significant difference. Nor was there a statistical difference of mean scores among the established demographic groups, except for high school seniors, whose scores dropped from 80% to 70%. In response to these findings, the authors explored how "academic gains" were defined and measured in relation to community service-learning outcomes. Despite a decline in posttest scores, a qualitative analysis of one student's journal, using the Dimensions of Learning Model, exhibited evidence of high levels of problem solving and reasoning through the completion of the project. Additionally, some students became consumed by the measure of their learning, and focused on getting a higher grade. One student became overwhelmed, and "spent more time worrying about what to do and how to do it than [he] did learning Spanish" (Roberts and Moon, 1997, p. 207). A pattern was established wherein some students became overly-focused on the aesthetic elements of the final product, neglecting the academic content of the project. Students who

continued to perform well, or made notable gains, under the community service-learning methodology, demonstrated attention and investment to both their service and academic agendas, and sought out the assistance of a more-capable peer or teacher in proofreading and developing their project. Additionally, students with academic gains from the service project either planned and/or implemented the sharing of their written work aloud to its intended audience.

Roberts and Moon (1997) strengthened their research analysis by exploring more than one dimension to academic development, acknowledging that quantitative evidence demonstrating students' knowledge and skills is not the only domain of students' cognitive growth. Their attention to service-learning's potential of developing higher-order thinking skills is noteworthy, but would benefit from the inclusion of more qualitative data. The authors do not acknowledge themselves as instruments in reviewing this data, and there is no mention of how data was coded, member-checking or triangulation in their qualitative analysis. The small sample size for the study was drawn from only one school, and no demographic data was provided, despite demographics data being one of the targeted areas of analysis. As a result, the external validity of the study is low.

The internal validity of the study is also low; maturation is accounted for with the use of a control group, but the control group is less than a fifth of the size of the treatment group. Control and treatment groups were self-selected, and there is no mention by the authors of how the school and sample for the study were selected. The instrumentation used in the study was not consistent from

pretest to posttest. The quantitative findings are not provided with mean scores and *p*-values. The published work also seems to be lacking a cited table, and there are large, identical sections of prose that appear twice throughout the conclusion of the article, which reflects poorly on its overall credibility, and the quality of its peer-review.

Continuing to explore how effective service learning is at engaging students in learning, and helping them learn course curriculum and content objectives, Jensen and Burr (2006) led a qualitative case study focused on a high school construction technology course. The high school was located in a lower middle-class, agricultural community of 14,000 people in the rural Northwest. Participants (N = 25) were of ages 14 to 18, and mostly male, with fifteen Caucasian students, nine Latino students, and one Asian student. A third of the students were either English Language Learners (ELL) or had an Individualized Education Program (IEP). The course was an elective for most students, but some were assigned the course by their school counselor.

The authors conducted the study during a two-week period, and collected data informally using pre- and post-surveys, rubrics, interviews, peer evaluations, and observations. At the beginning of the trimester, the teacher-researcher gave students a choice to learn the safe operation of the laboratory machinery: the traditional lecture and safety test, or a service-learning project. Students voted and were in consensus to do the project, so they then discussed in small groups what type of project they would like to do. Once they had identified the local elementary school's first grade class as the community they would like to help,

they visited the school site and interviewed the elementary students on what they would like for Christmas. The class discussed the findings of the survey, and made a design and manufacturing plan for 90 wooden toys that they would create and deliver for Christmas. The teacher used a “just-in-time” instructional method, and would train students in technological skills they needed as they came to tasks that required them. Students worked under a deadline, forming task groups and working cooperatively.

To determine the effects of the service-learning project on the participants’ mastery of content objectives, the teacher formally interviewed the students twice during the study on their knowledge of when and how to operate the laboratory machinery. He also made daily, informal observations, and assessed them as they completed the project. Seventy-six percent of the students were proficient in the safe operation of the machinery, and the results were equivalent with a previous class’s performance on the same unit under traditional methods. The teacher measured students’ perceptions of confidence with an eight-item pre- and post-survey, which asked students to rank their measure of confidence on each piece of machinery on a Likert-type scale. Students demonstrated a mean increase of one to two points on each item, which matched the noticed increase in the teacher’s student observations. Students did not exhibit a sense of increased content knowledge in a comparison of their mean survey scores.

The author measured commitment according to participation in class and amount of out-of-class time spent working on the project, as well as student attitudes and feelings concerning their work. The author found that 23 of the 25

students were committed to the service-learning project, compared with a previous class's engagement of 12 out of 24 under traditional methods. Within the service-learning class there were varying levels of commitment, and the author ranked 13 students as highly committed, 6 as moderately committed, 4 with low commitment, and 2 with no commitment. Jensen and Burr (2006) made a positive correlation between level of commitment and perceived confidence and content knowledge, and participants with the highest levels of commitment made the greatest gains on those measures. The two students with no commitment decreased their mean scores on the post-survey. The teacher verified the student perceptions with his own observations and interview data.

Although Jensen and Burr (2006) claimed this was a qualitative study, they presented mainly quantitative findings without statistical analysis. The survey instrument the authors constructed is described with no details on its reliability. Even as a quasi-experimental study, there is no formal comparison group. The authors frequently make comparisons to a former class sample, but this is not done systematically, and there is no discussion of potential significant differences between the two groups of students. While the chosen sample was representative of this teachers typical shop class, it's predominantly male population, ethnic composition, and the specific content focus of the course are not transferable to broader populations, which Jensen and Burr acknowledge. In terms of the qualitative analysis, supposedly the framework of the study, there was little discussion of how the authors coded and analyzed the interview data. The teacher-researcher collected the qualitative data exclusively, limiting the

confirmability of the findings in the absence of an outside evaluator. Despite the authors clear bias in this study, he did not acknowledge self as instrument, and offered no evidence of member-checking. One strong point of the methodology however, was the triangulation between interview, surveys, and observation data.

Overall, the findings relating academic achievement and service-learning are positive, but not conclusive. Scales et al. (2000) found no effect on students GPA or perceived scholastic competence, and Magarrey and Francis (2005) only found a significant increase in GPA for service students beginning with low scores. Kraft and Wheeler however did notice that participants of high quality service programs had higher GPAs, attendance, and reading and writing scores. Roberts and Moon (1997) quantitative analysis failed to demonstrate a quantitative difference in student achievement. The two analyses of principal data by Kielsmeier et al. (2004) and Scales et al. (2006) revealed the majority of principals felt service has a somewhat positive effect of their students academic performance, but only a third of respondents felt that service was highly important. While evidence demonstrating numeric gains in academic achievement was low, a number of the studies indicated greater cognitive growth by other means. Scales et al. (2000) and Jensen and Burr (2006) saw an increased commitment of students in class, and students demonstrated proficiency and confidence in applying their course content learning. Robert and Moon (1997) and Terry (2000) observed students developing higher-order thinking skills such as problem solving and reasoning though their service programs. Scales et al. (2006) observed the potential of service-learning to

narrow the differential for attendance, achievement motivation, school engagement, reading for pleasure, and most notably, bonding to school between low and high-SES students.

Personal Development

Adolescence presents a sense of crisis and confusion to the adolescent as they explore who they are and shape their identity (Nakkula, 2003). Equally important as becoming a productive citizen and good student, is developing a sense of self, and learning to control emotions and behaviors. This section identifies a small portion of the research that focused particularly on personal development. The findings of this section are not exclusive, and the previous sections have overlapped heavily through discussion of empowerment, efficacy, and competence. The first study, by Johnson and Notah (1999) looks to the outcomes of service on self-esteem and personal responsibility. The second and third studies look at how service can deter harmful behavior, with O' Donnell, Stueve, San Doval, Duran, Haber, Atnafou, Johnson, Grant, Murray, Juhn, Tang, and Piessens (1999) examining risky sexual behavior, and Hoffman and Xu (2002) researching delinquency. Conclusions from other sections will be drawn upon to offer evidence that service also promotes greater career development and resilience in youth.

In a mixed-methods, one-group pretest/posttest design, Johnson and Notah (1999) examined the effects of community service-learning on self-esteem and personal responsibility. The participants were 56 middle-school students from an urban environment, and 98% of the participants were Hispanic.

The study took students from one author's advisory class, and the other author's six science classes. The authors randomly assigned students to the classes by a computer system. They required students in these classes to complete one service project, which they integrated with their curriculum, and they required students to work either alone or in small groups during a nine-week period. Students first held a brainstorming session to identify community needs and project ideas, lasting from one to four days, depending on when the group came to consensus. The next phase, of six to ten days, all students planned and prepared for projects, and arranged contacts and project implementation. All students completed daily reflective journals in which they discussed thoughts and feelings pertaining to their project. Students completed three types of service, including direct, indirect, and advocacy. Service projects implementation ranged from 2 to 40 hours depending on the project.

Johnson and Notah (1999) collected pretest and posttest data on self-esteem and responsibility using the Coopersmith Self-Esteem Inventory and the Junior Index of Motivation Scale. The Coopersmith is a 58-item instrument to which students respond "like me" or "unlike me." The Junior Index contains 80 items to which students respond they strongly agree, agree, disagree, or strongly disagree. Each advisory student was given both instruments, and the science students were randomly assigned one of the instruments. Johnson and Notah used a *t*-test to compare the mean pretest and posttest scores on each instrument for population as a whole, and by gender and types of service. Changes in self-esteem indicated by the Coopersmith measure were not

statistically significant, but modest changes of two points in the mean scores were evident. The changes in responsibility indicated by the Junior Index did not differ significantly but the mean scores increased 1.5.

Other data collected from both classes included reflective journals, concluding narrative essay, student interviews, and field notes with anecdotal records and observations. Additionally, the authors randomly selected a group of 17 students to be interviewed. Johnson and Notah (1999) coded data for indicators of personal responsibility and self-esteem, and the qualitative data indicated substantial growth on both measures. No negative responses toward service were recorded. The qualitative results contradicted the quantitative results, and authors attributed this to participants' language barriers and challenging home lives.

The authors created strong triangulation of the qualitative data, but they provided no information on the way the data was evaluated and coded. Since the qualitative findings were not consistent with the quantitative results, and were not evaluated by an outside party, their confirmability is suspect. The study benefited from a fair sample size, in which students were randomly selected in non-elective coursework. The participants were almost all Hispanic however, creating findings that cannot be generalized to other populations. In terms of the service program's quality, the actual time spent doing service work was small and extremely variable, lasting between 2 and 40 hours. Students' planning of the project also created confounding variables. The researchers' use of a pre-experimental design did not provide a control group to compare the findings against and

account for maturation, and they gave no reliability levels for the testing instruments.

To determine the effectiveness of a Community Youth Service (CYS) program in reducing sexual risk behaviors for middle-school adolescents, O' Donnell, Stueve, San Doval, Duran, Haber, Atnafou, Johnson, Grant, Murray, Juhn, Tang, and Piessens (1999) performed a quantitative, nonequivalent control group design. They selected participants (*N* = 1,157) from two urban, public schools in Brooklyn, NY, serving economically disadvantaged students. The participants came from a total of 68 classrooms, of general and special education students, and were roughly balanced between males and females. The study targeted 7th and 8th grade students, with an average age between 12 and 13. The sample was predominantly African-American, and 16% identified as Hispanic, and 5% as other. Of the 1061 participants completing both surveys, 255 participated in CYS intervention, 222 in curriculum only intervention, and 584 served as controls.

O' Donnell et al. (1999) selected the intervention site because it was where the CYS program originated, and a similar comparison site school was selected as the control group. School sites were selected by the New York City Public School Central Board for having a large, almost exclusively minority student body ($N > 700$), 99% African-American and Hispanic. The schools chosen were also deemed as having high-risk health profiles, and high-risk academic profiles.

In the intervention school, classrooms (*n* = 35) were randomly assigned to receive either the core Reach for Health curriculum, or the curriculum enhanced by CYS. Students in the CYS program spent three hours per week in a community service-learning placement. Placements included four nursing homes, one neighborhood full-service health clinic, two childcare centers, and one senior center. Back in their health classes, students shared experiences in debriefing sessions used to reinforce skills in decision-making, communication, information seeking, and health advocacy.

Specific CYS assignments varied by grade; 7th grade students worked primarily at childcare centers; 8th graders received broader placement. Over the course of the school year, students were assigned to two different field placements. The 8th graders also received orientation lessons to prepare them for work in health care settings. Teachers accompanied students to placement sites, as well as nursing students or faculty from Medgar Evers College. From a total of 35 classrooms, 13 were randomly assigned the CYS condition, determined by time and logistics of escorting students. The Reach for Health curriculum was given to all 7th and 8th graders with or without CYS condition. It consisted of 40 core lessons per year, for both grades seven and eight. Teachers were trained in curriculum and observed during their teaching. Twenty-two classrooms were assigned the curriculum only condition.

Youth were asked four questions about sexual experience at baseline and follow-up, placing them on a continuum of risky sexual behavior. All items were measured through self-report pencil and paper questionnaires. Ethnicity was

measured through four separate items, and responses were scored into three categories. Percentages of youth reporting sexual behavior were generated for both intervention groups, and control groups, at baseline, and a six-month follow-up. O' Donnell et al. (1999) calculated delta scores from the baseline percentage minus the follow-up percentage, and reported the difference along with the baseline percentages to demonstrate aggregate changes over the year, in sexual behaviors, by treatment condition for the total sample and subgroups. Measures were taken for "ever had sex," "recent sex," "recent sex without condom," and "recent sex without birth control." O' Donnell et al. (1999) used MIXOR, a computer program for mixed effects regressional analysis, to take into account the effects of clustering.

The findings demonstrated reports of both lifetime and recent sex across all groups were higher at follow-up than at baseline, higher among 8th than 7th grade participants, and higher among males than females. For the whole sample, delta scores were slightly higher for students in the control condition than for curriculum only or CYS students on each of the outcome measures, meaning that they had larger increases in risky behavior. For curriculum only students the delta for risky sexual behavior for the whole sample was a decrease of -12.7, for without condom, and -5, for without birth control. For CYS participants in the same outcome measures, the delta was a decrease of -15.9, and -8.4. The greatest decrease in risky sexual behavior was for the 8th grade participants, with a delta of -30.8 and -27.7. Rates of risky behavior thus increased for the control group, but decreased for the curriculum and CYS groups. Students in the CYS

group were significantly less likely to report recent sex (p = .05), which indicated that the CYS program was most effective at reducing recent sex in 8th graders. Curriculum only instruction showed a positive effect, but with less statistical significance (p = .08).

The strengths of this study include a large sample size that targeted low-income, minority students, and a strong experimental design with randomized testing and a control group to account for history and maturation. The authors also made an appropriate choice not to equate the positive impact of the curriculum with a decrease in sexual behavior, but rather, they emphasized the importance of condoms and birth control in preventing risky sexual behavior. With a retention rate of 91% of the sample over a six-month period, there was not significant attrition to distort the results. The limitations of the study include some lurking variables, such as student age and the intensity of their service experience. In determining what age level the CYS program is most effective for, the authors were uncertain to what extent the broader, more intensive service placements of 8th graders had affected the measured outcomes, and the larger percentages of 8th graders that are sexually active resulted in greater delta change. Finally, the authors of the study acknowledged that there was a strong community partnership in place with the research team, and without such collaboration the success of this program might not be transferable.

While O' Donnell et al. (1999) focused on the potential of community service to decrease undesirable sexual behaviors in adolescents, Hoffman and Xu (2002) explored if it could be used to reduce delinquent behavior. They

collected data for correlational research from the 1992 National Educational Longitudinal Study, with a national sample of 12th grade participants (*N* = 11,560). Public schools (*n* = 10,186), private schools (*n* = 688), and Catholic schools (*n* = 736) were all represented. The sample was closely balanced between males and females, and included 8,034 Caucasians, 1,248 African Americans, 1,295 Hispanics, and 983 participants of other ethnicities.

The NELS began data collection on the participants in the 8th grade, following them from 1988 to 1992 with a 90% follow-up rate. The NELS (α = .80) relied on self-reported data to measure participants delinquent behavior, with items concerning the frequency of physical fights, suspensions, disciplinary transfers, arrests, and time spent in juvenile detention centers.

The key exogenous variable Hoffman and Xu (2002) measured for was involvement in community activities. They used a 6-point composite measure scale that assessed involvement in volunteer or community service, with responses coded as yes (1) or no (0), with a mean score of .63. The authors controlled for the effects of mandated service by court orders by coding it as a "no response." In response to the literature that suggests control variables such as school disorder, perceived lack of school safety, grades, work, and parent-child communication may affect involvement in activities, these items were measured and included in the analysis. Additionally, they included variables to account for individual characteristics such as gender, race/ethnicity, and socioeconomic status. And finally, the authors included the school-level variables

of percentage of minorities, student-teacher ratio, type of school, and school location.

Hoffman and Xu (2002) used a multilevel statistical analysis in four stages. They began their analysis with a simple variance components model, and calculated a mean level of delinquency (M = .215, SE = .0006) across schools, finding greater variation within schools (.214) than between them (.008). They then used a random effects model to allow the impact of activities on delinquency to vary across schools, to determine the degree to which the effects were uniform. In the third stage, they controlled for the individual and school-level variables to determine if the measured effects on the outcome remained. The final stage of the model was a cross-level interaction that explored which characteristics of the schools interacted with community activities to have the greatest effect on delinquency.

In the first model, the data indicated that community activities did have a significant negative impact on delinquent behavior (-.054, SE = .007, $p < .01$). After accounting for individual and school-level characteristics in the second model, a significant negative impact remained (-.020, SE = .007, $p < .01$). The study found the effects of community activities did not depend on certain school characteristics.

By accessing the NELS data, Hoffman and Xu's (2002) correlational study used objective data on a large, nationally representative sample. The NELS had a high reliability level (α = .80), which contributed to the internal validity of the study. The authors realized that the longitudinal nature of the NELS was limited

however, since data on participants' community activities was only collected in the 12th grade survey. Although they gave a thorough treatment to individual and school-level variables, they were careful not to assume causality. They admit that unobservable characteristics of the participants could have affected the results: students with more motivation and self-control might be more likely to participate in community activities and less likely to be delinquent. Also, the NELS did not collect socioeconomic data on the schools' local communities, which could have had a strong effect on students' desire to serve. The authors noted that the NELS did not collect data on community activities from students that had dropped out, which was relevant if those students had high rates of delinquency. Additionally, the delinquency data was self-reported, which created a potential bias, but the authors felt data from school records also contained equal if not greater bias, as well as error. In the context of the greater research literature, their findings are not highly reliable, because little research has been done to show similar results. Applying their findings to service-learning contexts is problematic because the authors do not distinguish between the types of service completed by participants. Therefore, "community activities" becomes a very vague exogenous variable, with low external validity.

Johnson and Notah (1999) concluded that service-learning has the potential to increase self esteem and personal responsibility, which coincides with Scales et al.'s (2000) findings that service students with high levels of commitment saw schools as places of personal development. Hoffman and Xu (2002) and O' Donnell et al. (1999) determined that service has a negative

impact on delinquency and risky sexual behavior, and these findings are consistent with Schmidt et al.'s (2007) negative correlation of service and behavioral problems. Addressed in previous section but pertinent here, Yamauchi et al. (2006) and Terry (2000) both suggested a positive impact of service on career development, in which youth have better networks, skills, and commitment to aid their future work. Kraft and Wheeler (2003) related this advantage from service as greater resilience, in which youth have developed caring relationships with adults and feel they have something valuable to offer others.

Service Design

The four articles featured in this final section reflect more on the structure of the service-program and instructional qualities of the teacher, in an attempt to better understand and predict outcomes for participants. A great limitation of service-learning research is the lack of clarity and consistency by educators and researchers in detailing whether students are engaged in basic community service, with no integrated academic component and structured reflection, or if they are completing community service-learning as defined in the first chapter. Furco (2002) addressed this with a comparison of the outcomes between students in community service, service-learning, and non-service programs. In the second study, McLellan and Youniss (2003) questioned whether there are different effects between mandatory and voluntary service programs. Finally, Seitsinger (2005) and England and Marcinkowski (2007) explored the

relationship between teacher attitude and service, and collected data on what participant outcomes teachers are attending to in the field.

Furco (2002) compared the effectiveness of service-learning with community service, and examined the effects of varying service designs on high school students' (*N* = 529) academic development, with a mixed-methods, nonequivalent control group design. He was unable to demonstrate a causal relationship between types of service and impact on students, but did identify recurring outcomes among the various service programs. Student participants came from large California schools with diverse ethnic populations. Furco also included teachers (*n* = 24), service coordinators (*n* = 3), site administrators (*n* = 2), and community agency representatives (*n* = 17).

The one-year study used 11 quantitative and qualitative measures to compare the educational development outcomes of three different categories of service programs: community service, service-learning, and service-based internships, on students' development across six educational domains. These domains were identified by the seminal studies of Goodlad, and Conrad and Hedin, and include the academic, career, personal, social, civic, and ethical sectors.

Furco (2002) selected school sites for the study that had offered the three types of established service programs for longer than two years. To be eligible for the study, classrooms and service programs at sites had to have clearly defined educational objectives for students, and have teachers who had been with the program for at least two years, with identified service placements and

structured reflection activities. The author formed the three experimental groups of community service (CS), service-learning (SL), and service-internship (SI) according to the following criteria: the program's primary intended beneficiary, degree of emphasis on service or learning, intended educational purposes, degree of service integration with curriculum, and nature of service activities. Furco selected ten additional classrooms comparable in grade level, subject matter, teacher experience, and student abilities as control groups with no service (NS). He included a total of 34 classrooms in the study, distributed as follows: CS=7; SL=9; SI=8; NS=10.

The author developed a 41-item attitudinal survey to measure student outcomes in six educational domains (α = .43 to .72). It included the Rosenburg Self-Esteem Test, the Pier-Harris Children's Self-Concept Scale, and a survey created by Conrad and Hedin. For the academic domain, the survey measured "changes in students' attitudes and motivations toward school and learning, understanding of relevance of academic content, and overall school performance;" in the social domain, the survey measured "changes in students' ability to work with others and attitudes toward those who are culturally and racially different;" and for the civic participation domain, the survey measured "changes in student awareness of societal issues and willingness to take on active roles in the community" (Furco, 2002, p.32).

Furco (2002) supplemented the survey with a series of qualitative instruments to form the Evaluation System for Experiential Education (ESEE). ESEE used a set of 11 data collection instruments, approaches, and protocols to

attempt to capture the particularities of each individual service program as well as the more generalizable data common to all. This included: a researcher designed pretest/posttest; student journals; semi-structured focus group interviews; content analysis of samples of student produced work; a student placement questionnaire; teachers' program goals and objectives; classroom site visits and observations; teacher focus group interviews; a teacher questionnaire; a community-based organization questionnaire; and formal and informal meetings with site administrators. Furco (2002) recorded data from these sources in meta-matrix cells to help identify the recurrence of themes across domains and groups with comparative analysis.

The study collected data over one year, and Furco (2002) administered a pretest to the treatment group (n = 158) and control group (n = 125) at the beginning of the school year prior to service activities. The author then administered a posttest to the same groups toward the end of the school year, once all service activities had been completed. Furco analyzed the data separately from each of the three different service type groups to test the null hypothesis on their impact on academic development. He used six ANCOVAs, one for each domain, to establish if there was a significant difference between those who had participated in types of service and those who had not. To control for initial differences, Furco used conditions such as gender, ethnicity, and the school site with grade level and students pretests domain scores as covariates.

Furco (2002) found that in the academic domain the adjusted mean of the service-learning group (M = 3.06) was significantly higher than the no-service

group (M = 2.86; $p < .05$). Interestingly, the adjusted mean of the community-service group (M = 3.05), was almost identical to the service-learning measurement, despite having no demonstrated academic component. In the social domain, the adjusted mean of the service-learning group (M = 2.93) was also significantly higher than the no-service group (M = 2.81; $p < .05$). In the civic domain, there was no significant difference between the different groups.

The qualitative analysis (n = 100) revealed contrasting results, which suggested that all groups, including the no-service control group, have the potential to produce positive outcomes, and the nature of the student is the most significant predictor of outcome, not whether or not they engaged in service. Based on randomly selected outcome statements, there were 91 positive responses from the no-service group in the academic domain, compared with only 82 positive responses from the service-learning group. The no-service group had 78 positive responses in the social domain, compared with 62 for the service-learning group. The only qualitative difference noted which suggested the service-learning group had experienced a greater positive impact was the civic domain, with 65 positive responses, compared to 14 in the no-service group. It is worth noting, that this is the inverse of the quantitative findings comparing service-learning to a non-service control group.

An obvious omission that warrants concern is that Furco (2002) provided no specific data on the demographics of the school or sample. Additionally, the schools were not chosen randomly, and focused only on treatment groups with established service programs older than two years. This creates confounding

variables that complicate the findings, such as teacher experience and program infrastructure, all of which weaken the external validity of the study. The reliability level of his 41-item attitudinal survey is also low (α = .43 to .72), which suggests the internal validity of the quantitative design was also weak. In the qualitative analysis, Furco was transparent in his coding and labeling of data, and his data is made more credible by the multiple perspectives he investigated. However, there is no evidence of his member-checking, which could have strengthened the results.

McLellan and Youniss (2003) collected data over eight years in a pre-experimental, comparative study of two school groups to determine if required service was as effective as voluntary service, and how the types of service completed differed in each arrangement. Their findings demonstrated that students were more likely to interact directly with those in need in a structured and supportive service program that was integrated with their academic curriculum. Their test sample (*N* = 783) included students of two private Catholic High Schools. Of the sample, there were 310 sophomores, 280 juniors, and 223 seniors. Fifty-two percent were female, and 48% male. The students came from mainly upper-middle income families, and 20% identified as African-American, Hispanic, or Asian.

Data was collected in a questionnaire form administered by teachers to students during a 55 minute English or Religion class. It was given at the beginning of the school year, and again at the end of the school year. Data was tracked on two cohorts of students, beginning in their sophomore year, and

continuing through their graduation. Over the course of the study, data was collected from 2,003 students. The data for the present study comes from one questionnaire completed in 1996, and it was selected because it provided the largest sample, and was representative of all the other collection times.

The two cohorts compared were from Schools A and B. School A required service that was integrated with their religious studies curriculum, and students had a service coordinator. School B required the same hours of service, but did not integrate it with curriculum and had no special staff support.

The three dependent variables in the study were: what type of service students did to fulfill their school requirement, whether students did voluntary service beyond school requirements, and what type of voluntary service they did. The types of service students pursued were divided into five categories: social service, working for a cause, teaching/coaching with the needy, teaching/coaching with the non-needy, and functionary work. Service was dichotomized for the study into service working for a cause and with those in need, such as low-income, inner-city or elderly service recipients, or functionary service to the non-needy. This type of service would include filing, sweeping, phone work, or working with affluent populations.

The results suggested that School A's structured service program students were more apt to do social service or teaching/coaching with people in need, where School B students were more likely to have done functionary service for those not in need. There were no differences found in the extra

voluntary service completed between the two. Both were twice as likely to do functionary service for the required component as compared to voluntary service.

A series of logistic regression analyses examined the data collected on required service type, presence or absence of volunteer service, and volunteer service type. Parallel logistic regression analyses were conducted to look at predictors of whether students conducted social service to meet their school requirements. A hierarchy of independent variables were used: parent education and gender; individual difference measure of propensity toward helpfulness; school attended; parent or best friend voluntary service type; level of involvement in community, church, or political organizations; and type of volunteer service. The two greatest predictors of variance in whether students performed social service for the needy, in both Schools A and B, was the structure of the service program and whether the students had previous social service volunteer experience, with R^2 values of 14% and 16%, and with p-values of .001 and .01. In the hierarchical logistic regression analysis of which students did volunteer service versus which did not, the predictors accounted for a cumulative 27% of the variance. All factors except parent education and belonging to a political or cause-related organization were significant predictors ($p \leq .01$).

The regression analysis is a strength of this study, and helps examine the background variables influencing students' propensity toward social service behavior, and isolate which predictors cause the greatest variance in probability. The research also examined a large sample size, with significant findings ($p \leq .01$). However, because the sample was taken from an affluent Catholic school

population, the results cannot be generalized to public K-12 schools. Additionally, the service requirements of School A were embedded in students' religious studies of social justice; therefore the structure of the treatment is not transferable to public school classrooms.

In further considering what constitutes a quality service-learning program, Seitsinger (2005) conducted a correlational study on teacher survey data from schools participating in the Project High Performance Learning Communities (Project HiPlaces). She examined what educational attitudes and beliefs were associated with service-learning and what the relationship was between teachers' educational attitudes and beliefs, service-learning, and standards-based instruction. This broader study collected data over the past two decades in 2,000 K-12 schools across the country, with an 80-90% staff response rate. Seitsinger drew data from a sample of core classroom teachers (N = 4,434) that participated during one academic year, and constructed her own sample with participants (n = 2,164) from 271 schools from that had completed survey items which assessed the 48 variables relevant to the research question. The educational background and experience of the teachers were representative of public middle school teachers nationally. The majority of the schools featured in the sample served students grades 6-8 (N = 143,877), and the ethnic composition of the schools included students that were 52% Caucasian, 15% African-American, 21% Hispanic, 4% Asian-American, 2% Native American, and 6% multiracial. Over half of the students served in the sample were eligible for

free or reduced-price lunches, and the percentage of schools that were urban, suburban, and rural were roughly equal.

Seitsinger (2005) collected data from the HiPlaces Assessment staff survey, and assessed the teacher's instructional practices with the Classroom Instructional Practice Scale (CIPS). The CIPS consists of 16 subscales, and, when scored for 82 items, has a high internal consistency ($\alpha > .95$). To evaluate service-learning, the author used the Community-Based Learning Opportunities subscale, which has eight items ($\alpha = .82$-$.92$). The items from this scale were independently mapped to the definition of service-learning by the Alliance for Service-Learning in Education Reform Standards (1993), with interrater reliability estimates in the .82 - .92 range. To assess teachers' attitudes and beliefs on traditional and reform-based educational practices, the author used the Attitudes Towards Educational Practices Scale (ATEPS), adapted from the CIPS, which contains 55 items on 13 subscales ($\alpha > .94$). Teachers reported on a Likert-type scale ranging from 1 to 5 on how strongly they agree that each practice is essential to effective education.

In her findings, Seitsinger (2005) reported that among the classroom practices teachers considered essential were citizenship, social competence, and critical thinking ($M = 4.28$; $SD = .53$); and community-based learning ($M = 3.91$; $SD = .77$). The author also ranked the participants into high (SL HI), moderate (SL MOD), and low (SL LO) levels of service-learning implementation and reported a correlation between the level of implementation and a difference in attitudes toward standards-based instruction. Teachers that implemented

service-learning the most frequently had significantly higher predispositions toward standards-based instructional practices than teachers with low levels of implementation ($p < .001$). The author compared the teachers' attitudes on cross-content area instruction (SL HI, $M = 4.60$, $SD = 1.03$; SL LO, $M = 2.76$, $SD = 1.04$), literacy instruction (SL HI, $M = 5.29$, $SD = .99$; SL LO, $M = 3.65$, $SD = 1.40$), and literacy: application and analysis (SL HI, $M = 4.12$, $SD = 1.23$; SL LO, $M = 2.48$, $SD = .96$).

Seitsinger (2005) did not imply a causal relationship between service-learning and standards-based instruction, but rather suggested that service learning is closely aligned with school-reform initiatives, and educational settings committed to engaging students in higher-order thinking. Her study is strongly supported with highly significant mean differences, and the instrumentation used carried high levels of internal consistency. The schools that provided data represented a wide diversity of students in terms of ethnicity, locality, and socioeconomic backgrounds. However, the authors acknowledged that the schools were all part of longstanding reform initiatives, and so the teachers classroom practices may not be representative of all middle school teachers, thereby weakening the data's external reliability. Additionally, teachers' self-reported attitudes on the importance and effectiveness of service opportunities did not include students' voice on the impact or way these practices are perceived.

Using Furco's (2002) service-learning outcome domains, England and Marcinkowski (2007) surveyed the nature, status, and effects of environmental

service-learning for secondary students, and added a seventh domain to Furco's evaluative framework—school participation and behavior. To identify secondary schools (*N* = 25) with established service-learning programs for the study, the authors relied on the Florida Learn and Serve K-12 and St. John's River Water Management District Legacy programs, as well as the Florida Department of Environmental Protection Agency in the Classroom and ParKnership Program. Fifteen of the sites were traditional schools, three sites were combined junior and senior high schools, and three were charter and magnet schools. The programs surveyed varied in their duration and standing, elective or mandatory status, scheduling, assessment, and service type (direct or indirect), reflective activities, amount of support students received, and the frequency with which teachers collected data on student outcomes.

The authors developed a quantitative survey instrument with 30 Likert-type items, as well as eight open-ended questions. They then entered the quantitative data into spreadsheets and constructed tables, and analyzed the content of the qualitative data. Only the study's quantitative results were published, and the relevant findings indicated the frequency with which schools (*n* = 16-18) collected data in the seven outcome domains.

With a maximum number of 100 "yes" responses, there were 60 "yes" responses for Academic Achievement and Success, followed by 46 "yes" responses for Civic Responsibility Development, and 38 "yes" responses for Social/Interpersonal development. This means that for civic and social development, the majority of secondary teachers in Florida leading service-

learning programs were not collecting any data on the outcomes in these areas. Teacher-reported program assessments, however, indicated student growth and learning in all seven of the outcome domains.

England and Marcinkowski (2007) acknowledged a significant limitation of their study was the data concerning the outcome domains was teacher-reported with no attempt to establish its reliability. However, with so few teachers collecting information in most of the outcome domains, the authors still report growth in all areas. Of the 59 schools contacted for the survey, less than half responded, which the authors also admitted compromised their attempt to get a system-wide analysis of service-learning in Florida's secondary schools. There is no mention of the internal consistency of their instrumentation, which leaves the question of how effective it was as a measurement tool. There is also no use of statistical methods evident in this research; all data is presented in raw, table form, and this strongly weakens the internal validity of the quantitative work. Unfortunately, the authors did not present any of the qualitative findings, or reveal any of their methodology in coding and analyzing its content. Again, this weakens the credibility of their findings. The study is at least useful in providing a sense of what some teachers are looking for as outcomes to their students' service, and demonstrates, unsurprisingly, that academic achievement is the highest priority.

Adding to England and Marcinkowski (2007) and Seitsinger's (2005) findings surrounding the interaction between teacher and service, Billig, Root, and Jesse (2005) bolstered the importance of the teacher's role in enhancing the

service program. They found that the more experience the teacher had, the greater impact she had on participants' civic knowledge and dispositions. This is reflected through Yates and Youniss (1998)'s conclusion that the interaction of service, class discussion, and reflection was essential, and surely dependent upon teacher facilitation. Additionally, Billig et al. found service programs lasting between one semester and one year were ideal to build civic knowledge and skills, and that student-selected service activities had the highest gains. Riedel (2002) found that teachers who structured participatory service programs—that are concerned with broader social concern, rather than personal virtue—are more effective at increasing students' feelings of civic obligations. McLellan and Youniss determined to best serve those in need, required service-learning is most effective.

While Furco (2002) studied the differences between community service and service-learning, Schmidt, Shumow, and Kackar (2007) examined outcome differences among types of service-learning placements, and found that students in direct service earned higher grades, and adolescents assisting organizations reported greater civic efficacy and civic knowledge. Billig et al.'s (2005) conclusions linked indirect service—in contrast to Schmidt et al.—with higher levels of academic engagement than civic or political action, but claimed direct civic or political service increased civic knowledge. Finally, Terry (2000) and Morgan and Streb's (2001) research on service design suggested that students benefit from greater voice and ownership in program planning and implementation.

Summary

This chapter served to critically review the most recent body of research surrounding community service-learning, particularly its effects on adolescents' civic, academic, and personal outcomes, and significant factors surrounding service design. The civic engagement section contained the wealth of community service-learning research, and indicated primarily positive effects. The research of the academic achievement section demonstrated limited findings on the quantitative gains for students as measured by GPA, attendance and test scores, but suggested the benefits of community service-learning surfaced in other areas. Higher-order thinking skills, and increased engagement, commitment, motivation, and pleasure in school were among these. The third section concluded that community service-learning enhanced personal development, as shown by improved self-esteem, personal responsibility, career development and resilience, and abated delinquency and unsafe sexual behavior. Finally, attention to service design suggested that the teacher, duration and type of service, and level of student ownership in community service-learning programs are significant predictors of outcomes. The fourth and final chapter will now outline the summary of the findings just presented, consider their implications for the secondary classroom, and conclude with suggestions for future community service-learning research.

CHAPTER FOUR: CONCLUSION

Introduction

Bringing youth back into the fold of civic society is a challenge educators face with a new generation disenfranchised by the American political process. Chapter One developed the rationale behind targeting this population with community service-learning as a way to promote greater civic engagement, academic achievement, and personal development. Service creates an educative experience for youth (Dewey, 1938) and allows students to complete the full learning cycle Kolb described (Pritchard & Whitehead, 2004). In service programs, students are able to discuss and hypothesize, as well as act and reflect on real problems their communities face. Immense developmental opportunities lie within the identity crisis adolescents experience. Reflection on their civic participation during this phase of advanced cognitive abilities can help students forge ideologies that integrate their personal identity with moral and political stances, encompassing the shared concerns and ideals of others. Erikson detailed the “developmental moratorium” that can interrupt an adolescent’s unhealthy process of development and give them a chance to escape a foreclosed identity, preventing them from moving forward in life without exploring enhanced relationships with others (Nakkula, 2003; Youniss and Yates, 1997).

Community service-learning also promotes cognitive growth as understood by constructivist learning theory, developed by Piaget (Singer & Revenson, 1996) and Vygotsky (Wertsch, 1985). Rogoff (2003) claimed learning is a cumulative

and mutually transformative process in which culture and people shape one another through active participation with their local environment. Community service-learning exemplifies this collaborative, problem-solving learning style that promotes social dialogue and is contextualized by its sociocultural environment.

Chapter Two established the philosophical origins for the promotion of service and citizenship in education. It demonstrated how more practical forms of education, relating to civic reform, agriculture, industry, and home economics, took root in higher learning alongside an esteemed, European, liberal arts curriculum. As the worlds of work and education merged, the African American community remained divided on what type of education best served their needs for greater social justice and civic participation. The Progressive Era and the educational framework of Dewey constructed a valued foundation of experiential education for service-learning, with the tenets of interaction and continuity resonant in the educative experience—the perfect means to both educate and bolster a participatory democracy (Dewey, 1938; Rocheleau, 2004).

The social work profession and federal government played an important hand during the mid-20th century in strengthening civic ideology, and creating models of community service. Presidents Roosevelt, Kennedy, Johnson, Bush, and Clinton were supportive of service as a national goal with measures such as the Civilian Conservation Corps, Peace Corps, War on Poverty, and the National Community and Service Trust Act. Many state governments followed suit with funds to implement service programs in K-12 schools, such as the Learn and Serve Program featured in the previous chapter. The rise in community service-

learning as a new pedagogy with a civic agenda captured the attention of educational researchers beginning in the 1980s, and the literature has expanded as it seeks to piece out what exactly is being learned.

At the turn of the 21st century, the research community as a whole was dissatisfied with the ambiguity of terms and methodology that prevented more consistent and conclusive findings on community service-learning's impact on secondary students. This paper has systematically reviewed the literature of the last decade, analyzing and critiquing the directions it has taken. Chapter Three organized 30 research articles into four focus areas: civic engagement, academic achievement, personal development, and service design. It presented a critical review of those studies, and now Chapter Four revisits the research, guided by the questions: what are the effects of community service-learning on adolescents' civic engagement, academic achievement, and personal development? Are there design elements that better enable these outcomes? The following summary of the findings examines the contemporary research and its limitations as a whole, based on the initial review from Chapter Three. The implications for the secondary classroom will then be discussed, followed by suggestions for further research.

Summary of Findings

The effects of community service-learning have been primarily a construct and measure of civic engagement, as evidenced by more than half of the research critically reviewed in this paper. The majority of the findings suggested that the civic outcomes of service-learning are significantly positive. Only Rutter

and Newmann (1989), Lakin and Mahoney (2006), Lee et al. (2007) Yamauchi et al. (2006) found no effect on some of the civic outcome variables measured, and even still, they managed to report gains in other areas. Yamauchi et al. and Lee et al. found a perceived increase in students' connection to their community and sense of efficacy to take action in it. Yates and Youniss (1998), Kahne and Sporte (2007), Kielsmeier et al. (2004), Schmidt et al. (2007), and Lee et al. (2007) reported increases in measures of students' levels of civic engagement, participation, knowledge and efficacy. Lakin and Mahoney, Melchior and Bailis (2002) and Metz and Youniss (2003) were consistent in their findings of an increase in youth's propensity to serve, with the greatest impact for students initially less-inclined to serve. McGuire and Gamble (2005) and Halfacre et al. (2006) both indicated a strengthened sense of social responsibility, and Terry (2000) found increases in student caring, commitment and empowerment.

The strongest studies controlled for the interaction variables that stem from the background characteristics of students, schools, classrooms, teachers, and neighborhoods. Kahne and Sporte (2007) and Schmidt et al. (2007) employed multiple regression models, Billig et al. (2005) used a MANOVA, and Melchior and Bailis (2002) used an ANCOVA. Another strength noted in the literature on civic engagement was its inclusiveness of low-income student and students of color. While few of the studies were what could be considered nationally representative, there was a balanced focus on Caucasians, Hispanics, and African Americans. There were however, low percentages of Asians and Native Americans in many of the samples.

The occasional omission of demographic data compromised the literature's external validity. Rutter and Newmann (1989), Riedel (2002), Melchior and Bailis (2002), Halfacre et al. (2006) reported no demographic data, and an even greater number failed to report the gender of their participants. Of 16 studies, only 4 indicated that the male to female ratio was roughly equal. Four studies had disproportionate numbers of female participants. Researchers have hypothesized that greater numbers of females complete service, and that gender may be an important predictor in service outcomes. The lack of better data on participant's gender is relevant and problematic. Five of the quantitative studies had sample sizes of 100 or less, and this again limits the generalizability of the research findings.

While the findings seem consistent and positive overall, a considerable amount of the quantitative research on civic engagement lacks the statistical rigor and reliability of even a quasi-experimental design. In the 16 research studies reviewed in the first section, only half control for maturation in their research with the use of a control group. Yamauchi et al. (2006) provided a control group, but did not control for the history of their participants by including a pretest in their research study. Another frequent limitation on the reliability of the quantitative research was the lack of data about instrumentation; 11 of the 14 quantitative studies either reported partial or no alpha scores on their instrumentation, or depended upon instruments with very low internal consistency ($.4 < \alpha < .8$). Instrumentation was always consistent from pretest to

posttest, but researchers typically used different instruments between studies making it harder to compare dependable results from one study to the next.

The focus of this paper is on the effects of service-learning, but seven of the studies on civic engagement studied the broader dependent variable of community service, and therefore offered less precise findings on the outcome of a structured service program with an integrated academic component and reflection. This is partly a result of reliance on national survey data; even the research of Sporte and Kahne (2007), which did focus on service-learning, could not specify details on the service programs completed because of the data source. For researchers like Kielsmeier et al. (2004) and Schmidt et al. (2007), national survey data allowed them to examine more nationally representative samples, but limited the precision of their research. Other factors that create doubt around the internal validity of the findings are high mortality rates—as much as 50%—without explanation in the work of Morgan and Streb (2001), Riedel (2002), Lee et al. (2007), and Halfacre et al. (2006).

In the absence of control groups, statistical significance, and reliability levels, the trust and confidence of naturalistic data depends on strong methodology. Most of the qualitative research reviewed lacked credibility and confirmability, and Rutter and Newmann (1989), Yates and Youniss (1998), Riedel (2002) are examples of this. Even though Rutter and Newmann are frequently cited, their research produced no significant civic outcomes, and they acknowledged none of the serious limitations on their work. They provided no explanation of how they coded and collected their qualitative data, and did not

member-check or acknowledge self as instrument. The only strength the above authors demonstrated in their qualitative methods was triangulation of data. For authors such as Yates and Youniss, who draw heavily upon their qualitative findings to make substantial conclusions, there was no outside evaluation of the data to reduce bias and create more confirmability in their work. Even Billig et al. (2005), whose leading author is a central figure in advancing service-learning research, exhibited similar limitations. The authors created strong triangulation of data, but there was no discussion of member-checking, author bias, or mention of who collected the qualitative information and the process by which it was coded.

Two examples of improvements upon the problematic trends in qualitative service-learning research on civic outcomes are the work of Terry (2000) and Yamauchi et al. (2006). Terry's qualitative methods clearly detailed triangulation, her coding process, the use of external evaluators, and member-checking with participants. Yamauchi et al. did not report member-checking, but their research was clear in its data analysis methods, and outside parties collected and audited the data for greater confirmability.

The six studies of the academic achievement section were positive overall, but not conclusive around students' quantitative gains. Studies by Scales et al. (2000) and Roberts and Moon (1997) found no gains for GPA and limited gains for students' quantitative scores, while Kraft and Wheeler (2003) and Magarrey and Francis (2005) indicated more significant gains in these areas, particularly for previously low-achieving youth. The perception of school

principals, analyzed by Kielsmeier et al. (2004) and Scales et al. (2006), suggested that the majority of principals were divided on how important they felt the pedagogy was in their schools, and only a third of respondents valued it highly. Scales et al. (2000) and Jensen and Burr (2006) reported students' increased classroom commitment, proficiency and confidence in course content. Robert and Moon (1997) and Terry (2000) observed the development of problem-solving and reasoning as examples higher-order thinking skills in students. Finally, Scales et al. (2006) indicated there was potential for service-learning to reduce the achievement gap between low and high-SES students.

The research on achievement has a much tighter focus on service-learning as opposed to other types of service, and all but Scales et al. (2006) are able to distinguish between service-learning and community service in their research. Three of the six studies benefited from large sample sizes, creating more representative findings. However two of the studies were rural in locality, one was suburban, and two studies provided no demographic details. Additionally, two of the six studies samples were overwhelmingly male. There is no indication Latinos, African Americans, Asian Americans, and other ethnic groups were proportionally represented. Therefore the external validity of this body of work is still in question.

All of the studies presented quantitative findings with methodology lacking in internal validity. Kraft and Wheeler (2003), Magarrey and Francis (2005), Roberts and Moon (1997), and Jensen and Burr (2006) reported results in the form of raw scores, without attaching any statistical analysis or significance.

These studies also suffered from the omission of alpha scores on instrumentation, and a notable absence of the regressional analysis or ANOVAs featured in stronger studies to explore the interaction of background variables. In the case of Kraft and Wheeler, although 50% of the sample was low-SES, the authors did not explore the differential impact it had on participant outcomes. The qualitative findings were also inadequate as discussed previously. Jensen and Burr claimed their study was qualitative, but reported no qualitative findings or methods. The teacher was the only collector and evaluator of data, and yet there was no discussion of bias or member-checking. Additionally the study had a duration of only two weeks, which is very short relative to most service learning programs.

The work of Scales et al. (2000) and Scales et al. (2006) stood out as the strongest, most credible work in the section. Both studies featured large samples, and ran ANOVAs or ANCOVAs to account for differences in the sample. Yet even these studies had strong limitations on their internal validity. Scales et al. (2000) experienced a 50% mortality rate, and the sample was non-representative due to the high educational background of parents, and the prior service experience of participants. Scales et al. (2006) used instrumentation with extremely low reliability creating doubt around the findings. As a whole, the academic achievement section did not feature significant and credible findings to warrant serious implications, and there was a notable absence of qualitative research. Only Terry's (2000) strong qualitative methodology previously

discussed under civic engagement compensated for the weakness of the mainly quantitative findings.

The findings on personal development established the potential of service learning to raise students' self-esteem (Johnson and Notah, 1998), reduce risky sexual behavior and delinquency (O' Donnell et al., 1999; Hoffman and Xu, 2002: Schmidt, et al., 2007), and promote greater career development (Terry, 2007; Yamauchi, 2006) and resilience in youth (Kraft and Wheeler, 2003). The quantitative methods of this small subset of the literature are much stronger than those of the previous two sections. They are also more concentrated on service-learning, with large sample sizes and better representation of African American and Hispanic populations, urban localities, and equal ratios of male to female. A clear lack across all sections still remains: Asian Americans, Native Americans, other ethnic groups are insufficiently represented.

O'Donnell et al. (2006) presented one of the finest quantitative studies of all the research reviewed. This on the account of clearly defined criteria for service-learning treatment, a low mortality rate, the accounting for both history and maturation with a pretest and posttest and strong control group, and the inclusion of regressional analysis to account for difference. Hoffman and Xu (2002) also used strong quantitative methods in their correlational research, which benefited from a sample with high external validity, and the use of instrumentation with a high alpha score. Johnson and Notah (1998) offered the weakest contribution to the section in their qualitative research, which had no transparency or confirmability in its methods. For their quantitative findings, no

alpha score was reported on the instrumentation, and the service treatment was highly variable, sometimes as brief as two hours. Kraft and Wheeler (2003) and Schmidt et al. (2007), as discussed previously, had little strength to offer this section in their related findings, but Yamauchi et al.'s (2006) research was a significant addition. Even though their findings on career development lack external validity due to the purposive sampling on native Hawaiians, their research is credible enough to prompt further research.

The final section of the research discussed elements of service design, with variable findings. Furco's (2002) contradictory qualitative and quantitative findings could not demonstrate whether service learning was equal to or more effective than community service or non-service curriculum. Schmidt et al. (2007) found that students in direct service earned higher grades, while Billig et al. (2005) found that students in indirect service had higher academic engagement. Perhaps if they had both used a consistent academic measure there might have been more reliability between the two. There was greater consistency between the two researchers in their conclusions that assisting organizations, or direct civic or political service related to greater civic efficacy and civic knowledge. Terry (2000) and Morgan and Streb's (2001) research on service design was compatible in its conclusion that greater voice and ownership by students in program planning and implementation was beneficial. McLellan and Youniss (2003) concluded that structured and required service programs were more effective at serving those in need. England and Marcinkowski (2007), Billig et al., and Seitsinger's (2005) findings suggested there was a significant interaction

between teacher and service program outcomes, in terms of experience, attitudes, and which program outcomes are emphasized. Yates and Youniss (1998) concluded that class discussion and reflection were an essential part of any effective service program, and Riedel (2002) indicated that programs with a participatory nature were more effective at increasing students' feelings of civic obligation than private service.

All the research of this section was either mixed-methods or quantitative, and all but one study concentrated directly on service-learning. Each study had a large sample size, and indicated ethnically and socioeconomically diverse samples. McLellan and Youniss's (2003) findings were strengthened by their use of regression analysis consistent with other strong quantitative studies, but the sample came from a Catholic school service-learning program connected to a religious course, and the conclusions were therefore non-representative. Only Furco (2002) used a quasi-experimental design, but his statistical methods were weak, relying on instrumentation with low alpha scores. His qualitative work was much stronger, with transparent coding and triangulation, but consistent with most of the qualitative data reviewed thus far, it lacked member-checking. England and Marcinkowski's (2007) research was exceptionally weak, with high variation in the program designs evaluated, and a complete reliance on teacher-reported data. The results were presented in raw form without statistical analysis or alpha scores for instrumentation. The authors claimed qualitative methods, but provided no information on how they conducted their analysis, and made conclusions that were based on unreported data.

Seitsinger's (2005) research was the strongest addition to this section, with a nationally representative sample, strong mean differences with high significance, and high alpha scores on her instrumentation. Although Morgan and Streb's research was weak, Terry's (2000) strong qualitative methods lend more credibility to the consistent findings on the importance of voice and leadership from students in designing service programs.

Classroom Implications

While the cumulative findings of the research reviewed do not always significantly point to positive outcomes in the civic, academic, and personal domains, the directionality of community service-learning's effects are consistently in the right direction. As the thorough critique of the research methods above suggested, failure to determine the benefits of community service-learning may still be the unfortunate consequence of inadequate research methodology, rather than evidence of an ineffective pedagogy. The research has shown community service-learning to be as effective if not more so than traditional methods, and as the next section will detail, the strengthening of further research holds promise to highlight more conclusive findings.

From the research on academic achievement, one can infer that gains in grade point averages, test scores, and attendance levels may not rise as a result of service (Scales et al. 2000; Roberts and Moon, 1997). Teachers may not even have the full faith of their principals that the community service-learning work they facilitate is worth its time and effort (Kielsmeier et al., 2004). For an educational system that is heavily dependent on quantitative gains on state

assessments, the hard data to support community service-learning is not quite there. Some educators would then question its worth in a public school system already pressed for resources, time, and students' attention. But for researchers such as Terry (2000) and Roberts and Moon, whose assumptions about what constitutes academic achievement went beyond the yardstick of grades, there were findings that suggested the kind of cognitive growth many educators struggle to create in classrooms with traditional methods. Higher order thinking skills, problem-solving and reasoning skills were evident in their work—but their findings are conditional. Roberts and Moon warned that the student may easily lose sight of their academic agenda in the course of their project, and Terry's students were in a gifted program with the skills of self-discipline already in place. This means the teacher in the typical classroom of mixed abilities must scaffold skill building in time management, self-discipline, group work, planning, and whatever social skills her students need to engage in a more constructive, social form of learning.

Morgan and Streb (2001) and Terry (2000) suggested that student voice, planning, and greater ownership enhances community service-learning outcomes. Student-selected projects were shown to be more effective than those the teacher selected (Billig et al., 2005). This implies that students and teachers must engage in a conversation about what the areas of need are in their community, providing a teaching opportunity to model the group decision making process. This student-directed learning can enhance the workings of a more democratic classroom, and students' commitment to the classroom, necessary to

sustain it, comes more intrinsically as an effect of the methodology (Jensen & Burr, 2006; Scales et al., 2000). Community service-learning also complements the instructional methods and best practices advocated by leaders in progressive education. The research on classroom instruction suggested that problem-based instruction, cooperative learning, and classroom discussion are all effective teaching strategies, and that providing student choice on project-based units are reflective of “best practice” (Arends, 1997; Daniels & Bizar, 2004).

Community service-learning can be a way to reach the lowest-achieving students, as Magarrey and Francis (2005) indicated those with the lowest grades demonstrated greater gains than service students that were already high achievers. The same holds true for community involvement; those who demonstrated the least inclination to serve in their community initially proved to have some of the greatest increases in volunteering (Metz and Youniss, 2003). If a teacher is not able to structure community service-learning for the whole class, perhaps he can provide opportunities for those struggling with motivation, attendance, and finding enjoyment in school—in other words, those who might benefit from it the most. As Scales et al. (2006) suggested, service-learning can make the difference in reducing the achievement gap between the high and low achievers in the classroom.

Yates and Youniss’s (1998) work indicated that the union of classroom and community through discussion and reflection is essential, and the teacher holds the responsibility for facilitating this work. Billig et al. (2005) concluded that the experience of the teacher in guiding students through service-learning is

significant to its success, but that should not deter beginning teachers. Rather, it is an incentive to seek out and collaborate with colleagues already actively involved in community service-learning instruction.

Suggestions for Further Research

Quantitative methods dominated the research reviewed in this paper, yet many of the articles featured consistently weak design and reports of their work. To create higher external validity in the findings, a full report of the demographics should be included, detailing the ratio of male to female participants, locality, and ethnicity. Future research must acknowledge the presumption that females are more likely to serve and benefit from service, and make a concerted effort to balance the gender of their participants and report it at such. Asian Americans and Native Americans are underrepresented in the samples compiled by current researchers, and further research needs to be done to explore the effects of community service-learning on larger school populations of both groups. For Native American participants, it would be valuable to have data on community service-learning to compare from both reservation schools and traditional, public schools.

While it is difficult to establish truly random sampling in working with K-12 participants, a greater attempt at randomized, larger, nationally representative samples is needed to create more generalizable findings. Since a true experimental design is unattainable, further research requires a greater number of quasi-experimental studies. Almost half of the studies reviewed in all sections did not provide for a control group, and this leaves the effects of maturation

uncontrolled for. The use of a pretest and posttest in a nonequivalent control group design is a strong method that educational researchers can use to increase the internal validity of their work, controlling for history and maturation.

Additionally, further research must apply statistical analysis more rigorously. Authors such as Magarrey and Francis (2005), Jensen and Burr (2006), Riedel (2002) and Morgan and Streb (2001) could have established greater confidence in their findings if they had reported the reliability levels of their instrumentation, and only depended on measures with alpha scores greater than .80. Furthermore, they needed to run their data through statistical analysis, and not provide raw data without reporting its statistical significance. Leading authors in the field, such as Scales et al. (2006) and Furco (2002) have demonstrated the necessity of ANOVAs, ANCOVAs, and MANOVAs in accounting for different variables that may affect outcomes. Kahne and Sporte (2007) and Hoffman and Xu (2002) also featured four-stage regressional modeling, which controls for different levels of interaction variables starting with individual characteristics such as gender, ethnicity, and age, and then moves through characteristics of the classroom, school, home and family, and neighborhood allowing a staged analysis. Further research would benefit from consistent use of these methods, or an explanation of why they are not used, since background variables and differences will affect all community service-learning research samples.

The qualitative research on community service-learning must also improve its methodology. Most authors consistently collected data from multiple sources,

in interview formats, observations, surveys, and test scores, developing substantial triangulation. Future research must strengthen the other areas that were significantly lacking. All qualitative research must first acknowledge the self as instrument, and be explicit about the assumptions and biases the authors bring to the work. Johnson and Notah (1998) could have increased the credibility of their findings by explaining the methods they used to code and analyze interview data and opening their collection and analysis of the data to an outside evaluator for audit. Jensen and Burr (2006) and Yates and Youniss (1998) needed to concede the personal bias in their findings as a result of their lone, unaudited review and analysis. Authors can engage the participant in raising the confirmability of their findings by consistently member-checking, ensuring that their interpretation of observations and interviews is correct and making participants less passive in the research process. Of all the qualitative research reviewed, Terry (2000) was the only author to member-check, and her qualitative methods can serve as an example for future research.

Authors such as Halfacre et al. (2006) and Scales et al. (2000) need to better account for the high mortality rates in their studies. Also, less reliance on national survey data would create more precision in the research findings and allow for more longitudinal data than annual surveys that change from year to year allow. Melchior and Bailis (2002) called for additional research that focuses on the input of service-learning research as well as the outcomes; program design and duration and intensity of service were identified as target areas that the authors believed need greater attention. Indeed not enough attention was

paid to the conflated variables present in even the same classroom's service options, much less the differences that exist between service programs of different schools and areas. Authors need to provide details on whether service is required or mandatory, self-selected, what type of service it is, and what the frequency, intensity, and duration of students' service is. Once these variables are made clear upfront, further research can work on producing more dependable results by comparing the efficacy of similarly structured programs, and better account for the variance of the findings.

Few studies reviewed here have even broached the subject of program length, and Jensen and Burr (2006) and Johnson and Notah (1999) examined service programs as brief as two weeks, and two hours. Program duration is one variable that has the most obvious potential for further research, as indicated by Billig et al. (2005). If a study could isolate the effects of duration between similar service programs, and examine the student outcomes at different intervals over a one-year period, it could greatly benefit the length of service design that remains highly variable.

Perry and Katula (2001) suggested that research about citizenship skills and behaviors is lacking in the literature, Attitudinal outcomes—such as self-esteem and civic responsibility—are most easily measured by scales and surveys, and as a result, the psychological effects of service have dominated the literature. Research that actually documents what youth are doing, such as how many high school seniors register to vote, how many join civic organizations, or how many volunteer for school government or public campaigns are all examples

of the types of political behaviors—rather than attitudes—that could be measured.

A trend in the service literature is the younger the participants, the scarcer the research. Lakin and Mahoney (2006) and Halfacre et al. (2006) were the only authors to work with elementary aged youth, with the youngest participants at a fifth grade level. There is a great opportunity for researchers to pioneer studies on the effects of community service-learning in the K-5 setting. Kahne and Sporte (2007) and McLellan and Youniss (2003) proposed there were greater gains for service participants with previous service experience. Therefore, the immediate and longitudinal effects of service could be better understood if youth participants were contacted at younger ages.

Both Hoffman and Xu (2002) and O' Donnell et al. (1999) had two of the strongest research designs with significant findings on service's negative impact on undesirable behaviors. Community service-learning's use as deterrent seems to be another area of research that is ripe for investigation. It adds the complexity of working with at-risk youth, who may be difficult to track as participants for long term studies due to unstable home environments, incarceration, drug abuse, and inconsistent school attendance. As Hoffman and Xu pointed out in their data from the National Educational Longitudinal Study, students that dropped out of school provided no data. Despite the challenges such research poses, the potential to help adolescents with self-destructive, addictive, or criminal behaviors is worthwhile, both to reform members of civic society, and to improve participants' personal health and well-being.

Finally, Terry (2000) and Yamauchi's (2006) credible qualitative findings concerning career development point to a highly relevant field of new community service-learning research. While vocational education programs allow youth to train in certain areas, they are limited in their scope to mainly industrial, cosmetic, culinary, or secretarial positions. If high school courses could implement community service-learning programs with a wider array of working options, perhaps the professional imaginings of adolescents could proliferate. Especially for students not college bound, this could be one more way for community service-learning to narrow the achievement gap discussed by Scales et al. (2000).

Conclusion

As defined in Chapter One, community service-learning is a combination of three things: a course curriculum with clear learning objectives, meaningful service to the local community, and regular reflection and critical analysis. These components are each valuable in their own right, but their sum is a greater teaching tool. Together these elements provide an educative experience of interaction and continuity that takes place in the social world of the adolescent, bridging the gap between books and the wider world that youth often find themselves in without direction. Founded on the principles of experiential education outlined by Dewey (1938) and Kolb (Pritchard & Whitehead, 2004), and informed by an understanding of adolescent identity as discussed by Erikson (Nakkula, 2003), community service-learning has a strong theoretical core that is not easily dismissed.

Chapter Two illustrated that human thought reaches far back into antiquity in its ruminations on the role of education as the forge of civic society. Over the past century, civic leaders have made service a national priority, and schools have become the front line of community action. Although community service has long been heralded as the bringer of civic renewal and personal virtue, its newest incarnation as a social constructivist pedagogy is still under relatively recent examination, especially surrounding its cognitive effects.

The critical review of the literature on community service-learning featured in the third chapter functioned to demonstrate the potential of service-learning as documented by the latest empirical studies. The literature was organized by four emergent themes: civic engagement, academic achievement, personal development, and service design. These themes served to focus the inquiry of this paper on the outcomes and design of service. The research on civic engagement found that community service-learning could enhance civic participation, commitment, knowledge and efficacy, and give students a greater sense of social responsibility, with an end result of more community service involvement. The findings on academic achievement indicated that while the effects of service on participants' GPA, attendance, and test scores is variable, there are other more qualitative benefits apparent: an increase in commitment and confidence in the classroom, the development of higher order thinking skills, and proficiency in content knowledge. Conclusions on personal development indicated that self-esteem and personal responsibility rise through service, students experience greater opportunities for career development and caring

relationships with adults, and there is a decrease in risky sexual behavior and delinquency. The research on service design indicated that student voice, ownership, and selection of activities can enhance the outcome of a program, and experienced teachers who provide structured participatory service opportunities of at least one semester, in both direct and indirect service, will enable students to a full range of skills and outcomes. This chapter summarized the findings of Chapter Three by section, discussed the secondary classroom implications, and made suggestions to strengthen future research.

As U.S. citizens work to restore the vibrancy of democracy in a time where young people's faith in government has been weakened, educators have a vital political role to play. They can structure civic learning opportunities in a variety of disciplines that engage students in participatory learning and meaningful community interaction. By addressing real needs in the community, adolescents are given an opportunity to demonstrate their worth to society, and become invested and hopeful as citizens. Through participatory service that extends care and concern for others, youth can develop higher levels of thinking, self-esteem, and resilience for a modern world laden with uncertainty. With the experiential education and reflection encouraged through community service-learning, each student realizes they are a valued and needed member of their local community.

REFERENCES

Arends, R. (1997). *Classroom instruction and management*. San Francisco, CA: McGraw Hill.

Billig, S., Root, S., & Jesse, D. (2005). *The impact of participation on service-learning on high school students' civic engagement.* College Park, Maryland: Center for Information and Research on Civic Learning and Engagement. (ERIC Document Reproduction Service No. ED 495215).

Billig, S. (2002). Adoption, implementation, and sustainability of K-12 service-learning. In A. Furco & S. Billig (Eds.), *Service-learning: The essence of pedagogy* (p. 3-22). Greenwich, CT: Information Age Publishing.

Billig, S. & Furco, A. (2002). Research agenda for K-12 service-learning: A proposal to the field. In A. Furco & S. Billig (Eds.), *Service-learning: The essence of pedagogy* (p. 3-22). Greenwich, CT: Information Age Publishing.

Campbell, D. (2000). Social capital and service learning. *Political science & politics, 33, 3,* 641-645.

Conrad, D. & Hedin, D. (1991). School-based community service: What we know from research and theory. *Phi delta kappan, 72,* 743-749.

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

Daniels, H. & Bizar, M. (2004). *Teaching the best practice way: Methods that matter, K-12*. Portland, ME: Stenhouse.

Dewey, J. (1938). *Experience & education.* New York: Simon & Schuster Inc.

Du Bois, W. (1903). The talented tenth. In H. Gates, Jr. & C. West (authors), *The future of the race* (pp. 133-158). New York, NY: Vintage Books.

England, Y. & Marcinkowski, T. (2007). Environmental service-learning programs in Florida high schools and colleges: Nature, status, and effects as determined by a statewide program census. *The journal of environmental education, 38, 4,* 51-60.

Flanagan, C. & Sherrod, L. (1998). Youth political development: An introduction. *Journal of social issues, 54, 3,* 447-456.

Furco, A. (2002). Is service-learning really better than community service? A study of high school service program outcomes. *Service-learning: The essence of the pedagogy.* Furco, A. & Billig, S. (eds.). Greenwich, CT: Information Age Publishing.

Halfacre, J., Chessin, D. & Chambless, M. (2006). Service-learning and student attitudes. *Academic exchange quarterly, Spring,* 226-231.

Hoffman, J. & Xu, P. (2002). School activities, community service, and delinquency. *Crime and delinquency, 48, 4,* 568-591.

Jensen, S. & Burr, K. (2006). Participation and learning relationships: A service-learning case study. *Journal of industrial teacher education, 43, 3,* 6-28.

Johnson, A. & Notah, D. (1999). Service-learning: History, literature review, and a pilot study of eighth graders. *The elementary school journal. 99, 5,* 453-467.

Kahne, J. & Sporte, S. (2007). *Developing citizens: The impact of civic learning opportunities on students' commitment to civic participation*. Chicago, IL:

Consortium on Chicago school research. (ERIC Document Reproduction Service No. ED 499374).

Kielsmeier, J., Scales, P., Roehikepartain, E., & Neal, M. (2004). Community service and service-learning in public schools. *Reclaiming children and youth, 13, 3,* 138-143.

Kraft, N. and Wheeler, J. (2003). Service learning and resilience in disaffected youth: A research study. *Deconstructing service-learning: Research exploring context, participation, and impacts.* Billig, S. and Eyler, J. (Eds.). Greenwich, CT: Information Age Publishing. p. 213-240.

Lakin, R. & Mahoney, A. (2006). Empowering youth to change their world: Identifying key components of a community service program to promote positive development. *Journal of school psychology, 44,* 513-531.

Magarrey, M. & Francis, R. (2005). Service learning and student academic success. *Academic exchange,* 245-248.

Lee, S., Olszewski-Kubilius, P., Donahue, R., & Weimholt, K. (2007). The effects of a service learning program on the development of civic attitudes and behaviors among academically talented adolescents. *Journal for the education of the gifted, 31, 2,* 165-197.

McGuire, J. & Gamble, W. (2005). Community service for youth: The value of psychological engagement over numbers of hours spent. *Journal of adolescence, 29,* 289-298.

McLellan, J. & Youniss, J. (2003). Two systems of youth service: Determinants of voluntary and required youth community service. *Journal of youth and adolescence, 32, 1,* 47-58.

Melchior, A., Bailis, L. (2002). Impact of service-learning on civic attitudes and behaviors of middle and high school youth: Findings from three national evaluations. *Service-learning: The essence of the pedagogy.* Furco, A. and Billig, S. (Eds.). Greenwich, CT: Information Age Publishing. p. 201-222.

Metz, E. & Youniss, J. (2003). A demonstration that school-based required service does not deter but heightens volunteerism. *PS: Political science and politics, 36, 2,* 281-286.

Morgan, W. & Streb, M. (2001). Building citizenship: How student voice in service-learning develops civic values. *Social science quarterly, 82, 1,* 154-169.

Nakkula, M. (2003). Identity and possibility: Adolescent development and the potential of schools. In M. Sadowski (Ed.), *Adolescents at school: perspectives on youth, identity, and education* (pp. 7-18). Cambridge, MA: Harvard Education Press.

O' Donnell, L., Stueve, A., Doval, A., Duran, R., Haber, D., Atnafou, R., Johnson, N., Grant, U., Murray, H., Juhn, G., Tang, J. & Piessens, P. (1999). The effectiveness of the reach for health community youth service learning program in reducing early and unprotected sex among urban middle school students. *American journal of public health, 89, 2,* 176-181.

Owen, D. (2000). Service-learning and political socialization. *Political science and politics, 33, 3,* 639-640.

Perry, J. & Katula, M. (2001). Does service affect citizenship? *Administration and Society, 33, 3,* 330-365.

Pritchard, F. & Whitehead, G. (2004). *Serve and learn: Implementing and evaluating service-learning in middle and high schools.* Mahwah, NJ: Lawrence Erblaum Associates, Inc.

Pritchard, I. (2002). Community service and service-learning in America: The state of the art. In A. Furco & S. Billig (Eds.), *Service-learning: The essence of pedagogy* (p. 3-22). Greenwich, CT: Information Age Publishing.

Riedel, E. (2002). The impact of high school community service programs on student's feelings of civic obligation. *American political research, 30, 5,* 499-527.

Roberts, L. & Moon, R. (1997). Community service learning methodology and academic growth in secondary school content disciplines: An action-research study. *The high school journal, 80, 3,* 202-209.

Rocheleau, J. (2004). Theoretical roots of service-learning: Progressive education and the development of citizenship. In B. Speck & S. Hoppe (Eds.), *Service-learning: History, theory, and issues* (p. 3-21). Westport, CT: Praeger.

Rogoff, B. (2003). *The cultural nature of human development.* Oxford: Oxford University Press.

Rousseau, J. (1968). *The social contract.* (M. Cranston, Trans.) London: Penguin Books. (Original date published 1762).

Rutter, R. & Newmann, F. (1985). *Educational Leadership*. *43, 4,* 65-71.

Terry, A. (2000). An early glimpse: Service learning from an adolescent perspective. *Journal of secondary gifted education, 11, 3,* 115-136.

Scales, P., Roehikepartain, E., Neal, M., Kielsmeier, J., & Benson, P. (2006). Reducing academic achievement gaps: The role of community service and service-learning. *Journal of experiential education, 29, 1,* 38-60.

Scales, P.C., Blyth, D.A., Berkas, T.H., & Kielsmeier, J.C. (2000). The effects of service-learning on middle school students' social responsibility and academic success. *Journal of early adolescence, 20, 3,* 332-358.

Schmidt, J., Shumow, L., & Kackar, H. (2007). Adolescents' participation in service activities and its impact on academic, behavioral, and civic outcomes. *Journal of youth adolescence, 36, 2,* 127-140.

Shumer, R. (2005). The wisdom of Delhi: An investigation of the most influential studies in K-12 service-learning research in the past 25 years. In K. Casey, G. Davidson, S. Billig, & N. Springer (Eds.), *Advancing knowledge in service-learning: Research to transform the field* (pp. 185-204). Greenwich, CT: Information Age Publishing.

Singer, D. & Revenson, T. (1996). *A Piaget primer: How a child thinks.* New York, NY: Plume.

Seitsinger, A. (2005). Service-learning and standards-based instruction in middle schools. *Journal of educational research, 99, 1,* 19-29.

Speck, B. & Hoppe, S. (2004). *Service-learning: History, theory, and issues.* Westport, CT: Praeger.

Spring, J. (2005). *The American school: From the puritans to no child left behind.* San Francisco, CA: Mc Graw Hill.

Wade, R. (1997). *Community Service-Learning: A guide to including service in the public school curriculum.* Albany, NY: State University of New York Press.

Wertsch, J. (1985). *Vygotsky and the social formation of the mind.* Cambridge, MA: Harvard University Press.

Van Benschoten, E. (2000). Youth-led civic organizing: Countering perceptions of apathy and redefining civic-engagement. *National civic review, 89, 4,* 301-307.

Yamauchi, L., Billig, S., Meyer, S., & Hofschire, L. (2006). Student outcomes associated with service-learning in a culturally relevant high school program. *Journal of prevention & intervention in the community*, *32, 1/2*, 149-164.

Yates, M. & Youniss, J. (1998). Community service and political identity development in adolescence. *Journal of Social Issues, 54, 3,* 495-512.

Youniss, J., McLellan, J. & Yates, M. (1997). What we know about engendering civic identity. *The American behavioral scientist, 40, 5,* 620-631.

Youniss, J. & Yates, M. (1997). *Community service and social responsibility in youth.* Chicago, IL: The University of Chicago Press.

Zieren, G. & Stoddard, P. (2004). Theoretical roots of service-learning: Progressive education and the development of citizenship. In B. Speck & S. Hoppe (Eds.), *Service-learning: History, theory, and issues* (pp. 23-42). Westport, CT: Praeger.