ABSTRACT

Title of Document: PARTICIPATORY EQUITY AND STUDENT OUTCOMES IN LIVING-LEARNING PROGRAMS OF DIFFERING THEMATIC TYPES

Matthew Edward Soldner
Doctor of Philosophy
2011

Directed By: Dr. Karen Kurotsuchi Inkelas
Associate Professor
College of Education
University of Maryland

This study evaluated participatory equity in varying thematic types of living-learning programs and, for a subset of student group × program type combinations found to be below equity, used latent mean modeling to determine whether statistically significant mean differences existed between the outcome scores of living-learning participants and their peers in traditional residence hall environments. This study employs a conceptual framework informed by Astin's (1991) IEO model and Pascarella and Terenzini's (1980) model of structural mediation in residential environments, and is based on data collected as part of the 2007 National Study of Living-Learning Programs.

First, a team of raters used descriptive content analytic techniques to identify a typology of living-learning programs consisting of 41 specific thematic types, based on those programs' stated goals and objectives. That typology was the basis for computing Hao's (2002) equity indices, which were used to determine whether students from different racial/ethnic groups or socioeconomic statuses were under- or over-represented in specific thematic types of living-learning programs, relative to their representation in living-learning programs overall. Twenty-two race/ethnicity × type combinations exhibited low levels of participatory equity, as did 13 socioeconomic status × type combinations.

Three group × type combinations were selected for latent mean modeling, including: (a) Asian/Pacific Islander students in disciplinary, general academic, honors programs; (b) White students in international/global programs; and (c) low-SES students in honors programs. The outcome of interest for Asian/Pacific Islander and low-SES students was ease of academic transition, and, for White students, diversity appreciation. Analyses revealed that although L/L participants reported higher mean scores on measures of several key living and learning environments, no statistically significant mean difference in outcome measures was observed.

In the face of the participatory inequities found in this national sample of living-learning programs, the primary implication for student affairs practitioners generally is that the exploration of equity in high-impact practices for students—and involvement and engagement opportunities for *all* members of the university community—is warranted. This implication is indicated for living-learning practitioners as well, who can also take findings vis-à-vis the relationship between key living and learning environments and specific student outcomes in to account when designing and improving the programs with which they work. Finally, living-learning practitioners should consider whether the relatively small differences in environment measures and the lack of detectable differences in outcome measures is driven by weak treatments, weak measures, or both.

PARTICIPATORY EQUITY AND STUDENT OUTCOMES IN LIVING-LEARNING PROGRAMS OF DIFFERING THEMATIC TYPES

by

Matthew Edward Soldner

A dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2011

Advisory Committee:

Associate Professor Karen Kurotsuchi Inkelas, Chair
Professor Alberto Cabrera
Associate Professor Emerita Marylu K. McEwen
Associate Professor Sharon Fries-Britt
Assistant Professor Julie Park

Table of Contents

List of Tables

List of Figures

List of Equations

Chapter 1 : Introduction

> While student engagement [in so-called "high impact" learning opportunities] is not a silver bullet, finding ways to get students to take part in the right kinds of activities helps to level the playing field, especially for those from low-income family backgrounds and others who have been historically underserved, increasing the odds that they will complete their program of study and benefit in the desired ways. The real question is whether we have the *will* to more consistently use what we know works in order to increase the odds that more students complete their program of study and benefit in the desired ways. (Kuh, Kinzie, Cruce, Shoup, & Gonyea, 2007, p. 70)

Above, Kuh et al. (2007) acknowledge one of higher education's most persistent concerns: once a student has been admitted to a college or university, there is no guarantee that he or she will experience success. Indeed, whether that success is narrowly tailored to mean graduation, or more broadly conceptualized of as the attainment of important student learning outcomes, analysis of the available evidence suggests that, for some groups of students, the battle to succeed in college is hard fought.

Graduation rate data are sobering: bachelor's degree-seekers from non-dominant racial/ethnic backgrounds and from lower socioeconomic strata consistently fare more poorly than their White or wealthier peers (Cook & Cordova, 2006; Terenzini, Cabrera, & Bernal, 2001), restricting later access to meaningful economic and social benefits (Kinzie, Palmer, Hayek, Hossler, Jacob, & Cummings, 2004; Leslie & Brinkman, 1988).

Recent investigations seem to echo that trend, at least vis-à-vis race. Focusing on those students who entered four-year colleges in 2001 and who indicated the intention to earn a bachelor's degree, Knapp, Kelly-Reid, and Ginder (2009) reported that 67% of students identified as Asian American and 60% of students identified as White, non-Hispanic had earned their degree by Spring, 2008, compared to 42% of students identified as Black, non-Hispanic, 48% of students identified as Hispanic, and 40% of students identified as American Indian/Alaskan Native.

Graduation notwithstanding, several critics have questioned whether college students are actually *learning*. Work by Arum and Roska (2011), for example, calls in to question whether undergraduates are developing the analytic and cognitive skills required to succeed in today's economy, noting meager gains in critical thinking between a student's first and second year. Evidence of growth in other skill domains, such as building the capacity to work across racial differences, developing an appreciation for liberal learning, clarifying personal values, and engaging with civic and political concerns is similarly scant (Bok, 2006).

Whether critique finds its genesis in the halls of government (e.g., The Secretary of Education's Commission on the Future of Higher Education [The Spellings Commission], 2006) or within the advocacy community (e.g., Association of American College and Universities [AAC&U], 2002; Kellogg Commission on the Future of State and Land-Grant Universities, 2001; National Association of Student Personnel Administrators [NASPA] & American College Personnel Association [ACPA], 2004; National Leadership Council, 2007), it shares two common themes: (a) the cognitive and affective competencies needed by college graduates are more rigorous (and numerous) today than at any point in the nation's history, and (b) little systematic evidence has been gathered to determine whether, how, and how equitably undergraduate education promotes those competencies. In the face of these challenges, the charge to institutions vis-à-vis student success seems clear: irrespective of background, once a student enters college, educators should expect that he or she would graduate having acquired the skills needed for personal and professional success.

To that end, some educators have suggested colleges and universities develop any number of "high impact" (p.7) activities and programs believed to be uniquely well suited to promote learning (National Survey of Student Engagement [NSSE], 2007). According to AAC&U, 10 practices in particular have been "widely tested and shown benefits for college students, especially those from historically underserved backgrounds" (National Leadership Council, 2007, p. 53), including (a) first-year seminars, (b) common intellectual experiences, (c) learning communities, (d) writing-intensive courses, (e) collaborative assignments and projects, (f) undergraduate research, (g) diversity/global learning, (h) service-learning/community-based learning, (i) internships, and (j) capstone courses and projects. Engagement in programs like these, their proponents argue, develop the intellectual, affective, and behavioral capacities that *all* students require to be successful during—and beyond—college (AAC&U, 2002; NASPA & ACPA, 2004; NSSE, 2007; The Spellings Commission, 2006).

However, Kuh et al. (2007) have argued that practitioners do not uniformly apply practices that research (or, in some cases, conventional wisdom) has suggested can improve the outcomes of undergraduate education. To be sure, much of the theory and research underlying the promising practices identified by AAC&U (2002) and others is far from new, being documented in classical texts such as Pascarella and Terenzini's (1991) *How College Affects Students: Findings and Insights from Twenty Years of Research*, Astin's (1993) *What Matters in College? Four Critical Years Revisited,* and Chickering and Gamson's (1991) *Applying the Seven Principles for Good Practice in Undergraduate Education* (National Leadership Council, 2007). Given that the demands facing today's college students—including the unique challenges put before those from non-dominant racial/ethnic or socio-economic groups—seem unlikely to abate without institutional intervention and that practitioners know (and have known) at least *something* of the programs and activities that may be helpful in promoting student success, Kuh et al.'s opening quotation begs two simple questions: are students equally accessing so-called best practices in undergraduate education and, if so, is there evidence to suggest they are actually benefiting?

This dissertation seeks to address a portion of those questions by focusing on the experience of students exposed to a single type of high impact program. Specifically, this research is concerned with residential implementations of learning communities, also known as living learning (L/L) programs. L/L programs situate topically-focused formal and informal student, faculty, and peer interactions within students' residential environment, expanding the site of learning beyond the classroom to include the residence hall, the dining hall, and anywhere else participants are actively engaged with program content (Lenning & Ebbers, 1999; Shapiro & Levine, 1999). In so doing, the student experiences what Schroeder (1994) has lauded as a "seamless" learning environment, where students and faculty can be immersed in the process of learning and are free of the traditional boundaries of the academic building and the credit hour. In the light of such praise, it is perhaps not surprising that although this research considers only one type of high impact program, the one it focuses on is undeniably popular: as I will describe in more detail below, the study from which this research is derived enrolled nearly 50 colleges and universities which, in 2007, offered almost 650 L/L programs to their students (Inkelas, Brower, & Associates, 2008).

L/L programs appear to be an example of a high impact program that has enjoyed substantial diffusion and, in the words of Jones, Levine Laufgraben and Morris (2006), even become “fashionable” (p. 263). The expansion of these programs may be the result of institutional responses to the critiques of undergraduate education identified earlier: learning communities, like those created in L/L programs, have been offered as possible solutions to the vexing problems identified by stakeholders and policymakers (AAC&U, 2002; Kuh, 2007; National Leadership Council, 2007; The Spellings Commission, 2006). Indeed, proponents of learning communities and L/L programs have argued for—and presented preliminary evidence supporting—these programs’ capacity to strengthen the types of outcomes critics have been demanding, including increased cognitive complexity, curiosity and problem solving ability, and citizenship (Inkelas, Johnson, et al., 2006; Laufgraben, Shapiro, & Associates, 2004; Lenning & Ebbers, 1999; Rowan-Kenyon, Soldner, & Inkelas, 2008; Shapiro & Levine, 1999). Similarly, the prevalence of L/L programs may be a result of institutional action to better serve racial/ethnic minority and lower socio-economic status (SES) students who enter college—and persist there—at lower rates than their more privileged peers (Astin & Oseguera, 2004; Cook & Cordova, 2006; Terenzini, Cabrera, & Bernal, 2001). Again, proponents have identified learning communities as a possible solution (Lenning & Ebbers, 1999), and researchers have documented positive relationships between learning community or L/L participation and retention (Stassen, 2003; Tinto, 2000). From these perspectives, the answer to the question “who is benefitting?” inspired by Kuh et al. (2007) is, at least preliminarily, “everyone.”

However, the implementation of L/L programs may do more than simply quell voices for improved educational quality. As a result, an alternative (or, minimally, a simultaneous) rationale for their diffusion exists: by serving as a mechanism for the recruitment of new, high-talent students, L/L programs may directly benefit institutional profiles. Ostensibly objective measures of institutional quality, most notably entrance examination scores of incoming first-year students, are an increasingly important determinant of college rankings and, as such, prestige (Ehrenberg, 2003; Monks & Ehrenberg, 1999; Pascarella et al., 2006). Touted by the popular press and college guidebooks alike as innovative routes to a more rewarding college experience, L/L programs have become attractive selling-points for the institutions that implement them (Bonisteel, 2006; Foderaro, 2005; Thomson Peterson's, 2006). By developing L/L programs that target a desirable sector of potential admits—like students who have distinguished themselves through their high school achievement—at least one institution has found a way to meet its reputational goals (Shapiro & Levine, 1999). In this case, the answer to the question “who is benefitting?” is substantially murkier. Some have suggested that, at least for L/L programs targeting enrichment of high-talent students, institutional interests may surmount those of students, especially those from racial/ethnic minority groups or those from low-SES backgrounds (Soldner, McCarron, & Inkelas, 2007).

The challenges of improving educational quality, student outcomes, and institutional stature are daunting. To the extent that they might be used to meet those challenges, learning communities and L/L programs have been billed as promising practices that may allow for (a) the improvement of educational quality and the support of certain groups of students, or (b) the positive shaping of institutions' incoming classes and the attainment of recruitment goals (Inkelas, Vogt, Longerbeam, Owen, & Johnson, 2006; Jones et al., 2006; Kuh et al., 2007; Lenning & Ebbers, 1999 Shapiro & Levine, 1999; The Spellings Commission, 2006). As a result, the once simple "who benefits?" question may be complicated by the imperative that guides an institution's decision to implement L/L programs. To be sure, there is no necessary indication that institutions have developed L/L programs with participatory equity in mind. However, given the challenges today's graduates face, determining whether *all students benefit equally* through the application of institutional resources seems more important today than it has ever been. With the results of that determination in hand, institutions can—irrespective of their initial motivations—make their own data-based judgments about the level of equity L/L programs are achieving and whether it is consistent with institutional goals.

To ground this research effort, I offer this introduction, which consists of five parts. Each is designed to provide the reader additional context for this work. First, I define what the term "living-learning program" means, both in theory and as practically implemented. Second, I present information about the extent of these programs' adoption. Third, I summarize oft-cited literature about the benefits of L/L participation that suggests a rationale for the creation of L/L programs based upon the positive outcomes that students can accrue. Fourth, I offer alternative rationales for implementing L/L programs that are not focused on benefiting students, but rather the institutions that enroll them. Finally, I state the study's research questions, and suggest ways in which L/L scholarship and practice can benefit through programs' examination.

Learning Communities and L/L Programs

As L/L programs have become more prominent nationally, they have evolved into a set of interventions that, while diverse in their approach, share a common heritage: the learning community movement (Inkelas, Soldner, Longerbeam, & Brown Leonard, 2008; Schoem, 2004). Scholars have described learning communities as sharing a number of common goals (Gabelnick, MacGregor, Matthews, & Smith, 1990; Lenning & Ebbers, 1999; Shapiro & Levine, 1999; Stassen, 2003).

First, learning communities are intimate, having a small number of participants and faculty, for the sake of promoting deeper, more meaningful peer-to-peer and student-to-faculty interaction than might be found outside the community. Second, they are centered on the exploration of a particular theme, which may be a specific, often interdisciplinary scholarly topic (e.g., environmental sustainability or global citizenship), a more general student concern (e.g., college transition for underrepresented students), or both. Finally, learning communities seek to integrate students' knowledge around that theme by implementing a particular curricular innovation tied to credit-bearing courses (e.g., linked courses taught solely for learning community members, or creating cohorts of learning community members through a series of general university courses) and active pedagogies (e.g., collaborative or problem-based learning).

L/L programs' particular distinctiveness comes from their intentional integration of the best elements of learning communities with students' residential experience. "The critical difference," Inkelas and Weisman (2003) noted, " … [is that participants] live together in a specific residence hall where they are provided with academic programming and services" (p. 335). This singular difference has opened up numerous avenues for broadening how, when, where, and with whom learning can occur. As practitioners have developed L/L programs, authors have cataloged a litany of possible programmatic features (Gabelnick et al., 1990; Inkelas & Weisman, 2003; Lenning & Ebbers, 1999; Shapiro & Levine, 1999; Stassen, 2003). Faculty members, for example, may teach courses and have their offices in the residence hall, and may incorporate co-curricular experiences, such as service-learning, study abroad, or internships, that take L/L participants off-campus. For their part, residential life staff may better tailor spaces for learning and student interaction, including the use of dedicated study space or equipment (e.g., computer labs or art studios) and offer theme or population-appropriate developmental programming. Finally, colleagues across campus may offer services that complement the goals of the L/L program, such as academic advising (n.b., Arms, Cabrera, and Brower, 2008), bringing their efforts into the learning community.

The Extent of L/L Programs

The number of L/L programs nationwide is unknown. An accounting by Smith (1994) in the early 1990s suggested that that "about sixty-five institutions in North America" (p. 243) were operating L/L programs. According to Smith, a third of those institutions had begun their programs in the previous 10 years. The Residential Learning Communities International Clearinghouse, maintained by Bowling Green State University, currently lists almost 250 L/L programs nation-wide (Midden, 2011). Rather than creating a simple tally of programs, the Clearinghouse allows practitioners to submit information about their programs and maintain a dynamic database of offerings from across the country. While this sort of system can provide a more contemporary accounting of L/L programs than does Smith's (1994) work, the work of Inkelas, Brower et al. (2008) demonstrates it remains incomplete.

In their on-going study of L/L programs, Inkelas, Brower, and Associates (2004) and Inkelas, Brower et al. (2008) sought to gather data about the effectiveness of L/L programs nationally. In 2004, 34 colleges and universities participated and provided data on 297 discrete L/L programs. In 2007, 52 institutions participated, and among them offered a total of 611 L/L programs. Although I cannot infer a growth rate of L/L programs using Inkelas et al.'s 2004 and 2007 studies, I can be sure that L/L programs have spread across the educational landscape. That these programs seem often found at large, public institutions is not surprising: recall that a goal of L/L programs is to build intimate communities for learners, a particular benefit for campuses interested in "making the big store small" (Shapiro & Levine, 1999, p. 4).

Irrespective of their true scope, this expansion of L/L programs means that, increasingly, institutions are funneling valuable resources their way. In 2007, roughly half of the programs cataloged by Inkelas and her colleagues provided specific information about their budgets; those that did cost their institutions a total of $8.02 million dollars, an average of $21,000 per program (Soldner & Szelenyi, 2008). Programs do not run themselves, of course, meaning there were human resource costs as well. Hundreds of people were tasked with spending some amount of their time operating L/L programs, including 332 housing/student affairs professionals and 192 faculty/staff from academic affairs units (Inkelas, Brower et al., 2008). Certainly, hundreds more educators are also involved, ranging from the Career Services staff member who offers a one-time résumé workshop to the faculty member who teaches a course every term. To be sure, L/L programs represent an investment to institutions that choose to implement them, one that many hope will solve important challenges facing undergraduate education (Jones et al., 2006).

L/L Programs' Benefits for Undergraduate Students

Institutions seeking to implement L/L programs can turn to a research literature that suggests three domains in which L/L participants may accrue enhanced outcomes when compared to their peers in traditional residence hall (TRH) environments: (a) intellectual/academic gains, (b) transition and persistence gains, and (c) psychosocial gains. Compared to their peers not participating in L/L programs, studies have suggested L/L members report better academic performance (Blimling, 1993; Stassen, 2003 Terenzini, Pascarella, & Blimling, 1996), greater levels of knowledge integration (Pike, 1999), and "significantly larger gains in intellectual orientation" (Pascarella & Terenzini, 1991, p. 245). L/L membership may also promote some forms of psychosocial growth—that is, an individual's increasingly complex understanding of all aspects of self in relation to the larger social world (McEwen, 2003)—including a student's autonomy (Leinwall, 2006) and expressed intention to work to solve civic/community concerns (Rowan-Kenyon et al., 2008). Finally, studies have suggested that L/L program participation can promote self-efficacy and self-esteem, thereby easing collegiate transition (Brower, 1997), and that L/L participants have greater rates of persistence within their major (Scholnick, 1996) and the larger institution (Pascarella & Terenzini, 1980; Stassen, 2003) than non-participants.

Given the range of benefits that the studies above have suggested students can accrue by participating in L/L programs, it is not surprising that Jones et al. (2006) have argued that many campus administrators have come to see them as “panaceas” (p. 263). In the light of the challenges facing academic professionals today, should not all campuses seek to implement this particular form of “high impact” educational practice for all students (Kuh et al., 2007)? Certainly, examples exist of just such a strategy: both Loyola College in Maryland (2008) and the University of California-Irvine (2008) have mandated L/L participation for the majority of their first-year, residential students. The evidence, however, may indicate a more measured approach.

First, it may well be that L/L programs, writ large, are not as effective as once believed. Inkelas, Vogt et al. (2006) have argued that the past evidence of L/L programs’ benefit was based on a mere “patchwork” (p. 40) of studies, conducted on small numbers of students, programs, and institutions. As a result, the generalizability of those findings to other students in other settings is in doubt. Perhaps of more concern is work by Pascarella and Terenzini (2005) contending that, irrespective of the merits of past L/L research, present research on today’s students and programs has painted a mixed picture of programs’ successes. This reevaluation represents a significant departure from the authors’ position a decade ago, in which Terenzini and Pascarella (1994) wrote “living learning-centers are not only a great idea—they actually work!” (p. 32).

Alternatively, it may be that not all types of L/L programs are equally well positioned to help students accrue each of the benefits attributed to them. Recall that all learning communities, including L/L programs, are focused on one or more central learning objectives (Lenning & Ebbers, 1999). Inkelas, Brower et al. (2004) have presented preliminary evidence that L/L programs can be incredibly diverse in terms of focus, cataloging 26 distinct themes. Subsequent research has suggested that a program's theme may be strongly related to the outcomes its students report (Inkelas, Longerbeam, Brown Leonard, & Soldner, 2005).

The notion that not all thematic types of L/L programs are equally as likely to promote all student outcomes has implications for both scholarship and practice. Take, for example, a hypothetical L/L program focused on multicultural issues. The researcher who has selected "first semester GPA" or "first year persistence" as his or her dependent variable may conclude that the program has no significant relationship to student success, while the researcher who has chosen "promoting positive diversity attitudes" as a dependent variable may uncover positive effects. The practitioner who anticipated a diversity-themed program would promote students' racial development but then assumed the program would promote academic achievement and/or persistence by virtue of its character as a L/L program may be disappointed with assessments showing developmental gains in diversity appreciation but little change in cumulative grade point average and/or graduation rates. Intentional implementation—and solid assessment—of L/L programs seems predicated upon matching an institution's outcomes of interest with those that a program of a given theme is poised to promote (Inkelas et al., 2005).

L/L Programs' Benefits for Institutions

Ostensibly, institutions have begun to offer L/L programs because of their presumed ability to enrich students' academic life and to remediate concerns about mediocre undergraduate experiences (Shapiro & Levine, 1999). It well may be that institutions are operating from this position, making L/L programs' expansion a logical outgrowth of institutions' desires to improve learning opportunities for all students. It is important to reiterate, however, that Inkelas, Vogt et al. (2006) cautioned that institutions are doing so on the basis of relatively slim research evidence, and Pascarella and Terenzini (2005) would contend that the evidence from which practitioners are proceeding is less conclusive than originally thought. Given the investment L/L programs to represent, what else might motivate institutions to act in this manner? Below, I offer two related alternatives that suggest that institutions might not be operating exclusively to benefit students.

First, institutions may simply be seeking to assuage the demands of critics, and uncritically implementing programs. Inkelas, Vogt et al. (2006) and Shapiro and Levine (1999) have acknowledged that L/L programs are one way institutions have responded to the demands of external stakeholders who are demanding greater accountability for improving the quality of undergraduate education nationwide. Although this does not mean that institutions are *not* interested in programs' ability to strengthen undergraduate education, it may signl institutions are conscious of these programs' dual utility: they may be educationally beneficial, and they provide a tangible reform to which institutions can point when external parties demand evidence of reform efforts. Indeed, one theory of organizational behavior—the New Institutionalism—posits that organizations uncritically adopt structures not because they are likely to be efficacious but because by doing so they maintain social legitimacy (Powell & DiMaggio, 1991; Scott, 1987, 2003). That both the popular press and college selection guides have cast L/L programs in a positive light may lend this innovation, and those institutions that have adopted it, increased credibility (Bonisteel, 2006; Foderaro, 2005; Thomson Peterson's, 2006).

Second, institutions may be using L/L programs as a mechanism for student recruitment. Beyond the legitimacy that operating these programs might confer to a college or university, institutions may hope to use L/Ls as the proverbial “carrots” for high-talent students. Shapiro and Levine (1999) cite one institution, the University of Maryland, which implemented one of its largest L/L programs as a way to raise the average SAT scores of its entering first-year class. This invitation-only program targets high-talent students that, with the extra incentive of membership in this select group, may be inclined to enroll at Maryland rather than one of its competitors. Given the centrality of SAT scores to external rankings of institutional prestige and the importance of rankings in today’s highly competitive admissions environment, institutional attempts at maximizing their student profile for organizational gain is understandable, if unpalatable (Ehrenberg, 2003; Meredith, 2004; Monks & Ehrenberg, 1999).

It is difficult, if not impossible, to know the consequences of these differing motivations upon educational practice. At least two questions come to mind, however. First, if institutions are implementing programs for “implementation’s sake,” then does this choice have consequences for the student outcomes they purport to facilitate? It becomes somewhat prophetic to return to Pascarella and Terenzini’s (2005) finding that, compared to studies conducted before 1990, recent research has suggested L/L programs yield fewer benefits than originally hoped. Might it be that, in their uncritical acceptance of L/L programs’ purported benefits, or in their rush to adopt a much-touted “best practice,” practitioners have failed to link the thematic focus of the programs they develop with the outcomes they desire their students will attain, as suggested by Inkelas et al. (2005)?

Second, if institutions are using L/L programs as a mechanism for recruiting high-talent students, does this restrict others' access to an educational "best practice," in particular those students who may be at increased risk? There is emergent evidence to suggest that, in at least one form of L/L program in which practitioners target high talent students, inequities in participation by virtue of race and socioeconomic status do exist. Soldner, McCarron, and Inkelas's (2007) study of first-year students in honors L/L programs found that these programs enrolled disproportionately low numbers of students who were African American or Latino, were first-generation college students, or came from families that made less than $50,000 per year. As a result, Soldner et al. suggested that honors L/L programs may represent the collegiate equivalent of the practice of tracking in elementary and secondary education, which has been derided not only for its lack of consistent benefits for tracked students, but also for the possible psychological harm low-track students accrue via stigmatization (for critiques, see Hallinan, 1994; Loveless, 1999; Oakes, 2005). Although not in a position to assess possible harms, Soldner et al. did conclude that underrepresented students who were able to participate in honors L/L programs attained benefits their non-participating peers did not. In this case, expanding access to honors L/L programs—or at least to the practices and pedagogies they employ—seemed to hold the potential to benefit students least well served by the status quo.

For those seeking to remove barriers to the success of students who identify as racial/ethnic minorities or who come from low-SES backgrounds (Cook & Cordova, 2006; Terenzini, Cabrera, & Bernal, 2001), or those who are demanding a more rigorous undergraduate education for all students (AAC&U, 2002; NASPA & ACPA, 2004;

NSSE; The Spellings Commission, 2006), that status quo is increasingly untenable. Although proponents have offered L/L programs as possible mechanisms for change (Lenning & Ebbers, 1999; Shapiro & Levine, 1999), questions about equity and efficacy (and the interaction between the two) remain. To date, the latter has only just begun to be comprehensively addressed, the former remains largely unexplored. Understanding both is important if, as suggested by Kuh et al. (2007, p. 70), we are to “take what we know works”—if, in fact, it does—and ensure those students who might stand to benefit have the opportunity to do so.

Questions for Research in This Study

This study has three purposes that, when brought together, address a substantive research question that seeks to better understand equity and student outcomes in L/L programs. First, it seeks to create a thematic typology of living-learning programs to facilitate comparisons between like programs. Second, it seeks to identify whether undergraduate students from different racial/ethnic and socioeconomic backgrounds participate equitably in L/L programs, relative to their presence within the larger L/L community. Third, for underrepresented students, it seeks to determine the benefits associated with participation in L/L programs versus living in a traditional residence hall (TRH) environment. Having done so, the study addresses the following substantive research question:

> *Do students from various racial/ethnic and SES groups who are underrepresented in a particular thematic type (or types) of L/L programs report better outcomes than their peers living in the TRH environment.*

To that end, three instrumental research questions are posed, specifically:

1. First, what type of thematic typology of L/L programs can be created based upon their stated goals and objectives?

2. Next, do students from different racial/ethnic backgrounds or socioeconomic statuses participate in various thematic types of L/L programs at a rate proportional to their presence in the larger L/L community?

3. Finally, considering the outcomes upon which L/L programs of different thematic types focus, do students underrepresented in a given thematic type (or types) of L/L program accrue differential benefits by participating in L/L programs when compared to their peers living in traditional residence halls, driven by six key measures representing peer, faculty, and hall environments and holding other characteristics constant?

Significance of the Study

The hope of this study is to benefit scholars who research the effectiveness of L/L programs, practitioners who implement and operate them, and students who participate in them. To do so, it does three things. First, it uses a team of raters to craft a rigorously designed thematic typology of L/L programs, replacing that developed previously by Inkelas et al. (2004, 2005). Second, it is the first large-scale study of equity in L/L programs, addressing in part Kuh et al.'s (2007) question as to whether will exists to offer high impact educational programs to all students. Finally, because this study is concerned with the unique experience of L/L participants who are members of groups otherwise underrepresented in these programs, it seeks to identify the benefits underrepresented students can attain through their membership in an L/L program compared to their peers living in traditional residence hall environments.

Benefits to higher education researchers. Initially, the strengths of this work provide methodological guidance for those who wish to conduct studies focused on the outcomes associated with participation in a wide range of educational interventions. For example, it demonstrates the use of multiple imputation to overcome missing data, confirmatory factor analysis to improve the specification of latent constructs, structural equation modeling to consider direct and indirect effects, multiple group analysis to demonstrate where effects are (or are not) invariant across groups, and latent mean modeling to attenuate measurement error and overcome restrictive assumptions associated with traditional analysis of variance or linear regression techniques. Its limitations provide fertile area for further research, and underscore the importance of careful planning during data collection and the need to use more advanced research designs to make stronger statements about treatment effects.

For researchers dedicated to exploring issues of social justice and who are ensuring equal access to high-quality educational experiences for students, this study demonstrates the application of a simple method for characterizing equity that can be applied to a wide range of outcomes. While it is used here to evaluate whether the share of a student population in a given type of L/L program is proportional to that population's representation within the larger L/L population, it could as easily be used to explore equity in college access, field of study, graduation, or employment. Additionally, it extends the literature on equity within what appears to be an increasingly popular feature of campus life, expanding upon Soldner et al.'s (2007) work, which focused solely on honors L/L programs. Taken together, these may spur additional research on equity and inclusion in other domains of campus life.

Finally, because it provides L/L researchers guidance in how they might more precisely consider the outcomes that they should reasonably expect to be associated with a particular form of L/L program, an updated typology can improve the quality of the scholarly narrative about programs' benefits and efficacy. As suggested by Inkelas et al. (2005), given the diversity in foci of today's L/L programs, the universal attribution of a non-specific set of outcomes to all programs may be problematic.

Benefits to student affairs practitioners. Although focused on L/L programs, it is hoped the chief benefit this work affords student affairs practitioners is that it provides impetus and license to question participatory equity throughout the university community. Given the magnitude of problems L/L programs have been called upon to solve and the level of resources dedicated to them, making visible trends in their implementation that might be eroding opportunity is not only an educational imperative but also an ethical and fiduciary one.

However, L/L programs are only a small part of the larger higher education enterprise, and evaluating equity within them is no more or less important than in the myriad other opportunities colleges afford their students. Who, for example, is participating in leadership development, student government, or service learning programs? Or, could it be that interventions targeting underrepresented minority students serving to remedy one inequity while ignoring others, such as social class? Indeed, while it may be that equity issues vis-à-vis students is the chief concern for many in student affairs, questions of fairness extend throughout the entirety of the university community, and this demonstrates how those questions can be explored through quantitative measurement.

Although it only addresses a subset of student populations and outcomes, a benefit of this research for residential educators is that it provides additional information about what facets of L/L and TRH environments appear to be most strongly related to student learning and development. In particular, six key living and learning environments are the focus of this work: (a) academic and vocational conversations with peers, (b) socio-cultural conversations with peers, (c) course-related faculty interaction, (d) non-course-related faculty mentorship, (e) hall academic climate, and (f) hall social climate.

While this study's findings should not be taken to suggest that one component of a residential environment—say, non-course-related faculty interaction—can be discounted in wholesale fashion in lieu of a potentially more potent lever for growth, they may suggest ways to prioritize finite resources. To the extent that this study's findings may contradict prevailing wisdom about what makes L/L programs "work," it may also call upon practitioners to wrestle with thorny questions about whether their operating theories are wrong, implementation of those theories is suboptimal, or, perhaps, both.

Finally, it should be noted that the six key living and learning environments listed above have analogs in non-residential settings. Indeed, that educationally powerful environments can exist campus-wide is the driving principle behind works such as *Learning Reconsidered* (American College Personnel Association & National Association of Student Personnel Administrators, 2004) and *Learning Reconsidered 2* (Keeling, 2006). As such, findings vis-à-vis any relationships between peer and/or faculty interactions (which are hardly unique to the residential setting) and outcomes attained by L/L and TRH students may prove useful to student affairs professionals who are interested in cultivating those same outcomes in non-residential settings. To the extent that they can be generalized to other campus settings, findings related to residence hall climate measures may be similarly helpful: presumably, practitioners can (and likely do) already work to shape other environments to be socially and academically supportive, and this study may help identify outcomes for which those efforts can be particularly impactful.

Benefits to students. To the extent that researchers or practitioners engage with the findings and implications of this work, students stand to benefit in three ways. Initially, current and potential L/L participants may benefit as program staff reshape admissions requirements and priorities in ways that increase participatory equity. In so doing, currently underserved student groups could access potential benefits that accrue to L/L participants and, to the extent to which students benefit from exposure from diverse living and learning environments, all L/L participants stand to benefit.

Second, although TRH environments are not the explicit focus of this work, this study's third research question does explore the relationship between key living and learning environments and student outcomes in non-L/L environments. As practitioners increase their knowledge about which components of TRH students' residential experiences can help promote learning and development, they can apply that knowledge to strengthen the default environment, thereby benefitting all resident students.

Third, even though this study focuses on participatory equity in only one form of high-impact program, it may serve to call attention to equity across an institution's entire slate of educational offerings. This may serve to increase access to a wide range of programs, benefitting students who would never participate in L/L programs but might choose to study abroad, engage in undergraduate research, or one of any number of other educationally-purposeful programs.

The importance of research on equity in L/L programs. To date, the extant literature on equity in L/L programs is entirely comprised by the work of Soldner et al. (2007), a narrowly tailored investigation that considered only residential honors programs. As such, nothing is known about equity in other forms of L/L programs, or even whether the authors' findings were idiosyncratic to a given year, type of program, or data collection. Until additional work on L/L programs is conducted, making anything but the most qualified statements about equity and L/L programs seems ill-advised.

There is reason to believe, however, that the present study will identify both that additional participatory inequities exist and that some groups of students are missing out on opportunities to accrue meaningful learning and development gains. In a NSSE-based study of high-impact practices—which included learning communities but not L/Ls—Kuh (2008) concluded that first generation and African American college students participated at notably low rates, and that a relationship existed between the number of high-impact practices in which a student participated and four key outcomes: (a) GPA; (b) first-year persistence; (c) deep learning; and (d) general, personal, and practical gains.

Like the work of Soldner et al. (2007), Kuh's (2008) effort appears to be the only one of its kind. As a result, the present study stands to extend what is known about student participation in high-impact (and related) programs and the outcomes they attain. Notably, the absence of equity-focused research is in marked contrast to elementary and secondary education, where a robust related literature does exist, addressing such topics as tracking on the basis of ability (Hallinan, 1994; Loveless, 1999; Oakes, 2005), access to AP course-taking and advanced curricula (Darity, Castellino, Tyson, Cobb, & McMillen, 2001), and dual enrollment programs (Museus, Lutovsky, & Colbeck, 2007). Too little is known, it seems, about students' equitable participation in programs that could stand to benefit those for whom college success is far from guaranteed.

Definitions Used in this Research

The following operational definitions were employed in this research.

- *Living-learning (L/L) programs* were defined as those interventions "that involve undergraduate students who live together in a discrete portion of a residence hall (or the entire residence hall) and participate in academic and/or extra-curricular programming designed especially for them" (Inkelas et al., 2004, p. I-1). Importantly, L/L programs were "self-identified" by institutional participants in the NSLLP.
- A *traditional residence hall (TRH) environment* was one that does not host a L/L program.
- In this study, *race/ethnicity* was defined as students' identification with one or more of the following identifiers: (a) African American/Black, (b) Asian or Pacific Islander, (c) American Indian or Alaskan Native, (d) Hispanic/Latino, or (e) White/Caucasian. Students who identified with more than one racial/ethnic category were categorized as "Multiracial/multiethnic." Because the racial/ethnic identification of students who responded "Race/ethnicity not included above" cannot be determined with any degree of certainty, these students were excluded from analysis.

- In this study, *SES* was defined as a composite of students' self-reported family income and their parents' educational attainment, equally weighted. *Low, medium, or high SES* represented a student's membership in one of three SES composite groups separated by the 33rd and 67th percentile.
- For a given group, *participatory equity* was defined in this study as having an equity index greater than 0.80 (Bensimon et al., 2003). More detail on the equity index is provided in Chapters 2 and 3.
- *Underrepresented students in a type of L/L program*, or "underrepresented students" more generally, are students of a given race/ethnicity or SES background belonging to a group that does not have participatory equity within a particular thematic type of L/L program.

This Research in Context: The National Study of Living Learning Programs

This research took place within the context of a much larger effort, the National Study of Living-Learning Programs (NSLLP). The NSLLP is a multi-year, multi-institutional study of the efficacy of L/L programs led by principal investigator Dr. Karen Kurotsuchi Inkelas and supported by a research team based at the University of Maryland. The present research uses two NSLLP datasets collected in 2007: one consisting of information provided by resident students about their collegiate experiences, and another consisting of a detailed census of participating institutions' L/L programs and their characteristics. Because I reference them frequently in this dissertation, and review several studies stemming from the NSLLP's initial administration, I describe the 2004 and 2007 NSLLP in some detail below.

The NSLLP began in 2001 with a commissioned grant from the Association of College and University Housing Officers-International (ACUHO-I). The ACUHO-I grant funded the development of an initial survey instrument, piloted at four institutions in Spring, 2003, and the testing of data collection techniques (Inkelas et al., 2004). In 2004, the NSLLP expanded to include 34 institutions located in 24 states and the District of Columbia. At that time, an external survey research firm (MSIResearch) was engaged to assist the NSLLP team with a variety of administrative tasks, including (a) assisting in \school recruitment and management, (b) translating paper survey instruments to web-based forms, (c) soliciting the participation of student respondents via electronic mail, (d) gathering information from participating institutions about their L/L offerings and the characteristics of those programs, and (e) data clean-up and data file preparation.

The 2004 NSLLP consisted of two elements: (a) a student data collection, and (b) L/L program data collection. Participating institutions were asked to provide a full or random sample of students residing in L/L programs, as well as a comparison sample of students living in TRH environments that were matched on the basis of race, gender, academic class standing, and residence hall. A total of 71,728 students were selected and invited to complete the Residence Environment Survey (RES) via the Web. The RES is a 258-item survey covering a wide range of student characteristics and experiences, including (a) demographic information, (b) pre-college experiences and expectations, (c) engagement with residential and non-residential collegiate environments, and (d) self-reported gains on a variety of important collegiate student learning outcomes. A total of 23,910 students responded, yielding an overall response rate of 33.3%. The sample consisted of 12,241 students participating in an L/L program (a 36.5% response rate) and 11,669 students who were living in a TRH environment (a 30.6% response rate).

The second component of the 2004 NSLLP focused on gathering information about the L/L programs in which L/L participants lived. Administrative contacts at each institution were asked to complete the Living Learning Program Survey (LLPS) via the Web, a 23-item instrument that gathered information such as (a) program goals and objectives, (b) program staffing and funding, (c) use of faculty and staff, and (d) optional and required activities for participants. A total of 297 programs registered with the NSLLP, of which 268 (90.2%) completed the LLPS.

In 2007, the NSLLP was repeated with support of the National Science Foundation, ACUHO-I, the National Association of Student Personnel Administrators (NASPA), and ACPA: College Student Educators-International. As in 2004, the principal investigator retained a survey research firm (Survey Sciences Group [SSG]) to assist with a variety of administrative tasks related to data collection, as described above. The 2007 NSLLP consisted of a longitudinal follow-up with 2004 participants, a new baseline student data collection, and a new census of participating institutions' L/L programs. This dissertation relies solely upon data collected as part of the new student baseline and the new L/L program census in 2007. A total of 47 colleges and universities offering 654 programs participated in the 2007 NSLLP baseline study, including 38 classified by the Carnegie Foundation as doctoral extensive or intensive institutions, 8 classified as master's institutions, and 1 classified as a baccalaureate institution. Each was a four-year, predominantly White institution, and many were the flagship campuses of their home state's university system. More detailed information about the 2007 NSLLP appears Chapter 3, in which the methods of this study are outlined.

Chapter Conclusion

It seems clear that L/L programs have captured the imaginations of collegiate educators interested in improving the undergraduate experience and, perhaps, administrators seeking to strengthen their institution's position within the higher education marketplace (Inkelas & Weisman, 2003; Shapiro & Levine, 1999; Soldner, McCarron, & Inkelas, 2007). In this respect, they hold the prospect of realizing the goal set forth by Kuh et al. (2007) of providing high impact educational experiences to large numbers of students, particularly those who are underserved in current educational environments. Unfortunately, emergent scholarship has questioned these programs' equity, challenging whether institutions are making this purported "best practice" equally available to all students, especially those who might be in the greatest need of academic assistance (Soldner et al., 2007). Furthermore, despite their proliferation, practitioners still know relatively little about their efficacy, especially for particular groups of students (Inkelas et al., 2004; Inkelas, Brower et al., 2008; Pascarella & Terenzini, 1991, 2005). Complicating matters, these programs vary widely in their thematic focus, suggesting that, by design, not all programs will be equally as adept at achieving all the goals that might be set before them (Inkelas et al., 2005; Lenning & Ebbers, 1999; Longerbeam, 2005; Pascarella & Terenzini, 2005; Shapiro & Levine, 1999).

In short, L/L scholarship is still evolving, and much remains to be examined. This dissertation represents another in a series of organized efforts to understand this intervention more fully, deepening researchers' theoretical grasp of residence education efforts and strengthening practitioners' efforts to harness those efforts for the benefit of all students. In the next Chapter, I review a wide range of literature designed to ground my effort.

Chapter 2 : Review of the Literature

This study has three purposes. First, it seeks to use a team of raters to develop a thematic typology of living-learning programs on the basis of those programs' stated goals and objectives. After categorizing L/L programs in the study in to thematic groupings, it seeks to determine whether all students—particularly students from traditionally non-dominant racial/ethnic or socioeconomic backgrounds—have access to L/L programs, as evidenced through their equitable rates of participation in L/L programs of different types. Finally, it attempts to quantify the benefits underrepresented students not participating in L/L programs might stand to gain through L/L membership.

In this Chapter, I present a review of the literature that grounds this research study. The review is in two parts, the first dedicated to analyses related to equity in L/L programs, and the second focused on assessing L/L program efficacy. Part one begins with the introduction of a preliminary conceptual framework for evaluating equity in educational settings that is growing in its frequency of use. Then, I review how authors have applied that framework to questions of equity in L/L participation.

In part two of this review, I turn to a growing body of literature related to the consequences of L/L program participation. I begin by outlining two conceptual frameworks that have particular salience for the investigation of this issue. The first presents a generic, linear model of college impact. The second, more narrowly tailored, considers how L/L programs change the relationships among educationally powerful environments already inherent in campus residence halls. Next, I review the literature related to a methodological issue that researchers should consider when designing L/L-focused inquiry: the influence of programs' thematic type. Studies leading to, or stemming from, investigations of type are then outlined. The second part of this chapter concludes with my review of published studies that have explored the outcomes associated with L/L participation. Three classes of outcomes corresponding broadly to the benefits traditionally ascribed to L/L participation are considered: (a) cognitive outcomes, (b) transition and persistence outcomes, and (c) psychosocial outcomes. Within each class, I begin by outlining findings related to the main effects of L/L participation and then turn to the identification of independent/predictor variables related to studies' outcomes of interest.

Assessing Equity in L/L Programs

As noted in Chapter 1, the assessment of participatory equity within L/L programs is in its infancy. There is evidence, however, to suggest that both researchers and practitioners should be concerned about whether this purportedly beneficial intervention is reaching students from all socio-demographic groups in an equitable manner (Kuh, 2008; Soldner et al. 2007). Next, I review the literature related to equitable participation in L/L programs. To begin, I introduce a conceptual framework for quantifying equity. Then, I review the only study to date explicitly exploring participatory equity within L/L programs, an example of that conceptual frame put to practical use.

A conceptual framework for analysis: Hao's Equity Index. Hao's (2002) Equity Index provides researchers an indicator of whether members of a given group have attained a particular outcome at a rate comparable to their presence in the larger population. The numerator of the EI expresses the proportion of the total population of "outcome attainers" represented by members of the group of interest. The denominator of the EI is the proportion of the group of interest within the total population. The formula for the EI appears below.

Equation 2.1

The Equity Index (Hao, 2002)

$$EI = \frac{{Group\ of\ Interest, Attained\ Outcome}/{Population, Attained\ Outcome}}{{Group\ of\ Interest, Total}/{Population, Total}}$$

Bensimon et al. (2003) provide an example of the index's use by computing an EI for Latino students' enrollment in post-secondary education. Their example is set in a high school with a graduating class of 1000, of which 400 are Latino students. As a result, the denominator of the EI is 0.40. From this fictitious institution, 450 students go on to college, 45 of which are Latino. As a result, the numerator of the EI is 0.10. Dividing the numerator by the denominator yields an EI of 0.25, as demonstrated below.

Equation 2.2

An Example of Using the Academic Equity Index (Bensimon et al., 2003)

$$EI = \frac{{45}/{450}}{{400}/{1000}} = \frac{0.10}{0.40} = 0.25$$

An EI score of 1.0 indicates equity; that is, the group is attaining the related outcome at a rate proportional to its presence in the larger population. Bensimon et al. characterize EI scores of between 0.8 and 0.9 as meaning the group's attainment is "almost at equity" (p. 14). Increasingly lower scores are indicative, then, of increasing inequity on a given outcome.

The use of the EI framework in L/L research. To date, only one study has used equity indices in the study of L/L programs: Soldner et al.'s (2007) investigation of honors L/L programs. Honors L/L programs fuse the concepts of collegiate honors programs and L/L programs to create residential environments that are specifically designed for enhancing the educational experience of high-talent undergraduate students (Inkelas et al., 2004; Soldner et al.). Soldner et al. questioned whether honors L/L programs might, as interventions designed to provide further enrichment to academically gifted college students, exhibit characteristics and outcomes similar to the analogous effort at the K-12 level to direct high talent students to more advanced classes. This practice, known as tracking, had been roundly critiqued by authors for disproportionately favoring students from more privileged racial and socioeconomic positions, perpetuating existing social inequality, and failing to provide uniform benefits to any student group (Hallinan, 1994; Loveless, 1999; Oakes, 2005).

Soldner et al.'s (2007) work focused on the 23 honors L/L programs that had participated in the 2004 NSLLP. To begin, they computed equity indices for non-honors L/Ls and honors L/Ls based upon three different demographic variables: race, parental education, and family income. Although no inequities were noted among non-honors L/L programs on any variable, honors L/Ls exhibited inequities on each. African American students were found to be underrepresented in honors L/L programs (EI=0.347), as were Latino students (EI=0.469), students whose families made less than $50,000 per year (EI=0.735), and students whose households were not headed by at least one college graduate (EI=0.611).White students (EI=1.072), students whose families made more than $100,000 per year (EI=1.179), and students whose households were headed by at least one post-baccalaureate degree holder (EI=1.337), were, to varying degrees, overrepresented.

Next, Soldner et al. (2007) selected only those students exhibiting at least one marker of under-representation in their sample (that is, African American students, Latino students, lower-income students, and students from families with histories of lower educational attainment) for further analysis. To quantify gains on a series of collegiate outcomes that these students could be predicted to attain by participating in an honors L/L program, the authors constructed regression equations that controlled for race, gender, parental education, income, high school performance, and outcome pre-tests. Although underrepresented students were not predicted to differentially benefit from honors L/L programs on measures of academic transition, social transition, and sense of belonging, gains were predicted on growth in critical thinking (SD = .117) and college GPA (SD = .473). The authors found that participation in honors L/L programs was predicted to slightly depress students' scores on a measure of diversity appreciation (-0.052 SD), which they hypothesized might be due to a lack of structural diversity within the programs themselves.

Finally, after concluding that inequities existed, resulting in tangible benefits being withheld from students in underrepresented groups, Soldner et al. (2007) explored factors related to the administrative practices of honors L/L programs that might have resulted in participatory inequity. Driven by critiques from the K-12 literature that standardized test scores served as gatekeepers to enrichment programs in elementary and secondary education settings (Hallinan, 1994; Loveless, 1999; Oakes, 2005), the authors explored whether programs' admissions criteria might be related to the participatory equity they achieved. Their analyses of programs' use of SAT scores, either as an exclusive indicator of merit or as part of a more inclusive strategy including other markers, suggested that those programs that relied only upon SAT scores typically evidenced the poorest racial equity indices. Of even greater concern to Soldner et al. was their finding that institutions purporting to use SAT scores as their sole criteria for admission to honors L/L programs often had large numbers of White students with comparatively low SAT scores participating in honors L/Ls, while racial/ethnic minority students with comparatively high SAT scores were not participating, living instead in TRH environments.

Although Soldner et al.'s (2007) work represents a first step in investigating equity issues associated with L/L programs, and benefits from its drawing together of data from a variety of institutions, it has several limitations. First, although there is no evidence of response bias in the 2004 NSLLP (Inkelas et al., 2004), any undetected bias diminishes the precision of EI estimates. Given the particularly low EI indices associated with different racial/ethnic and socioeconomic groups, however, it seems unlikely that the inequities identified would disappear entirely. Second, the authors used a particularly parsimonious regression model to estimate only the direct effects of honors L/L participation. It is likely that indirect effect models that incorporated a wider range of collegiate environments, such as those suggested by Pike (1999), would yield more accurate conclusions. Finally, Soldner et al.'s work only considered honors L/L programs. Researchers have yet to disaggregate the status of equity in other types of L/L programs, and as such it is impossible to determine whether Soldner et al.'s findings were idiosyncratic or indicative of the state of equity in L/L programs generally.

Other treatments of equity in L/L research. To date, only one other study appears to have explored issues of equity in L/L program participation. Using data collected by the NSLLP in 2004 and completed before Soldner et al. (2007), Johnson (2007) explored L/L participation among women who were majoring in science, technology, engineering, or mathematics (STEM). She noted that women-only STEM-focused L/L programs evidenced particularly low rates of participation by non-White students and argued that “women of color are not accessing resources such as STEM LL programs that could facilitate their success in these majors” (p. 133).

Although Johnson (2007) did not calculate equity indices herself, it is possible to estimate them based upon her data (see Table 3, p. 110). Johnson’s conclusion vis-à-vis women-only STEM-focused L/L program, reached by an observation of percentage participation, is borne out by the EI calculation: only 17.6% of participants in these programs are women of color, while women of color make up 24.2% of the L/L population in her sample (EI = .727). In contrast, women of color were near equity in co-ed STEM programs (EI = .229 / .242 = .946) and slightly above equity in all other L/L programs (EI = .258 / .242 = 1.066). From this analytic perspective, it may be more appropriate to say women of color in Johnson’s sample were not accessing *women only* STEM L/L programs at expected rates, not STEM programs (or L/L programs) more generally. Nonetheless, as the first discernable example of equity-interested research in the L/L scholarship, Johnson’s work represented an important first foray in to this line of analysis.

Assessing the Outcomes of L/L Programs

Proponents of L/L programs have commented upon the thin nature of the empirical evidence supporting this type of intervention (Inkelas & Weisman, 2003; Jones et al., 2006), and some have suggested that the evidence is trending in the direction of less positive outcomes, rather than more (Pascarella & Terenzini, 2005). In the next section of this literature review, I address recent scholarship related to the outcomes of L/L participation. Excluded from this review are works that focus on learning communities, which are non-residential. I begin by introducing two conceptual frameworks that past researchers have used to guide their efforts. Then, I review the literature related to three broad classes of L/L outcomes: (a) cognitive, (b) transition and persistence, and (c) psychosocial. At its conclusion, the section summarizes the state of L/L outcomes literature in preparation for Chapter 3, a discussion of this study's proposed methods.

Conceptual frameworks for analysis. Shields and Tajalli (2006) characterized conceptual frameworks as "intermediate theory" (p. 313). That is, when applied to the investigation of a particular phenomenon, clearly articulated conceptual frameworks identify for the reader and the researcher how the various concepts and constructs under investigation eventually come together to form a coherent whole. Below, I identify two conceptual frameworks that guide this study's investigation of the outcomes associated with L/L participation. The first, Astin's (1991) model of college impact, is a generic model that is applicable to a wide range of student outcomes that may come about through students' engagement in any facet of the university environment. The second, Pascarella and Terenzini's (1980) model of structural mediation, is narrowly targeted to the influence of residential education efforts.

Astin's IEO framework. Astin's (1991) inputs-environments-outcomes (IEO) model offers a generic model for understanding student change while in college, as well as a method for quantifying the relationships between that change and potentially causative factors. In his IEO model, Astin argued that the outcomes of college are a function both of a student's entry characteristics and his or her experiences on campus. By relying upon statistical procedures that can control for the influence of entry characteristics and can partial out the differential contribution of collegiate environments that are successively more proximate to the outcome of interest, the IEO model allows the researcher to draw more definitive conclusions about the net effect of college and the strength of relationship between that effect and a student's experiences.

A fundamental assumption of Astin's (1991) IEO model is that who a student is when he or she enters college will have some influence on the outcomes he or she attains. Researchers who are primarily interested in the influence of environments, then, must address possibly confounding entry characteristics. Because the IEO model is generic in terms of the outcomes to which it can be applied, existing theory and research related to the outcome of interest—like that reviewed in the next section of this chapter—is used to provide the researcher insight as to which inputs are most likely to be influential and merit statistical control. Because it should be the variable most strongly correlated with the outcome of interest in any conceptual framework, a pre-test of that outcome is an important feature in any analysis using the IEO model.

Typically, researchers using the IEO model are interested in the influence of one or more college environments (Astin, 1991), net of differences caused by student inputs. Astin identified six broad categories of environments, organizing them from most distal in their influence on the outcome of interest to most proximal: (a) residence, academic major, and financial aid, (b) institutional characteristics, (c) curricular features, (d) faculty behaviors and attitudes, (e) peers, and (f) student academic, faculty, peer, and co-curricular involvement. This first set of environments is often referred to as *bridge variables*, because while they are features of the college environment that will subsequently shape the student's experience, they are the consequence of student pre-college characteristics (such as socioeconomic status) or the results of decisions made before entry (such as a student's interests). Those environments that follow may include any element of the student experience, but for reasons described more fully below their sequencing within the model should reflect the proximal ordering suggested by Astin or relevant theory.

The final element of the IEO model is the actual outcome of interest (Astin, 1991). Astin noted that collegiate outcomes could be located within a three-dimensional taxonomy: (a) cognitive vs. non-cognitive, (b) psychological vs. behavioral, and (c) short-term/during college vs. long-term/after college. Just as the distinction between input and environment is sometimes blurry, so too is the line between environment and outcome. *Intermediate outcomes* are those consequences of college attendance that are also environments in and of themselves (take, for example, Astin's example of friendship groupings).

After developing the initial conceptual framework, a researcher employing Astin's (1991) IEO model uses ordinary least squares regression in a hierarchical (that is, block entry) manner to evaluate it. To do so, the researcher first enters student entry characteristics into the model, followed by bridge variables, blocks of environmental variables in order of their increasing proximity to the outcome of interest, and any intermediate outcomes that are to be explored. By building the regression model in this manner, the researcher can identify the influence of a given variable on the outcome of interest, while holding other predictor variables that preceded it constant.

Astin’s (1991) IEO model has been explicitly referenced by a number of researchers interested in the influence of L/L participation on student outcomes, such as Stassen (2003), Inkelas and Weisman (2003), Longerbeam (2005), Pasque and Murphy (2005), and Rowan-Kenyon et al. (2008). Although it is the dominant conceptual frame among studies reviewed below, it is important to note that because Astin’s IEO model offers the researcher significant latitude to include a wide array of input and environmental variables and to explore any student outcome, how authors have chosen to operationalize it varies widely.

Astin’s (1991) IEO model has several notable strengths. First, it explicitly acknowledges not only the contribution of a student’s collegiate experience to his or her growth, but also the background and pre-college characteristics that he or she brings to campus. Its scheme of ordering environmental variables from most distal to the outcome of interest to the most proximal offers two benefits. Initially, doing so enforces temporality, accommodating both longitudinal designs as well as designs that are cross-sectional but employ quasi-pretests (as described in Pascarella, 2001) or ask the respondent to provide retrospective data. Second, the distal to proximal arrangement called for by Astin, when implemented using block entry multiple regression, tends to systematically overestimate the contributory power of variables appearing early in the model and underestimate those appearing later. These more conservative estimates of the influence of college environments minimize the chance of attributing an effect to a particular intervention when one does not exist. Finally, the IEO model allows the researcher to estimate both the contribution of every independent variable in the model to the eventual prediction of the dependent variable and its significance. As a result, it offers the potential to gain significant insight into how a system of variables tend to work together to explain variance within measures of a particular outcome of interest.

Despite its benefits, a critique of Astin's (1991) IEO model may exist, at least in terms of its use to assess the efficacy of L/L programs. In their classical form, IEO models employ linear regression and, as such, estimate only the direct effects of an independent variable upon the dependent variable of interest. What, however, of indirect effects? As noted earlier, L/L programs are complex interventions that, it has been argued, promote growth by intentionally shaping students' peer, faculty, and residence hall environments (Lenning & Ebbers, 1999; Shapiro & Levine, 1999). As a result, the most effective method for assessing the total effects of L/L program participation may be one that considers not only their direct relationship to student gains but also their indirect relationships, moderated by the collegiate environments that L/L programs purport to influence. Indeed, evidence for this notion has existed for almost 30 years.

Pascarella and Terenzini's structural mediation model. Support for L/L programs' indirect relationship with student outcomes was first noted in a study conducted by Pascarella and Terenzini in 1980. In that work, the authors employed a three-step process to conclude "structural and organizational characteristics" (p. 351) of L/L programs shifted the peer and faculty environments in which students engaged, ultimately mediating the direct effects of participation. I describe their research and resulting conceptual framework in more detail below.

In the 1976-1977 academic year, Pascarella and Terenzini (1980) tracked the experiences of one institution's first year students, with the hope of developing a greater understanding of residential environments' contribution to the attainment of important student outcomes. They collected data in two phases. Before students' arrival on campus, Pascarella and Terenzini surveyed a random sample of 1,905 new students to gather demographic data and information about students' pre-college expectations. The authors attained a response rate of 76.5%, or 1,457 students. In March of the following year, respondents were re-surveyed to gather data on their collegiate experiences to date and scores on several outcome measures. At the conclusion of the second collection, the authors had received a total of 773 valid responses. Sixty-five (8.4%) respondents reported they had lived in an L/L program, while 708 (91.6%) reported they lived in TRH environments. Although Pascarella and Terenzini did not present descriptive statistics about the respondents, the researchers noted that no statistically significant demographic differences existed between the students originally surveyed and the final sample.

The authors' analysis took place in three stages (Pascarella & Terenzini, 1980). First, they constructed regression equations to determine the effect of L/L participation on 10 "process variables" (p. 348). These variables included indicators of the amount and quality of formal and informal faculty interaction, a measure of students' perceptions of faculty members' concern for teaching and student development, and indicators of students' frequency and quality of peer interactions. After controlling for student background characteristics, Pascarella and Terenzini found statistically significant relationships between L/L participation and only three process variables: (a) interaction with faculty, (b) faculty members' concern with teaching and students' development, and (c) peer group interaction.

In the second phase of their research, Pascarella and Terenzini (1980) constructed regression equations to determine the effect of L/L participation on eight outcome variables, including (a) persistence, (b) intellectual development, (c) personal development, (d) first-year GPA, (e) intellectual press, (f) sense of academic community, (g) nonacademic press, and (h) sense of nonacademic community. After controlling for student background characteristics, statistically significant relationships were found between L/L participation and five student outcomes: (a) persistence, (b) intellectual development, (c) personal development, (d) nonacademic press, and (e) nonacademic sense of community.

In the final phase of their research, Pascarella and Terenzini (1980) combined the results of their two previous analyses. New regression equations for the five outcome variables found to have a statistically significant relationship with L/L participation were constructed, each consisting of student background characteristics, a variable indicating L/L participation, and the three process variables found in phase one to have had a statistically significant relationship with L/L participation. Analysis revealed that the once significant relationships between L/L participation and the five outcome variables of interest disappeared, only to be replaced by significant predictive relationships between the process and dependent variables. As a result, Pascarella and Terenzini concluded that process variables mediated the effect of residence arrangement. That is, "organizational differences [being in a L/L vs. a TRH environment] … create substantially different interpersonal environments" (p. 351), which, ultimately, influence student outcomes.

The lesson that researchers and practitioners can infer from the work of Pascarella and Terenzini (1980) makes an important contribution to our understanding of how L/L programs operate. Although it was the case that L/L programs had a direct effect on student outcomes (phase two of their research), it was also the case that L/L programs had a direct effect on faculty and peer environments that are believed to contribute to learning (phase one of their research). Indeed, at least within the context of their study, Pascarella and Terenzini demonstrated that L/L programs' direct influence on environments was more powerful than their direct influence on student outcomes. These findings seem to make a powerful case for the consideration of L/L programs' direct and indirect effects on learning outcomes, demanding analyses beyond those traditionally associated with Astin's (1991) IEO model.

Guiding further inquiry: Programs' thematic focus. As noted above, the consideration of Pascarella and Terenzini's (1980) study can refine researchers' thinking about how to best apply Astin's (1991) IEO model to an exploration of the outcomes associated with L/L participation. Other methodological advice comes in the form of a series of studies that explore the relationship between L/L programs' themes and the outcomes their participants attained.

Clarke, Miser, and Roberts's analysis of theme. Clarke, Miser, and Roberts (1988) were among the first L/L researchers to have considered the influence of a program's theme on the eventual outcomes reported by participants. In their single-institution study of 197 first-year students, the authors employed a 2 × 2 × 2 ANCOVA design to assess group differences among eight distinct residential programs that varied on three dimensions—designation as an L/L, the presence of a formal theme, and the involvement of faculty—on several modified CSEQ scales, holding a measure of students' locus of control constant. Although the authors reported a number of findings, only those related to theme are of particular relevance here. Clarke et al.'s analysis identified a statistically significant main effect for theme on four of the outcomes under investigation, with students who lived in a community with a formal theme reporting (a) more time spent in study groups, (b) a greater number of meaningful conversations with peers, (c) greater arts appreciation, and, interestingly, (d) less satisfaction with friendship relationships.

Several limitations of Clarke et al.'s (1988) work are worthy of mention. First, the authors provided little information about their sample, and how respondents were distributed between programs. Second, scant details were offered about the programs themselves, yielding questions about what was truly being studied. Finally, the authors did not report any psychometric information about their instrumentation, a locally derived version of an existing measure. Flaws aside, the authors' summative contention that the "specific components of residential programs appeared to generate effects consistent with their intended purposes" (p.11) yields an important hypothesis for researchers: if the focus of an institution's living-learning program is on career themes, for example, it might be reasonable to expect career-related outcomes. The opposite may also be true: if a program does not focus on a particular theme, expecting that type of outcome may be unreasonable.

Unfortunately, Clarke et al. (1988) did not ultimately test this hypothesis. Although the authors were interested in program theme as an independent variable, their measure of it was simple: a program either had a theme or it did not. A dichotomous treatment of theme (i.e., either a program had a theme or it did not) did not allow Clarke and his colleagues to evaluate the relationship between specific themes and specific outcomes. It would be more than 15 years until researchers resolved this dilemma.

Inkelas et al.'s thematic typology of L/L programs. The first empirically-derived, thematic typology of L/L programs appeared in the final report accompanying the 2004 NSLLP (Inkelas et al., 2004). As part of that study, administrative contacts at the 34 participating colleges and universities in the NSLLP were asked to complete the Living-Learning Programs Survey (LLPS), a 23-item Web survey that gathered information about the administrative, organizational, and programmatic features of each L/L program that the institutions offered. Inkelas et al. (2004) surveyed a total of 297 programs, and received in response 268 (90.2%) completed instruments.

To develop their thematic typology, Inkelas et al. (2004) employed basic content analytic techniques. According to the authors, 50-word descriptions of programs' "theme, goals, and objectives" (p. 12) were read by a single rater and possible category names induced. For programs without descriptions, program titles were reviewed to determine whether sufficient information was available to definitively classify a program within an existing category. The authors coded a total of 254 programs in this manner, initially yielding 26 discrete types. After review, it was determined that those types could be collapsed into 14 primary categories, several that subsumed one or more thematic sub-types. As a result, the following primary categories emerged:

1. Disciplinary programs, including the sub-types of business, education, engineering and computer science, health science, humanities, general science, and social science programs.

2. Cultural programs, including the sub-types of international/global, language, and multicultural/diversity programs.
3. Transition programs, including the sub-types of new student transition and career/major exploration programs.
4. Honors programs.
5. Fine and creative arts programs.
6. Civic and social leadership programs, including the sub-types of civic engagement, leadership, and service-learning/social justice programs.
7. Women's programs, including the sub-types of women in leadership and women in math, science, and engineering programs.
8. Wellness/healthy living programs.
9. General academic programs.
10. Residential colleges.
11. Multi-disciplinary programs.
12. Upper-division programs.
13. Research programs.
14. Outdoor recreation programs.

Inkelas et al.'s (2004) typology is significant because it moves well beyond cursory notions of theme or focus mentioned by learning community or L/L program theorists (e.g., Lenning & Ebbers, 1999). By cataloging the diversity of purposes to which colleges and universities have applied L/L programs, Inkelas et al.'s work demonstrates, in part, the breadth of their expansion within the academy. The thematic typology is also of value because it does provide definitions of each secondary type, allowing new users a way to readily classify their particular programs of interest. New research that employs their typology as a grouping variable, for example, is made possible, opening the door for comparison of different programs sharing an identical theme, or “benchmarking” (p. IV-3).

Of course, Inkelas et al.’s (2004) typology has some limitations. First, constant practitioner innovation makes it impossible for any thematic typology to be completely inclusive of all program types. Therefore, the development of thematically novel programs will always spur the need for the typology to be revised. Second, Inkelas et al.’s effort was guided by a pragmatic need to quickly generate useful practitioner information, and as such did not employ the most robust set of content analytic techniques. Neuendorf (2002), for example, prescribes a method for analysis that not only promotes validity through an iterative process of consensual coding by multiple raters, but also seeks to demonstrate reliability through statistical analyses of the results of raters’ coding efforts. Inkelas et al.’s typology appears to have been developed by a single rater, making the computation of inter-rater reliability estimates moot. Although the researchers did allow institutions to review the categorization of their programs before the typology was originally published, there is no evidence that any feedback—positive or negative—was ever received (Inkelas et al., 2004).

The use of Inkelas et al.’s typology in L/L research. After detailing the development of their thematic typology, Inkelas et al. (2005) explored each of the 2004 NSLLP’s learning outcomes and mean differences in students’ scores based upon the thematic type of the program in which they participated. Using the 26 secondary types identified by their work as the independent variable, the authors computed one ANOVA for each outcome. The authors' portrayal of their statistical findings was muted, perhaps because given the large number of groups their results were not wholly surprising: significant differences were found by thematic type for each outcome at the $p \leq .001$ level.

That differences existed, Inkelas et al. (2005) noted, was not the most compelling story. What was notable, however, was that those differences that were found to have existed were what the reader might have logically expected. New student transition programs, for example, were the programs associated with the highest mean scores on smooth social transition, and the second highest mean scores on smooth academic transition. Honors programs were those most frequently associated with high mean scores on measures of students' enjoyment of challenging intellectual pursuits and satisfaction with courses, and second highest on the development of critical thinking abilities. Finally, likely reflecting the developmental progression of their participants, upper-division programs (those reserved for juniors and seniors) were associated with highest scores on measures of self-perceived growth in cognitive complexity, appreciation for liberal learning, and personal philosophy.

Before considering the implication of Inkelas et al.'s (2005) work, several cautions are of note. First, the authors' exploration of outcomes is cross-sectional in nature, and as such while it is possible to talk about an *association* between theme and outcomes, it cannot be said that theme *promoted* a given outcome. Second, although the 2004 NSLLP included proxy pre-tests for many of the outcomes considered by Inkelas et al., those pre-tests were not used to adjust students' outcome scores. As a result, the magnitude of the association between theme and outcomes is likely overstated. Finally, Inkelas et al.'s work does not consider the possible confounding influence of selection effects: that one or more unobserved variables simultaneously drives L/L participation and the outcome of interest, thereby confounding the results of their study.

Despite these shortcomings, Inkelas et al.'s (2005) findings have implications for practitioners looking to target their interventions to address certain student concerns or enhance particular student outcomes. It also contains an important methodological caveat for researchers interested in the efficacy of L/L programs: programs are working when they are promoting student outcomes consonant with their emphasis, but they are not failing to work when outcomes unrelated to their purpose are not being supported (Inkelas et al.). Although a measure of smooth social transition to college might be particularly appropriate to assess the efficacy of new student programs, for example, it might not be nearly as telling about the gains students are experiencing in a program focused on wellness. Inkelas et al.'s findings suggest that it is unreasonable to expect that all L/L programs "do" all things. As a result, it seems, evaluation of L/L programs should be conducted through lenses that might reasonably be expected to focus upon where these programs devote their energy.

In this light, Clarke et al.'s (1988) finding that, when compared to programs with no themes, thematically driven L/L programs promoted participation in study groups and increased the frequency of meaningful peer conversations is entirely logical: themes can provide an organizing principle around which student behavior and programmatic efforts can coalesce. Based upon the work of Inkelas et al. (2005), themes also appear to provide a foundation upon which meaningful student learning outcomes can be built.

Cognitive outcomes of L/L participation. Recent research on the influence of L/L participation on cognitive growth, by far the most voluminous area of study, has yielded mixed results. This may be due to the diversity of outcomes that could potentially fall under the broad umbrella of "cognitive gains." Below, I briefly introduce several studies that have focused on the cognitive outcomes of L/L participation. Most follow a predictable pattern: the use of analysis of variance (or a related technique) to establish whether differences exist between students in L/L programs and those in the TRH environment, followed by the use of block-entry linear regression to identify relationships between the dependent variable and several predictor variables, including but not limited to (a) student background characteristics, (b) academic and curricular environments, (c) campus or residential environments, (d) co-curricular environments, (e) direct peer interactions, and (f) thematic characteristics of L/L programs. As such, I begin with findings that relate to the main effect of L/L participation, and then turn to the specific influence of the independent variables listed above. Because the same works are often referenced throughout this (and later) sections, critique is offered the first time a study is mentioned in this review.

Main effects of L/L participation on GPA. Several researchers focused on cognitive gains have chosen to operationalize them as simple increases in students' GPAs. Recent research has suggested that at least some GPA benefit may accrue to L/L participants. Stassen's (2003) comparison of two cohorts (n=3948 and 3580) of L/L and non-L/L participants at the University of Massachusetts-Amherst predicted GPA gains for L/L participants ranging from 0.121 to 0.217 points, holding background characteristics constant, depending upon program of enrollment. Similarly, Pasque and

Murphy's (2005) study of students (*N*=2415) enrolled in seven different L/L programs at the University of Michigan found that L/L participants earned statistically significantly higher GPAs when compared to their TRH peers (0.23 units on a six-point, ordinal, GPA scale).

A notable strength of Stassen's work is that the source of her data on students' GPA came from an unbiased source: campus data systems. In contrast, Pasque and Murphy relied upon students' self-reported GPAs, which they note may have been less than wholly accurate. Neither address selection effects, and Pasque and Murphy provide little detail about the specific programs they explored in their study or students' pre-entry achievement. In Stassen's work, both are addressed. The programs in which L/L effects were most consistently demonstrated were either highly selective or were open on a "first-come, first-served" (p. 588) basis. To the extent that these motivational characteristics are associated both with selection in to the program and GPAs, the estimated L/L effect is biased. Stassen also noted significant pre-entry differences between L/L participants and those in TRH environments, further suggesting these groups were asymmetrical from the outset.

This GPA benefit may not accrue uniformly, however, if it accrues at all. In Edwards and McKelfresh's (2002) single-institution comparison of 342 first-year students, 81 (approximately 24%) of whom participated in a L/L program designed for natural science majors, the mean GPA for male L/L participants was .397 grade points higher than the mean for non-participants, a statistically significant difference. This trend reversed itself for women, however, with the authors noting that non-L/L women had significantly higher GPAs (by .016 units) than did the women in their sample who participated in an L/L program. Later work by Purdie (2007) also cast doubt on the relationship between L/L participation and GPA gains. Using four years of campus data, Purdie compared three interventions: L/Ls, classroom-based first-year experience programs, and residential freshman interest groups. Net of student background characteristics, only participation in residential freshman interest groups was associated with higher first-semester GPAs.

While their focus on a single institution complicates generalizability, Edwards and McKelfresh's (2002) and Purdie's (2007) work has notable strengths. First, like Stassen (2003), both used information on student outcomes came from administrative data. Edwards and McKelfresh's work is particularly significant because since approximately half of their L/L population was assigned there, rather than opting to participate, and as such selection bias may have been partially attenuated.

Main effects of L/L participation on other cognitive outcomes. Other recent research has employed more nuanced indicators of cognitive benefits, resulting in a dizzying array of findings. That they have rarely used the same indicators or expressed their findings in similar formats makes getting a true sense of L/L programs' potential cognitive benefits even more complicated. Nevertheless, to familiarize the reader with the scope of current L/L scholarship, below I introduce several studies and their findings regarding the main effects of L/L participation on what their authors have variously construed as "cognitive outcomes."

Pike (1999) explored 626 first-year students' gains in general education and intellectual development, as defined by student scores on CSEQ subscales, comparing L/L participants (25% of the sample) to their TRH peers. Although L/L participants reported statistically significantly greater gains in general education than their TRH peers (with a reported effect size of .27), no significant difference between the groups was found using Pike's measure of intellectual development. After controlling for student background characteristics and environmental measures, the observed difference in general education gains disappeared.

Pike carefully identifies the limitations of his work, including generalizability to other settings, dangers of inferring causality from cross-sectional studies, and selection bias. Unfortunately, his choice of a measured variable path analysis, rather than full-fledged structural equation modeling, does not allow us to gain insight in to the measurement characteristics of his scales, and the scale scores used in his analysis are contaminated with measurement bias.

In a single institution study, Inkelas and Weisman (2003) contrasted the experience of 833 L/L participants in three different types of programs—academic transition programs, honors programs, and discipline-based programs—to that of 1277 students living in TRH environments, looking for differences in these groups' enjoyment of (a) challenging intellectual pursuits and (b) learning new or different perspectives. The authors conducted two between-groups analyses, the first contrasting all L/L students against their TRH peers, and the second considering each residence arrangement separately. In the two group analysis, Inkelas and Weisman found that L/L students outperformed their TRH peers on both measures. When the authors disaggregated L/L students by type and compared their outcomes alongside students in the TRH, statistically significant between-group differences remained, with the outcome means of honors programs typically exceeding that of other types and the TRH environment.

Although the authors appropriate identify one potential measurement concern—the use of student self-report data—they do not address another: the possibility that their outcome scales performed poorly. The highest item-factor loading reported on the authors scale of challenging intellectual pursuits was .54, and, for the scale of learning new or different perspectives, .53. As a result, I infer that the *highest* loading items are approximately 75% measurement error (i.e., one minus the squared loading). Stronger scales—or methods like structural equation modeling that can attenuate measurement error—would increase confidence in the authors' findings.

Pasque and Murphy (2005) were interested in a construct they referred to as intellectual engagement, which encompassed such concepts as "[motivation] to learn new things" (p. 431) and "[looking] at all sides of a disagreement" (p. 431), as well as GPA (described above). The authors found that L/L participation was positively related to higher student scores on their measure, with students living in TRH environments scoring significantly lower than their L/L peers, holding background factors constant. Unfortunately, the authors provide little detail about the outcome scale itself. Although Pasque and Murphy indicate that it was identified after exploratory factor analysis, neither the constituent items nor their loadings are provided. While they do indicate that items with loadings below .35 were dropped from their analysis, that does little to characterize the true strength of their measure.

In a three-institution study, Inkelas, Johnson et al. (2006) compared 1744 L/L participants to 2314 of their peers living in TRH environments, examining students' self-reported gains in their cognitive complexity as well as their growth in appreciation of liberal learning (defined as an interest in the tenets of a traditional, liberal education).

Although the authors noted no mean differences in the two groups' self-reported growth in cognitive complexity, they interpreted their findings to suggest that students who participated in L/L programs experienced greater growth in liberal learning than their TRH peers, albeit to a small degree.

Because the authors followed-up their means analysis with regressions predicting student outcome scores, there is at least some evidence to suggest that Inkelas, Johnson et al.'s (2006) findings may have been muted by important covariates. For example, in two of the authors' three samples, OLS models predicting cognitive complexity revealed that academic class level—uncontrolled in the authors' tests of means—was strongly associated with the dependent variable. (Note that class level was related to liberal learning in only one of three samples the authors considered, and differences in liberal learning *were* statistically significant between the L/L and TRH groups.) Controlling for class level, or other important student-level variables, might have yielded different findings.

Finally, Kohl (2009) contrasted the development of critical thinking ability of three groups of NSLLP participants: those in the TRH environment, participants in honors L/L programs, and participants in L/L programs focused on civic engagement or social leadership. Using participation in an honors L/L program as his reference group, Kohl noted no statistically significant difference in predicted critical thinking scores for civic engagement/social leadership participants, but a significant, negative effect for residence in the TRH environment.

To be sure, there is some evidence to suggest that L/L participants may demonstrate statistically greater cognitive outcomes than their TRH peers, even when controlling for potentially confounding background characteristics. However, concerns still plague the studies from which that evidence comes, such as those raised by Inkelas and Weisman (2003) that most efforts represent the experience of small groups of students, in a small number of programs, at a small number of institutions. In combination with potentially suboptimal analytic methods, in particular OLS models which assume structural invariance between independent and dependent variables across groups, it seems safest to say the scholarship is far from conclusive.

Despite their shortcomings, however, these studies can represent a springboard for larger research programs that can produce findings that are more generalizable. One way in which they do so is by identifying for future researchers important constructs to consider in new research. Below, I summarize findings, drawn from the studies above, about the relationship between six classes of predictor variables and cognitive outcomes, including (a) student background characteristics, (b) academic and curricular environments, (c) co-curricular environments, (d) direct peer interactions, (e) campus and residential environments, and (f) thematic or structural characteristics of L/L programs.

Related student background characteristics. Studies specifically focused on L/L participants have suggested several student background characteristics may be positively related to their cognitive gains. The constructs researchers have considered most frequently include (a) gender, (b) race, (c) pre-college achievement, and (d) pre-college expectations.

The role of gender in the development of cognitive outcomes among L/L participants is mixed within the research literature. Pike (1999), for example, found that gender had no direct or total effect on first-year students’ general education or intellectual development gains. However, when cognitive outcomes have been operationalized as GPA gains (Stassen, 2003) or growth in liberal learning (Inkelas, Johnson et al., 2006), being female has been associated with better outcome scores. Women did not fare so well in Inkelas and Weisman's (2003) investigation of students' enjoyment of challenging intellectual pursuits, where, at least in the case of women in discipline-focused L/L programs, men reported significantly higher outcome scores. It may be that some outcomes take time to meaningfully emerge and are not measurable within first-year populations, thereby complicating the identification of difference (Pike), or that there are gender-related differences in outcome attainment in certain thematic types of L/L programs (Inkelas & Weisman, 2003; Inkelas, Johnson et al. 2006).

Researchers have also found inconsistent relationships between race and L/L participants' attainment of cognitive outcomes. Like gender before it, Pike (1999) found no direct or total effects of racial minority status on first-year students' general education and intellectual development gains. Similarly, Stassen (2003) found race to be largely unrelated to students' cumulative GPAs in their first year. Other studies have found race to be significantly related to their cognitive outcomes of interest, including a negative relationship between being an Asian American student in a honors program and the enjoyment of challenging intellectual pursuits (Inkelas & Weisman, 2003), and positive associations between being Latino in a discipline-focused program and the enjoyment of learning new perspectives (Inkelas & Weisman, 2003) and being a racial minority and reporting growth in liberal learning (Inkelas, Johnson et al., 2006). Again, as noted above, whether these findings are indicative of "true" relationships between constructs, evidence of differentially powerful educational environments, or simply phenomenon of measurement remains unclear, motivating further research.

Pre-college measures of cognitive ability have tended to demonstrate more consistently positive relationships with collegiate cognitive outcomes. Pasque and Murphy (2005), for example, noted a strong positive relationship between high school GPA and students' academic achievement and engagement. Similarly, Inkelas and Weisman (2003) found that, for at least some L/L students, SAT scores were positively related to the enjoyment of challenging intellectual pursuits and to the enjoyment of learning new perspectives, and Inkelas, Johnson et al. (2006) noted positive relationships between SAT scores and students' growth in appreciation of liberal learning. Only Pike's (1999) work fails to corroborate this apparent trend, noting no relationship between students' ACT scores and students' general education and intellectual development gains. Pike's finding may be driven by his use of a first-year only sample, or, in the light of recent concerns about the predictive validity of ACT composite scores (Bettinger, Evans, & Pope, 2011), the genuine absence of a relationship.

Finally, measures of students' pre-college expectations about the cognitive growth they would experience show some connection to the outcomes they eventually attain. Inkelas and Weisman (2003), for example, found that pre-college inclination to learn about new ideas was positively related to honors L/L and discipline-based L/L participants' subsequent enjoyment of challenging intellectual pursuits and, not surprisingly, all L/L participants' collegiate enjoyment of new perspectives. In a later study, Inkelas, Johnson et al. (2006) found that students' pre-college rating of the importance of intellectually-focused dialog with their peers was positively related to students' later reports of cognitive complexity and growth in liberal learning. Importantly, in both studies student expectations were measured retrospectively, and prospective measurement might have yielded different results.

Related academic and curricular environments. Academic and curricular environments have demonstrated a mixed relationship with L/L students' ratings of their cognitive growth. Students' class standing, a variable that might reasonably be included in any cognitive study as a proxy for developmental change, appears less frequently in L/L research because many authors have focused on solely first-year students. One study that included students of all class levels noted that academic class level demonstrated a generally positive relationship to students' self-reported cognitive complexity and growth in liberal learning, consistent with a developmental interpretation of this outcome (Inkelas, Johnson et al., 2006). Other variables that have been more frequently considered include major field of study and faculty/student interaction.

Students' choice of major has been inconsistently related to various measures of their cognitive gains. Among studies that relied upon L/L-only samples, no relationships were found between major and enjoyment of learning about new perspectives (Inkelas & Weisman, 2003) or developing cognitive complexity (Inkelas, Johnson et al., 2006). However, a positive relationship between being a liberal arts major participating in transition-focused L/L programs and greater enjoyment of challenging intellectual pursuits, at least when compared to their undecided peers, has been documented (Inkelas & Weisman, 2003). In a design that mixed L/L and non-L/L students, Stassen (2003) found that enrollment in engineering, relative to being undeclared, was negatively related to GPA. Net of measurement problems, these findings may be due genuine disciplinary differences, unobserved variables that are related to both choice of major and student outcomes (e.g., motivation), or a dose effect (e.g., first or second year students, although they may have declared a major, may not have had sufficient exposure to a discipline to have been affected by it).

The broader college impact literature has attributed positive effects to student-faculty interaction (Pascarella & Terenzini, 1991, 2005), a finding generally mirrored in the L/L research focused on cognitive outcomes. Pike (1999), for example, found that interaction with faculty was positively related to students' subsequent gains in intellectual development. Similarly, Inkelas and Weisman (2003) noted that across a number (although not all) of L/L programs of different types, course-related faculty interaction was statistically significantly associated with students' enjoyment of challenging intellectual pursuits.

Related co-curricular environments. The generally accepted relationship between co-curricular involvement and students' cognitive growth (Pascarella & Terenzini, 1991, 2005) has been inconsistently evidenced in L/L research. Pike (1999) identified a linkage between involvement in arts, music, and theater, and general education gains, driven perhaps by a shared notion of liberal education. Inkelas, Johnson, et al. (2006) found a positive relationship between participation in Greek-letter organizations and students' growth in liberal learning in one of the three campuses participating in their research, and Inkelas and Weisman (2003) noted a positive relationship between community service participation and the enjoyment of diverse perspectives. However, other studies have noted little connection between participation in student organizations and the cognitive gains of L/L participants, including their general education or intellectual development gains (Pike), their enjoyment of challenging intellectual pursuits (Inkelas & Weisman, 2003), or their cognitive complexity (Inkelas, Johnson et al., 2006).

That findings are not more consistently supportive of co-curricular involvement's positive relationship with gains of this type is surprising, given theorists' and proponents' contention that it is the integration of co-curricular experience with the curriculum that, in part, contributes to L/L programs' efficacy (Inkelas & Weisman, 2003; Lenning & Ebbers, 1999; Schoem, 2004; Shapiro & Levine, 1999). However, it may well be that existing measures do a poor job of capturing either this integration or outcomes of interest, the quality of this integration is uneven, students fail to accurately report their efforts, or measures of frequency do not take in to account effort or quality. It may also be that co-curricular involvement has no causal impact on student outcomes, because one or more unobserved factors simultaneously compel students to engage in co-curricular activities and influence learning and development.

Related direct peer interaction. Prior research literature has suggested that peer interaction can be related to students' attainment of cognitive outcomes (Astin, 1993; Pascarella & Terenzini, 1991, 2005). Pike (1999) found that conversations with peers about "significant issues" were related to students' general education and intellectual development scores, and that L/L participants reported more significant peer conversations than their TRH peers. Similarly, what Inkelas and Weisman (2003) categorized as socio-cultural conversations (those that involved possibly contentious social or political issues, or issues related to multiculturalism and human difference) with peers were positively related to most, although not all, L/L students' enjoyment of challenging intellectual pursuits and learning about new or different perspectives. This same type of conversation has also been linked to scales assessing growth in liberal learning (Inkelas, Johnson et al., 2006). Because these findings are so intuitive on face, remembering they suggest association and not causation seems important: whether conversations drive growth or growth enables conversation (or whether something unmeasured promotes both) is not discernable from the extant literature.

The evidence concerning the relationship between academically-focused conversations and cognitive outcomes is less strong. Although Pike (1999) found that conversations with peers that integrated information learned in class bore a positive relationship to general education and intellectual development measures, and that L/L participants were more likely than non-participants to engage in such integration, neither Inkelas and Weisman's (2003) nor Inkelas, Johnson et al.'s (2006) studies detected a relationship between peer conversations focused on academic or vocational issues and students' self-reported cognitive growth. Given their somewhat counterintuitive findings, the extent to which the Inkelas-related findings are a function of measurement is worthy of consideration. Although Inkelas, Johnson et al. did not publish their measurement scale, Inkelas and Weisman does provide this detail, noting it included discussions of class material, "discussions with students whose personal values were very different from [their] own," (p. 361), and concerns about classes and assignments. The second item in particular lacks some degree of face validity, potentially confounding the authors' findings.

Related residential environments. Consistent with general theories of person-environment fit (Walsh, 1978), some L/L researchers have explored whether students' residential environments (or their perceptions of them) have a significant relationship with their cognitive outcomes. To date, research has yet to strongly link residential environments and cognitive outcomes. Inkelas and Weisman (2003), for example, found no relationship between L/L participants' perceptions of their residence halls' level of social or academic support and their enjoyment of challenging intellectual pursuits, and a relationship between students' perceptions of their hall as being socially or academically supportive and their enjoyment of learning from diverse perspectives was noted in only one of the three types of programs they investigated (i.e., transition programs). Similarly, Inkelas, Johnson et al.'s (2006) multi-institutional study found only limited evidence for the connection between perceived hall social support and cognitive complexity, and between hall academic or social support and growth in liberal learning.

These findings may suggest that it is not solely students' perception of aggregate environments, but rather (or perhaps also) their level of engagement with it (n.b., Pace, 1984). Indeed, relationships between above concerning sociocultural peer conversations—something students actually do, as opposed to environments which they merely perceive—and a subset of cognitive outcomes could be seen to lend additional credence to this argument. In that vein, Pike (1999) noted that L/L participants reported greater levels of residence hall involvement than their TRH peers, and that residence hall involvement was positively related to students' gain in intellectual development. No relationship was found, however, to students' general education gains.

Related thematic differences among L/L programs. Inkelas and Weisman's (2003) approach to evaluating the effects of thematic type on cognitive outcomes suggested they believed type had both direct and indirect effects on student outcomes.

Their three-part analysis evaluated theme-related differences in (a) students' ratings of eight "key living-learning environments and perceptions" (p. 355), (b) students' rating of transition (described elsewhere) and cognitive outcome measures, and (c) the relationships between "key" environments and student outcomes. In each analysis, the authors compared students living in the TRH environment to peers in one of three types of L/L programs: (a) honors programs, (b) disciplinary programs, and (c) transition programs.

Initially, Inkelas and Weisman (2003) used omnibus analysis of variance tests, followed by post hoc analyses, to identify differences between their four groups of students. Irrespective of program type, no statistically significant differences were found in L/L participants' ratings of key living-learning environments, suggesting that students experienced those environments similarly. Theme-related differences were found among L/L participants' outcome scores, however, with honors programs typically demonstrating statistically significantly higher means than other types of L/Ls. Minimally, then, one might infer program type has a direct effect on student outcomes. Of course, whether an analysis that controlled for student entry characteristics would have attenuated the effect of honors participation—or might have revealed true between-group differences in students' responses to the authors' environmental measures—is unknown.

Next, Inkelas and Weisman (2003) sought to determine whether the relationships between key living-learning environments and student outcomes were similar across types. If similarity was found, one might infer that type did not differentially influence those environments and, by extension, students' outcomes. By constructing separate regression equations for each program type and each outcome, the authors could estimate the strength and significance of each environment-outcome relationship, holding student characteristics constant.

Inkelas and Weisman's (2003) regression analyses offered preliminary evidence that different environment-outcome relationships did exist between types of programs. For example, students' perceptions that their residence hall was socially supportive were significantly predictive of the academic transition of students in transition and honors programs, but not for students in disciplinary programs or in TRH environments. Similarly, discussion of socio-cultural issues was not significantly related to transition-L/L students' enjoyment of challenging intellectual pursuits, but was for all other students in the study. Thematic type, then, may affect student outcomes in two ways: (a) by its direct effect on student outcomes, as shown by the authors' ANOVA tests, and (b) by its moderation of the strength of the relationships between outcomes and key living and learning environments.

Transition and persistence outcomes of L/L participation. Several recent studies have focused on the transition and persistence outcomes associated with L/L participation. Those presented in more detail below represent a wide range of conceptual frameworks and analytical techniques, and contribute uniquely to the further study of the educational benefits of L/L participation through the methods they employed. I begin below by outlining the main effects of L/L participation associated with students' academic transition to college and first-year student persistence. Then, I present researchers' findings about the relationship between transition, persistence, and six classes of independent variables, including those representing: (a) student background characteristics, (b) academic and curricular environments, (c) campus and residential environments, (d) co-curricular environments, (e) direct peer interactions, and (f) program's thematic type.

The main effects of L/L participation on ease of academic transition. Three studies, all based on either the 2004 or 2007 NSLLP or its antecedent pilots, have demonstrated that L/L participants report higher mean scores on the RES-B's measure of smooth academic transition than their TRH peers. Early work by Inkelas and Weisman's (2003) using NSLLP pilot data noted just such an effect when L/L participants were explored in omnibus fashion, but concluded it may have been due to outcomes reported by students participating in a transition-focused L/L program (notably, early evidence of the importance of theme).

National results reported in the 2004 and 2007 NSLLP data collections also suggested higher mean scores on smooth academic transition for L/L participants than non-participants (Inkelas et al., 2004; Inkelas, Brower et al., 2008). Although effect sizes are not reported, the magnitude of difference appears small: the absolute mean differences the noted in 2004 and 2007 were .10 and .06, respectively, on a six point scale. Were those differences to have been calculated while controlling for other potentially confounding variables, the effect might have been smaller yet.

Later work by Inkelas, Daver, Vogt, and Brown Leonard (2007) focused on subset of college attendees, first-generation students. Using 2004 NSLLP data, the authors noted that first-generation students who participated in L/L programs reported a small (η^2= .03) but detectably smoother transition than their TRH peers, controlling for students' pre-college confidence in handling college-level work. Regression analyses identified several student, curricular, and residential characteristics related to academic transition, discussed below.

The main effects of L/L participation on persistence. While the literature reviewed above consistently suggests that L/L participants report smoother academic transitions than their TRH peers, findings vis-à-vis first-year persistence are mixed. In an early work, Pike, Schroeder, and Berry (1997) found no statistically significant difference in L/L participants' likelihood of persistence and that of their non-participating peers, net of student background characteristics, measures of integration, institutional commitment, and academic achievement. However, L/L membership was associated with statistically significantly greater rates of faculty-student interaction, institutional commitment, and social integration (controlling for the influence of other variables), which were related to students' persistence decisions.

In contrast, both Edwards and McKelfresh (2003) and Stassen (2003) suggested L/L participation was positively related to persistence, at least for some student populations. In a study focused on natural science students, Edwards and McKelfresh detected statistically significant differences in the predicted probability of persistence for students identified as racial minorities, who were predicted to have an institutional persistence rate of 76% within a TRH environment but an 89% persistence rate if they participated in a L/L program. Stassen concluded that, after holding several student background characteristics constant and irrespective of major, L/L participants were between 33% and 60% less likely to withdraw than non-participants.

Although each of the studies in this set controlled for student background characteristics in some regard—typically gender, racial/ethnic minority status, and a measure of pre-college ability—only Pike et al. (1997) controlled for facets of the college environment beyond L/L participation. This may explain why Pike's findings vis-à-vis the benefits of L/L participation are muted in comparison to those of both Edwards and McKelfresh (2002) and Stassen (2003). None of the authors attempted to control for characteristics that might have driven students' participation in an L/L program. To the extent that those characteristics would have also been positively related to L/L participation, I would expect the L/L effect on persistence is actually less than that described above.

Researchers have frequently implicated student background characteristics in frameworks that seek to evaluate L/L participants' transition and persistence outcomes. Five variables are commonly considered, including (a) gender, (b) race, (c) measures of entering ability, (d) parental resources, and (e) pre-college attitudes.

Conventional wisdom notwithstanding (e.g., Conger & Long, 2010), L/L research has yet to demonstrate a relationship between gender and persistence or transition. Although Pike et al. (1997) identified a direct negative relationship between being female and persistence, that relationship was wholly offset by indirect positive relationships through other facets of their model (i.e., social integration and institutional commitment). Later work found no direct relationship between gender and persistence (Stassen, 2003) or smooth academic transition whatsoever (Inkelas & Weisman, 2003; Inkelas et al., 2007), net of other characteristics in the authors' models.

Scholarship vis-à-vis one's status as a member of a racial/ethnic minority group is more consistent. Among L/L participants, Pike et al. (1997) concluded racial/ethnic minority status had a statistically significant negative effect on persistence. Stassen's

(2003) later work is generally consistent with that finding, but her use of separate race categories revealed additional nuance: identifying as African American was significantly related to lower odds of departure after students' first year, while identifying as Asian American was significantly related to higher odds (both compared to White students). Unfortunately, because of Stassen's use of a single OLS model, I cannot determine whether her finding would have held across L/L and TRH groups, or whether it would have mirrored Pike et al.'s work.

There is limited evidence to suggest a relationship also exists between race and L/L students' academic transition: in one of three programs studied, Inkelas and Weisman (2003) noted a positive relationship between smooth academic transition and identification as African American or Latino. Inkelas et al.'s (2007) later work on first-generation students' academic transition found no detectable relationship between race and academic transition, although they noted that, for L/L participants, generation status in the United States was significantly positively related to that outcome. Given that no study reviewed here explored potential interactions between generational status, race/ethnicity, and transition/persistence, Inkelas et al.'s work hints at a potentially fruitful line of research.

Perhaps not surprisingly, the extant literature suggests a student's pre-entry ability and/or aptitude is positively related to their subsequent persistence outcomes. Pike et al. (1997) found a small yet still significant net positive relationship between ACT composite and persistence, driven by an indirect relationship through academic achievement. Later, Stassen (2003) concluded both high school GPA and SAT math scores were related to decreased odds of voluntary withdrawal.

These constructs' relationship to students' perceived academic transition is much less clear. Inkelas and Weisman (2003) found no relationship between any of the measures considered above and L/L students' academic transition to college, and Inkelas et al.'s (2007) later work that incorporated both L/L and TRH students found no such relationship for either group. At least four tenable explanations for this finding exist: there is no "true" relationship, students are poor judges of their academic transition, students experiencing poor transition are likely to be study non-respondents, or those with the worst transition had dropped out before being sampled.

The influence of parents and parental characteristics on persistence and outcomes is mixed in the L/L literature. Among L/L participants, parental encouragement has been found to be positively related to persistence (Pike et al., 1997). Parental education has exhibited a similar effect for L/L students participating in honors programs (Inkelas & Weisman, 2003). Although parental encouragement and parental education appear conceptually distinct, both may converge around the valuing of traditional notions of academic achievement, a marker of a particular form of cultural capital. Net of other characteristics, financial capital may be less related to transition outcomes. Inkelas et al. (2007) noted that, for at least the first-generation students who were the focus of their study, no relationship existed between parental income, financial aid packaging, and students' academic transition.

Finally, the work of Inkelas and Weisman (2003) and Inkelas et al. (2007) provided empirical evidence of the association between students' confidence in their ability to transition to college and subsequent outcomes. Net of a series of other student background characteristics and environmental measures, the authors found a strong positive association between pre-college confidence and post-entry transition scores in L/L populations. While Inkelas and Weisman's work did not include a TRH comparison group, Inkelas et al.'s did. Contrary to what might be expected given Astin's (1991) IEO model, they detected no relationship between the pre-college expectations of TRH participants and their subsequent ratings of academic transition. It may be that something in the TRH environment inhibits students' realization of their pre-entry expectations, or that something in the L/L environment makes that realization more likely. Alternatively, it may be that L/L participants responded in socially-desirably ways, either on the pre-test or outcome measure.

Related academic and curricular environments. Several academic and curricular environments have been explored in studies of L/L participants' college transition and persistence, including (a) major field of study, (b) academic achievement, and (c) formal and informal academic integration, including faculty-student interactions.

Recent studies have suggested no relationship between L/L participants' major field of study and their academic transition or first-year persistence. Inkelas and Weisman (2003), distinguishing between science or mathematics majors, liberal arts majors, professional or technical majors, and undecided students, found no relationship between field of study and smooth collegiate transition. Stassen's (2003) investigation of persistence also failed to identify a statistically significant relationship with major, which she subdivided into humanities and social sciences, natural sciences, math, and engineering, applied professions, and "pre" majors (e.g., pre-medicine). While field of study may signal meaningful underlying differences in student preferences that might be thought to be relevant to the study of student outcomes (n.b., Walsh, 1978, and his discussion of Holland’s work), it may be that the predominance of first-year students in L/L studies results mutes any true effect due to misclassification (e.g., a student was majoring in math at the time of the study but later realized he was better suited to education) or a lack of time to be socialized within a discipline.

Finally, studies of both persistence and transition have considered measures of academic integration, driven by both formal and informal engagement with institutions academic systems, generally operationalized as academic performance and faculty/staff interaction, respectively. Within the context of L/L programs, Pike et al. (1997) considered the role of academic achievement in L/L participants’ first-year persistence, a feature not considered by Stassen (2003) in her similar investigation of the same topic. Not surprisingly, Pike and his colleagues found that achievement, as measured by cumulative GPA at the end of a student’s first year, was strongly positively related to persistence.

L/L research has also generally supported the notion that student-faculty contact can lead to outcomes that might promote persistence. Interactions around courses and course content seem particularly valuable: both Inkelas and Weisman (2003) and Inkelas et al. (2007) found strong positive associations between course-related faculty interaction and students’ smooth academic transition. Informal faculty interaction, however, yielded no discernable relationships (Inkelas and Weisman, 2003). Before concluding that non-course-related faculty interaction is of no value, however, a caution is appropriate: the mean score on this measure was 5.8 on a scale ranging from 4 to 16. It may well be that the relative infrequency of informal student-faculty interaction (e.g., engaging in social settings or discussing personal problems) precluded a beneficial effect, if any, from being detected.

Finally, because most studies reviewed above are focused on persistence through a L/L participant's first year, academic class level is rarely included in persistence-focused studies. Both Inkelas and Weisman (2003) and Inkelas et al. (2007), however, have considered this variable in their analyses of students' ratings of their academic transition to college. Among participants in L/L programs focused on academic transition, Inkelas & Weisman noted a significant negative relationship to class year. This is unsurprising, given its interpretation: sophomores, juniors, or seniors would only likely be in a transition program if they were experiencing a problem in their adjustment to college, something that would seem to predispose them to poor scores on a measure of smooth academic transition. In their analysis of first-generation students in all types of L/L programs and in the TRH environment, Inkelas et al. (2007) noted no relationship between class level and perceived academic transition.

Related campus and residential environments. L/L participants' perceptions of their campus environments may influence their subsequent persistence and transition outcomes. Pike et al. (1997), for example, found that institutional commitment (as defined by their institutional satisfaction and their beliefs about whether they should remain at the institution) had a direct, positive relationship to L/L participants' first year persistence decisions.

The relationship between students' perception of their residence hall as academically and socially supportive and the smoothness of their academic transition remains unclear. The work of Inkelas and Weisman (2003) generally supports the contention. They found that an academically supportive residence hall environment was positively related to the academic transition of students in most types of L/L programs, as well as students in the TRH environment. Similarly, at least for students in transition and honors L/L programs, the authors noted a significantly significant relationship between perceptions that one's residence hall was socially supportive and ratings of academic transition. For first generation students, however, Inkelas et al. (2007) detected no relationship between residence hall climate and smoothness of academic transition, irrespective of L/L participation. The extent to which this latter finding is a function of the population under study is unclear and can only be addressed by further research.

Related co-curricular environments. Contrary to existing research that posits a generally beneficial relationship between persistence and co-curricular engagement or other markers of social integration (see, for example, Tinto, 1993), research among L/L participants suggests a neutral to negative relationship. Pike et al. (1997), for example, found no net significant relationship between social integration and persistence, as social integration's positive indirect effect on institutional commitment but negative indirect effect on academic achievement offset one another. Similarly, Inkelas and Weisman (2003) found no statistically significant relationship between involvement in four different co-curricular environments—student clubs, Greek-letter organizations, community service, or on or off campus work—and students' smooth academic transition. And, among the first-generation students in their study, Inkelas et al. (2007) noted a *negative* relationship between TRH students' community service participation and ease of academic transition.

One interpretation of these findings is that Tinto's (1993) conceptualization of an institution's academic and social systems as separate is, indeed, the case: co-curricular engagement might smooth social transitions, but it does little to ease academic ones. Alternatively, it may be the case that because the works cited above did not distinguish between engagement that was related to L/L participation and that which was unrelated, it failed to identify benefits associated with engagement that—in theory—seeks to blend students' academic and social worlds.

Related direct peer interaction. The general literature on college impact (Pascarella & Terenzini, 1991, 2005) has suggested that direct peer interaction can have a positive relationship to students' academic transition to college, so long as that interaction is consistent with academic success. Inkelas and Weisman (2003) found that academic conversations with peers were positively related to a smoother academic transition for students in honors L/L programs and in the TRH environment, with the former reporting more frequent conversations about academic issues outside of class than the latter. Peer conversations about socio-cultural issues, however, were found to be either negative (for honors L/L participants) or non-significant (for all other students). Later work by Inkelas et al. (2007) focused on first-year, first-generation college students found no relationship between either type of peer conversation and academic transition, irrespective of L/L participation.

Related thematic differences among L/L programs. As was the case in their analysis of cognitive outcomes, Inkelas and Weisman (2003) examined the effects of thematically different types of L/L programs on students' academic transition outcomes through a two-step process. After concluding there were main effects for L/L participation and thematic type on their measure smooth of academic transition, the authors sought to explore relationships between student background characteristics, key living and learning environments and academic transition scores, moderated by thematic type. Using a series of type-wise regressions, the authors identified significant variation in the relationship of key student characteristics or living and learning environments to academic transition across program themes, mirroring their findings vis-à-vis cognitive outcomes. Although the product of a single-institution study, this suggested thematic type might act to moderate any number of environment-outcome relationships, and that further type-wise exploration of student outcomes would be indicated.

Acknowledging that type might exert an effect on student outcomes, the technical reports accompanying the 2004 and 2007 NSLLP data collections explored theme-wise differences in students' smooth academic transition, to somewhat contradictory ends (Inkelas et al., 2004; Inkelas, Brower et al., 2008). In 2004, a statistically significant difference in mean transition score by thematic type of L/L programs was noted. However, in the 2007 collection, no such difference was detected. (Indeed, in 2007, smooth academic transition was the *only* outcome measure, excluding a set of ancillary measures related to alcohol use, where no type × outcome interaction was observed.) In the light of the sensitivity of the test due to the large number of comparison groups (i.e., types) and sample size, this latter finding is particularly surprising. Whether this contradiction is a function of measurement error in 2004 or 2007 or true changes in either students or programs cannot be readily ascertained.

The psychosocial outcomes of L/L participation. Compared to the cognitive or transition/persistence outcomes, relatively little has been published regarding the relationship between psychosocial outcomes and L/L participation. In fact, only one recent study appears in a peer-reviewed journal: Rowan-Kenyon et al.'s (2008) investigation of civic engagement. A review of recent doctoral dissertations and theses, however, yielded several additional studies for review. Due to the variability of outcomes within this class, only those studies that report on constructs somewhat analogous to those assessable in the present research are reviewed here.

Citizenship outcomes. Rowan-Kenyon et al. (2008) studied the citizenship outcomes associated with residence in one of three environments: (a) a L/L program focused on civic engagement, (b) a L/L program not focused on civic engagement, and (c) the TRH environment. A total of 1474 students were sampled from the 2004 NSLLP dataset, including all respondents from civic engagement themed programs and similarly-sized random samples of L/L students not in a civic engagement themed program and TRH students.

Omnibus ANOVA indicated a statistically significant difference existed between the three groups' means on the NSLLP's civic engagement measure (Rowan-Kenyon et al., 2008). Subsequent post-hoc analyses revealed that the omnibus test's significance was associated with a higher mean score of the civic engagement program group, when compared in a pair-wise fashion to the other two. Due to the possibly confounding influence of students' pre-college expectations about engagement in their collegiate community, as operationalized by a measure representing their rating of co-curricular involvement as important, this variable was included in a follow-up ANCOVA analysis. Although the main effect of program type remained, whether this approach sufficiently addressed selection bias on outcome scores remains an open question.

Rowan-Kenyon et al.'s (2008) conceptual framework posited that the development of civic engagement took place in two phases. First, student background characteristics (i.e., gender, race, high school ability, and socioeconomic status), pre-college ratings of the importance of involvement and academic and psychosocial confidence, residence arrangement (i.e., civic engagement L/L, non-civic engagement L/L, or TRH), co-curricular involvement, and interpersonal experiences drove three intermediate outcomes: growth in critical thinking ability, personal philosophy, and interpersonal self-confidence. Growth on these three factors, in turn, was posited to contributed to civic engagement.

Rowan-Kenyon et al. (2008) found that gender was the only predictive background variable, evidencing a significant positive relationship between being female and civic engagement. Each pre-test was statistically significant, with the importance of involvement and confidence in psychosocial skills evidencing a positive association with civic engagement, and confidence in academic skills evidencing a negative one. Although type of residence arrangement was not statistically significant in Rowan-Kenyon et al.'s regression analyses, several variables representing distinct co-curricular experiences and interpersonal experiences were positively related to the dependent variable, including (a) involvement in student government, (b) involvement in one-time community service, (c) involvement in on-going community service, (d) discussion of socio-cultural issues with peers, and (e) positive interactions with diverse others. Finally, as hypothesized, each intermediate outcome was related to students' scores on the NSLLP civic engagement measure.

At least two additional limitations to Rowan-Kenyon et al.'s (2008) work are of note. First, as the authors are quick to mention, the cross-sectional nature of NSLLP data prevents true causal statements. This seems to make the authors' interpretation of the relationship between the dependent variable and intermediate outcomes particularly problematic: it is impossible to say with any certainty whether civic engagement is driven by, or drives, critical thinking ability, personal philosophy, and interpersonal self-confidence. Second, the authors' analytic strategy not only forces us to assume that the relationships between predictors in the model and the dependent variable are invariant across each residential environment, but also obscures indirect effects of residential environment on other variables in the model.

Diversity appreciation outcomes. The 2004 NSLLP appears to represent the first time diversity appreciation was explored as an outcome of L/L participation (Inkelas et al., 2004). The NSLLP technical report accompanying the 2004 collection failed to detect a difference in mean diversity appreciation scores between L/L participants and their TRH peers. As noted earlier, no effort was made in the technical report to adjust mean differences on the basis of potentially important covariates; more complex analyses might have yielded wholly different findings. Within L/L programs, however, a statistically significant effect was found for type. Unfortunately, because no post hoc tests were conducted, the specific source of variation cannot be determined.

Using the same data set, Longerbeam (2005) explored the relationship between L/L program participation and students' openness to diversity, focusing on student background characteristics and behaviors, campus and academic environments, and program characteristics. Among the background and pre-test variables Longerbeam included in her model, being female and rating diversity activities as important exhibited significant, positive relationships to openness to diversity, while students' scores on the SAT or ACT exhibited significant, negative relationships. Positive campus diversity climates and the perception that residence halls had supportive academic and social climates were found to be positively related to students' openness to diversity.

Two markers of students' academic environments were found to have statistically significant relationships with the dependent variable: major and class level (Longerbeam, 2005). Students majoring in soft/applied fields, those practice-focused disciplines which are typified by multiple paradigmatic perspectives (Biglan, 1973a, 1973b), evidenced greater openness to diversity than their peers in hard/pure fields, those disciplines which are typified by a dominant paradigm and use them to drive further inquiry (Biglan). Consistent with a developmental interpretation of the construct, Longerbeam also noted that first and second year students, as opposed to seniors, evidenced lower levels of openness to diversity.

Finally, three indicators of L/L participants' interactions with others on campus bore significant, positive relationships to openness to diversity, including faculty mentorship, students' involvement in cross-cultural clubs, and their discussion of socio-cultural issues with their peers (Longerbeam, 2005). Although not assessed by Longerbeam, that openness to diversity (or an unobserved factor) influenced students' involvement in cross-cultural clubs and peer discussions of socio-cultural issues remains possible.

Relatively few variables representing L/L programs' structural characteristics were related to participants' openness to diversity (Longerbeam, 2005). Only the use of undergraduate peers as mentors and offering cultural outings as a program activity were found to be statistically significant predictors of openness to diversity, both with positive effects. Similar to the critique noted above, whether students who were appreciated diversity chose programs that included cultural outings or cultural outings influenced openness to diversity cannot be determined from her design. Finally, two of Longerbeam's three intermediate outcomes emerged as predictive of openness to diversity by the regression's final block. Both critical thinking and analysis abilities and sense of civic engagement evidenced a significant, positive relationship to the dependent variable. Again, because the intermediate outcomes included in Longerbeam's model were contemporaneously assessed with the dependent variable, directionality between these constructs is difficult to establish.

After the conclusion of Longerbeam's work, the 2007 NSLLP was fielded. In its technical report, Inkelas, Brower et al. (2008) again noted no main effect for L/L participation on diversity appreciation outcomes. Within L/L programs, however, statistically significant variation by program type was noted. Unfortunately, as was the case in 2004, mean scores were not adjusted by potentially important covariates, thereby making it possible any true main effect of participation was obscured. Similarly, no post hoc tests were conducted within L/L programs to determine the cause of the significant main effect of thematic type.

Outcomes related to Chickering's vectors. Leinwall's (2006) dissertation research also sought to assess psychosocial outcomes associated with L/L participation. Rather than focusing on a single aspect of development like the other studies reviewed here, Leinwall examined multiple facets of psychosocial growth using the Student

Developmental Task and Lifestyle Assessment (SDTLA), a measure of development along three of Chickering's (1969; Chickering & Reisser, 1993) seven vectors: developing autonomy, developing mature interpersonal relationships, and establishing purpose. Three questions were explored, including whether differences on measures of psychosocial development existed between L/L participants and their TRH peers, whether L/L participants exhibited gender-related developmental differences, and whether developmental differences existed between students in different types of L/L programs.

To answer her research questions, Leinwall (2006) sampled 292 sophomore L/L participants drawn from 12 structurally similar, but thematically varied, L/L programs at the University of Maryland, all of which targeted academically talented students. A comparison sample of 200 sophomore students was constructed by randomly selecting non-participating resident students who met the GPA and SAT/ACT criteria established for admission to the L/L program being studied. Slightly more than 40% of L/L participants and slightly more than 50% of TRH students completed the SDTLA, a sufficient number to evaluate Leinwall's first two research questions but not enough to consider the differential influence of thematic type.

Leinwall (2006) then computed 2 (gender) × 2 (residential environment) ANOVAs to identify group mean differences on each of her three outcomes of interest. Leinwall found no significant main effect for gender on either developing autonomy, developing mature interpersonal relationships, or establishing purpose. A significant main effect was found for residence arrangement on one outcome with L/L participants reporting higher levels of autonomy than their TRH peers. Tests of gender × residence arrangement interactions yielded no significant findings. Interpretation of Leinwall's findings is complicated by the lack of statistical controls evidenced in her data analysis plan. Although her comparison group was randomly selected from among students with academic credentials similar to her L/L participants, nothing was done (e.g., the use of ANCOVA or MANCOVA models) to attenuate the potentially confounding effect of students' background characteristics on their SDTLA scores.

Self-efficacy outcomes. Kamin's (2009) master's thesis explored the relationship between sophomore students' participation in an L/L and their academic self-efficacy, or their belief that they could do well academically in college. Using NSLLP data collected in 2007, she found that L/L participants reported higher mean efficacy scores than their TRH peers, with a medium effect size. Later multivariate analysis explored relationships between student background characteristics, a pre-test measure of efficacy, and key living and learning environments. Notably, Kamin's work estimated two separate OLS models, one for L/L participants and one for their non-participating peers. As a result, this allowed the author to test whether the strength of relationships between independent variables and the dependent variable varied between groups.

Net of background characteristics, two key living and learning environments—academic/vocational peer interactions and hall social climate—were positively related to self-efficacy (Kamin, 2009). No statistically significant differences in the unstandardized coefficients associated with these predictors and self-efficacy were noted between the L/L and TRH groups. The strength of Kamin's two-group analysis is, at least in part, offset by a limitation that was seemingly impossible to avoid: because she focused on sophomore students' academic self-efficacy, her sample did not include TRH or L/L participants whose self-efficacy was so low in their first year that they did not persist in postsecondary education, thereby being impossible to sample.

Chapter Conclusion

On balance, there is evidence to suggest that participation in L/L programs can be beneficial for undergraduate students. Existing research has pointed to gains in academic performance and development (Inkelas, Johnson, et al., 2006; Inkelas & Weisman, 2003; Pasque & Murphy, 2005; Pike, 1999; Stassen, 2003), transition and persistence (Edwards & McKelfresh, 2002; Inkelas &Weisman, 2003; Pike, Schroeder, & Berry, 1997; Stassen, 2003), and psychosocial growth (Leinwall, 2006; Longerbeam, 2005; Rowan-Kenyon et al., 2008) for L/L participants when compared to their peers in TRH environments. However, in addition to my individual critiques of these studies' methods, other methodological concerns exist, including those related to the assessment of L/L programs' indirect effects (Pascarella & Terenzini, 1980) and the failure to consider the relationship between programs' unique thematic type and the outcomes they are likely to attain (Inkelas et al., 2005). Researchers do know *something* about L/L programs' efficacy, it seems, but there is more to learn – and better ways to do so.

Methodological concerns about the research from which the reputation of these programs has stemmed notwithstanding, one thing seems clear: L/L programs are enduring features of campus communities upon which institutions have pinned at least some of their hopes for strengthening undergraduate education (Jones et al., 2006; Soldner & Szelényi, 2008). Concerns may exist, however, with why—and how—decision-makers have done so. It may well be, as Jones et al. have suggested, that institutions have been prompted to develop L/L programs because the conventional wisdom has painted them as cure-alls for what ails higher education in the United States. Or, as outlined in Chapter 1, institutions may also be implementing L/L programs for the sake of garnering status, prestige, and legitimacy in a higher education environment already fraught with competition (Ehrenberg, 2003; Monks & Ehrenberg, 1999; Powell & DiMaggio, 1991; Shapiro & Levine, 1999). Unfortunately, what is truly driving institutional decision-making about L/L programs is beyond the scope of this work and, perhaps, unknowable.

Although researchers and practitioners may not have clarity around why college and universities implement L/L programs, their choices may have implications on the extent to which students from different groups can benefit from L/L programs. Recall that Kuh et al. (2007) asked whether institutions "have the will to more consistently use what we know works in order to increase the odds that more students complete their program of study and benefit in the desired ways" (p. 70). Although there may be some lingering debate as to whether it is known whether L/L programs work, it seems how institutions choose to implement them is an example of their will to work for the benefit of all students, particularly those who are underserved. At least one study, that of Soldner et al. (2007), has suggested that one thematic type of L/L program—the honors L/L program—has been implemented in a way that restricts its benefits to students who, in general, are already poised to reap the benefits of postsecondary education. Soldner et al.'s finding that those underrepresented students who were able to participate in honors L/L programs accrued benefits beyond those of their same-race or same-SES peers in traditional residence hall environments demonstrates the costs associated with institutions' apparent lack of will. It is important to note, however, that honors L/L programs represent only one form of L/L program. Researchers know nothing of the equity, or differential efficacy, of the disaggregated L/L universe.

Who, then, do L/L programs benefit, and to what extent? These questions inspire this research, which seeks to (a) disaggregate L/L programs by their thematic type, (b) use the lens of thematic type and the Hao (2001) equity index equation to broaden investigations of equity beyond honors L/L programs, and (c) build upon what is known about how to more accurately assess L/L programs' efficacy, so that the consequences of participatory inequity for underserved students can be described. In the next chapter, the specific research methods used in this study are detailed. In so doing, this dissertation aspires not to be the last word concerning L/L programs, but, hopefully, the first to spark a potentially transformative dialog about equitable institutional use of high-impact undergraduate education programs.

Chapter 3 : Research Methods

Purpose Statement

The purpose of this study was three-fold. First, to facilitate comparisons across like types of programs, it sought to develop a thematic typology of L/L programs based upon their stated goals of objectives. Next, inspired by a desire to extend Soldner et al.'s (2007) study of participatory equity within Honors L/L programs, this study attempted to determine whether students from traditionally non-dominant racial/ethnic and socioeconomic backgrounds demonstrated equitable participation rates across a wider range of L/L program types. Finally, for students belonging to groups underrepresented in L/L programs, it sought to determine the benefits associated with participation in L/L programs versus living in a traditional residence hall (TRH) environment.

Substantive and Instrumental Research Questions

As noted earlier, this study's substantive research question asks:

Do students from various racial/ethnic and SES groups who are underrepresented in a particular thematic type (or types) of L/L programs report better outcomes than their peers living in the TRH environment?

To address that question, three instrumental research questions are offered. Initially, to facilitate comparisons between thematic types of L/L programs:

1. First, what type of thematic typology of L/L programs could be created, based upon those programs' stated goals and objectives.

Then, to evaluate questions related to participatory equity and student outcomes:

2. Did students from different racial/ethnic backgrounds or socioeconomic statuses participate in various thematic types L/L programs at a rate proportional to their presence in the larger L/L community?

3. Holding other student characteristics constant and driven by six key measures representing peer, faculty, and hall environments, did students underrepresented in a given thematic type (or types) of L/L program accrue differential benefits by participating in L/L programs when compared to their peers living in TRH environments?

Conceptual Frameworks

Chapter 2 offered detailed descriptions of three conceptual frameworks used in this research: (a) Hao's (2002) equity index, (b) Astin's (1991) IEO model, and (c) Pascarella and Terenzini's (1980) structural mediation model. Below, I describe the particular use of each to answer this study's second and third research questions.

Conceptual framework applied to instrumental research question two. Research question two—whether participatory equity exists in different types of L/L programs—was assessed using Hao's equity index. I computed nine separate indices for each thematic type of L/L program identified in the 2007 NSLLP dataset by research question one: six based upon students' racial/ethnic identifications (i.e., African American, American Indian or Alaskan Native, Asian/Pacific Islander, Latino, Multiracial, or White) and three based upon a composite measure of students' socioeconomic status (i.e., high, medium, or low).

In this study, the IE's numerator was calculated by dividing the number of individuals in a specific race/SES group participating in a particular type of L/L program by the total number of students participating in that specific type of L/L program. I calculated the denominator by dividing the number of individuals in a particular race/SES group participating in all L/L programs by the total number of all students participating in all L/L programs. The general format for the equity indices used in this research appears below.

Equation 3.1

Generic Equity Index Used in This Study

$$EI = \frac{Students\ in\ group\ in\ one\ L/L\ type \big/ All\ students\ in\ that\ L/L\ type}{Students\ in\ group\ in\ all\ L/Ls \big/ All\ students\ in\ all\ L/Ls}$$

As noted earlier, the resulting equity index can range from zero to one. A value of one indicates perfect proportional representation. Bensimon et al. (2003) have defined values above .80 as “almost at equity” (p. 14). For the present study, indices at or below .80 were considered evidence of inequity.

Conceptual frameworks applied to instrumental research question three. Research question three—whether students from racial/ethnic or socioeconomic groups underrepresented in certain thematic types of L/L programs accrued differential benefits by virtue of their participation when compared to their TRH peers—was evaluated using Astin’s (1991) IEO model, adjusted to consider the influence of different residence arrangements, as suggested by Pascarella and Terenzini (1980). Unlike traditional IEO models, which typically employ block-entry linear regression to determine the direct effects of one or more independent variables on the outcome of interest, the model used here employed structured latent mean modeling. This approach, described in more detail below, counts among its benefits the ability to model both direct and indirect effects, the absence of statistical assumptions that limit other linear approaches (i.e., univariate or multivariate analysis of variance or covariance), and greater statistical power (Hancock, 2003; Thompson & Green, 2006).

A simplified representation of the model appears in Figure 3.1, and Table 3.1 summarizes all constructs in the model and their associated variables (note that not all variables may remain in the model, should confirmatory factor analysis suggest they be removed). Not depicted in the figure for the sake of clarity, but present in the models analyzed, are covariances between each of the six key living and learning environments. Consistent with the IEO model, I hypothesized that both students' pre-entry characteristics and their engagement with important collegiate environments directly influenced their attainment of important outcomes (Astin, 1991). Similarly, as suggested by Astin, the model hypothesized that pre-entry characteristics also had the capacity to influence outcome attainment in an indirect manner, attenuating student engagement with key living and learning environments.

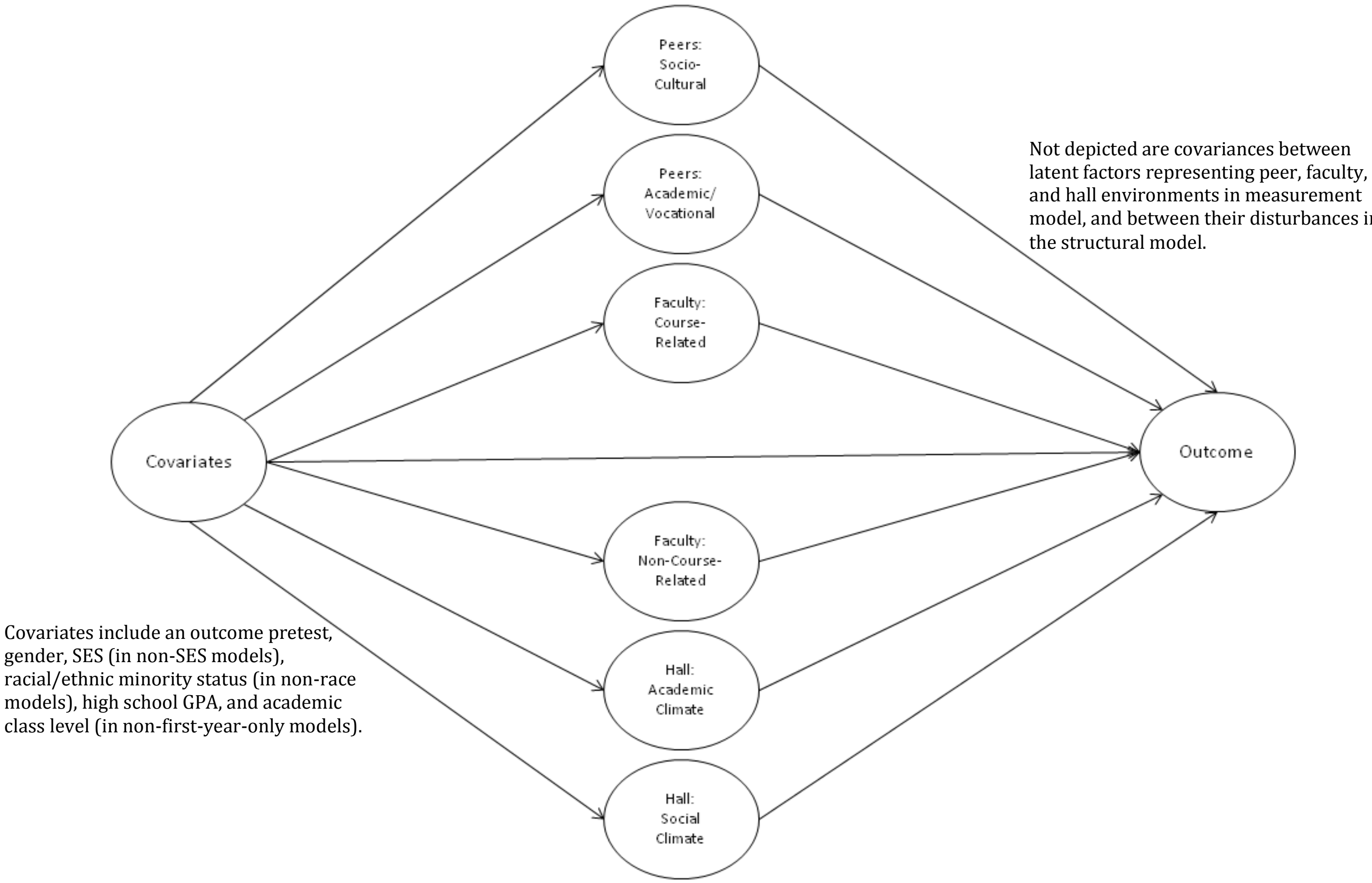


Figure 3.1

Simplified Model of Residential Arrangement's Influence on Student Learning Outcomes

Table 3.1

Summary of Model Elements, Prior to Confirmatory Factor Analysis

Element	Variable
Covariates	
	Gender
	High school GPA
	Race/ethnicity, in non-race/ethnicity-based models
	Socio-economic status, in non-SES-based models
	Academic class year, in models not constrained to first-year students only
	Proxy-pretest of outcome measure
Peers: Socio-cultural conversations	
	Discussions with students whose political opinions were very different
	Discussions with students whose religious beliefs were very different
	Discussed major social issues, such as peace, human rights, and social justice
	Discussed your views about multiculturalism and diversity
	Discussions with students whose personal values were very different
Peers: Academic-vocational	
	Shared your concerns about classes or assignments
	Discussed something learned in class
	Talked about current news events
Faculty: Course-related	
	Visited informally [with a faculty member] before or after class
	Made an appointment to meet in [a faculty member's] office
	Asked [a faculty member] for information related to a course you were taking
	Worked on a research project [with a faculty member]
Faculty: Non-Course-related	
	Discussed personal problems or concerns [with a faculty member]
	Discussed academic problems or concerns [with a faculty member]
	Visited informally [with a faculty member] during a social occasion
Hall: Academic Climate	
	My residence hall clearly supports my academic achievement
	Most students in my residence environment (RE) study a lot
	It's easy for students to form study groups in my RE
	Staff in my RE spend a great deal of time helping students succeed academically
Hall: Social Climate	
	Students in my RE are concerned with helping and supporting one another
	I find that students in my RE have an appreciation for people from different religious backgrounds
	Life in my RE is intellectually stimulating
	I find that students in my RE have an appreciation for people from different races or ethnic groups
	I would recommend this RE to a friend
	I see students with different backgrounds having a lot of interaction with one another in my residence hall
	I have enough peer support in my RE to succeed academically
Outcome	
	Varies

Rationale for Selection of Model Elements

Research question three posited a specific model of how students' background characteristics, residential arrangement, and key living and learning environments interacted to promote student outcome attainment (see Figure 3.1). As depicted above, that model consists of several elements, roughly corresponding to Astin's (1991) notion of inputs, environments, and outcomes. My rationale for including each element in the model, including relevant citations from the literature, is described below. Also described is how each element will be operationalized using items selected from the NSLLP's student survey.

Inputs. Six student characteristics were selected for inclusion in the model: (a) a proxy-pretest, designed to capture students' pre-college disposition toward the student learning outcome under investigation, (b) gender, (c) race/ethnicity, (d) socioeconomic status (a composite of family income and parental education), (e) high school GPA, (f) academic class level.

The first variable considered an input within Astin's (1991) framework was a single-item proxy-pretest (PRETEST) designed to assess students' pre-college motivation to experience growth on the student learning outcome of interest. In the absence of a true, parallel, pre-college measure of the dependent variable, proxy-pretests—students' retrospective ratings of their pre-college selves relative to that outcome—are preferable to no statistical control of any sort (Pascarella, 2001). Each proxy-pretest was of the same general form: "Looking back to before you started college, how important did you think it would be to …" Because the specific pretest to be used in each model was identified through the process of investigating the study's second research question, it is not described here. Instead, it is presented in Chapter 4, paralleling the chronology of my work.

As noted in Chapter 2, recent empirical research has suggested a connection between being female and increased intellectual development (Inkelas, Johnson et al., 2006; Inkelas & Weisman, 2003; Pasque & Murphy, 2005; Stassen, 2003), and a relationship between being female and reporting more openness to diversity (Longerbeam, 2005). Due to their low presence in the total dataset (less than .01%), transgendered students were excluded in analyses conducted as part of the third research question. As a result, gender was be entered into the model as an indicator variable, with students identifying as male coded as “0” and those identifying as female coded as “1” (FEMALE).

Some studies have suggested that race/ethnicity bears upon intellectual development, although patterns of significance do not consistently favor one racial/ethnic group over another (Inkelas & Weisman, 2003; Inkelas, Johnson et al., 2006; Pasque & Murphy, 2005). Similarly, racial minority status has been negatively related to persistence by some (Pike et al., 1997), while others have found mixed relationships between racial/ethnic group membership and students’ persistence and/or transition (Inkelas & Weisman, 2003; Stassen, 2003). White students were coded as “1” and students identifying as non-White were coded as “0” (WHITE).

Perhaps not surprisingly, pre-college academic characteristics appear to bear a relationship to each class of outcomes under investigation in this study. For example, studies have drawn a positive relationship between high school GPA (Pasque & Murphy, 2005) and intellectual development. Similarly, Stassen (2003) found a positive relationship between GPA and persistence. Respondents reported their GPA on a six point scale (i.e., 1=”A+ or A,” 2=”A- or B+,” 3=”B,” 4=”B- or C+,” 5=”C or C-,” or 6=”D+ or lower), or indicated their high school did not use grades. Nineteen students (five White, two African American, six Asian/Pacific Islander, and six low-SES) in the sample marked the “no high school GPA” option, subsequently recoded as missing (GPA).

The role of parents has been explored vis-à-vis transition and persistence within L/L programs. Findings have generally supported the contention that greater levels of parental encouragement and socioeconomic resources result in smoother transitions (Inkelas & Weisman, 2003) and greater likelihood of persistence (Pike et al., 1997). Including parental SES alongside the background variables listed above was designed to extend researchers’ knowledge of the relationships between this characteristic and other learning outcomes, including those that are academic or psychosocial in nature. Because SES is a *function* of characteristics like income and education, not an innate, latent factor that can be measured by income and education, an SES composite was created (see Kline, 2006 for a discussion of formative, rather than reflective, measurement).

The composite was constructed from two NSLLP variables: (a) family income, measured on a 9 point scale (i.e., 1=”less than $25,000” to 9=”$200,000 or more”), and (b) the highest reported level of parental education, measured on a 6 point scale (i.e., 1=“High school or less” to 6=“Doctorate or professional degree”). The parental education variable was multiplied by 1.5 to transform its range and then added to the family income measure. The resulting composite was then split into terciles to create the low-, medium-, and high-SES groups that would later be used for equity calculations (SES).

The last student characteristic considered was a student's academic class level. Researchers' previous studies have identified relationships between students' class standing and their self-reported intellectual growth (Inkelas, Johnson et al., 2006) and openness to diversity (Longerbeam, 2005). As noted in the review of the literature, these findings are not surprising: presumably, class level acts as a proxy for maturation. For the purpose of this study, class level was scaled from 1 (first-year student) to 4 (senior) (CLASSLEV).

Key living and learning environments. Given the centrality of students' residential experience to this research, six important living and learning environments—two each related to residence hall climate, faculty interaction, and peer interaction—were included in the model. Literature reviewed in Chapter 2 of this dissertation suggested each of these environments are, in some way, related to students' attainment of important student learning outcomes.

Past research has related students' perceptions of academically and socially supportive residence halls with their subsequent intellectual development, albeit somewhat inconsistently (Inkelas, Johnson et al., 2006). Similarly, both academically and socially supportive halls have been linked to students' reports of more smooth academic transitions to college (Inkelas & Weisman, 2003). Perhaps not surprisingly, Longerbeam (2005) noted a positive relationship between students' perception that their residence hall was academically and socially supportive and their subsequent openness to diversity.

For this purpose of study, students' perceptions of their residence hall climates were believed to be latent factors represented by a series of indicator variables. As is described in more detail below, preliminary exploratory factor analyses of 2007 NSLLP data suggested which RES items were most strongly related to these latent constructs (Inkelas, Brower et al., 2008). They included students' ratings of agreement (four points ranging from "strongly disagree" to "strongly agree") with the following questions:

- *Students' perception that their hall is academically supportive (*prior *α=.800)*
    - "My residence hall clearly supports my academic achievement." (ACADH1)
    - "Most students in my residence environment study a lot." (ACADH2)
    - "I think it's easy for students to form study groups in my residence environment." (ACADH3)
    - "I think the staff in my residence environment spend a great deal of time helping students succeed academically." (ACADH4)
- *Students' perception that their hall is socially supportive (*prior *α=.878)*
    - "Students in my residence environment are concerned with helping and supporting one another." (SOCH1)
    - "I find that students in my residence environment have an appreciation for people from different religious backgrounds." (SOCH2)
    - "Life in my residence environment is intellectually stimulating." (SOCH3)

- "I find that students in my residence environment have an appreciation for people from different races or ethnic groups." (SOCH4)
- "I would recommend this residence environment to a friend." (SOCH5)
- "I see students with different backgrounds having a lot of interaction with one another in my residence hall." (SOCH6)
- "I have enough peer support in my residence environment to succeed academically." (SOCH7)

Students' interactions with faculty and peers were also included in the proposed model. Faculty interaction, in general, has been linked to L/L participants' intellectual development (Pike, 1999), as has course-related faculty interaction (Inkelas & Weisman, 2003). Inkelas and Weisman noted that course-related faculty interaction was also positively related to students' smooth academic transition to college. Similarly, generic peer interaction has been linked to L/L participants' intellectual development (Pike, 1999), as have specific forms of that interaction, particularly socio-cultural conversations (Inkelas, Johnson et al., 2006; Inkelas & Weisman, 2003). Finally, Inkelas and Weisman also noted a positive relationship between a different form of peer interaction—that which centered on academic and vocational issues—and students' academic transition.

Similar to students' perceptions of the academic and social support found in their residence hall, students' ratings of their faculty and peer environments were also believed to be latent factors represented by indicator variables. They includes students' ratings of the frequency (four points, ranging from "never" to "once or more a week") of the following occurrences:

- *Perception of course-related faculty interaction (*prior *α=.740)*
    - "Visited informally [with a faculty member] before or after class" (CRSEF1)
    - "Made an appointment to meet in [a faculty member's] office" (CRSEF2)
    - "Asked [a faculty member] for information related to a course you were taking" (CRSFE3)
    - "Worked on a research project [with a faculty member]" (CRSEF4)
- *Perception of non-course-related faculty mentorship (*prior *α=.738)*
    - "Discussed personal problems or concerns [with a faculty member]" (MENTF1)
    - "Discussed academic problems or concerns [with a faculty member]" (MENTF2)
    - "Visited informally [with a faculty member] during a social occasion" (MENTF3)
- *Perception of academic/vocational conversations with peers (*prior *α=.806)*
    - "Shared your concerns about classes or assignments [with peers]" (ACADP1)
    - "Discussed something learned in class [with peers]" (ACADP2)

  - "Talked about current news events [with peers]" (ACADP3)
- *Perception of socio-cultural conversations with peers (*prior *α=.885)*
  - "Held discussions with students whose political opinions were very different than your own" (SOCP1)
  - "Held discussions with students whose religious beliefs were very different than your own" (SOCP2)
  - "Discussed major social issues, such as peace, human rights, and social justice [with peers]" (SOCP3)
  - "Discussed your views about multiculturalism and diversity [with peers]" (SOCP4)
  - "Held discussions with students whose personal values were very different than your own" (SOCP5)

Outcomes. The final element of the model depicted in Figure 3.1 is the student learning outcome of interest. Each NSLLP outcome is typically represented by a factor consisting of three to four items (Inkelas, Brower et al., 2008). Because the specific outcome to be evaluated in each model was identified through the process of investigating the study's second research question, it is not described here. Instead, the psychometric properties and constituent items of the factors representing the outcome or outcomes are presented in Chapter 4, paralleling the chronology of my work.

Summary. The model described above was informed by Astin's (1991) IEO model, which suggests that postsecondary outcomes are a function of both students' pre-college characteristics, including demography and existing levels of development, and the college environment itself. To that end, the model evaluates the relationships among student characteristics such as gender, race/ethnicity, high school performance, SES, a proxy-pretest and the outcome of interest. Similarly, key living and learning environments are included in the model, specifically those identified by Pascarella and Terenzini's (1980) Structural Mediation Model—which suggested peer, faculty, academic, and social environments were driving outcomes observed in living-learning programs—and later L/L research detailed in Chapter 2. Pre-college characteristics are posited to bear upon these environments, and, in turn, each environment is hypothesized to have an association with student outcomes and other environments in the model.

As described above, the initial model consisted of a maximum of 36 manifest variables, used individually or as indicators of latent factors: (a) four input variables (not counting a race/ethnicity or SES constant), (b) 27 environment variables, (c) the proxy-pretest variable, and (d) up to four outcome variables. This count directly influenced the minimum sample sizes sought for each analysis. Hancock and Mueller (2007) have suggested that samples contain at least five cases per variable, but preferably ten or more. As such, a minimum sample size of 180 cases per analysis was set as an *a priori* threshold for analysis. The consequences of this choice—a need to combine program types on the basis of their shared identification of a most highly rated learning outcome—are described further in Chapter 4.

The Sample

This study used two datasets from the 2007 NSLLP's baseline data collection: (a) a student dataset containing information provided by the study's student respondents about their college experiences, and (b) a program dataset containing information provided by administrators at each participating institution detailing their L/L programs. I describe both datasets below.

The student dataset. Each of the institutions taking part in the 2007 NSLLP was instructed to select either a full or a random sample of students participating in their L/L programs, along with a comparison sample of resident students not participating in any L/L program. The full sample consisted of 109,679 students, of which 48,426 (44.2%) were L/L participants and 61,253 (55.8%) were students who lived in TRH environments. E-mail contact information for potential respondents was provided to the data collection contractor, who, beginning in January, 2007, sent up to three messages to students soliciting their participation in the study. Each message contained instructions on how to access the Web-based survey instrument (described below) using a confidential, unique identification number and outlined incentives for participation offered by students' home institutions, if any.

At the close of data collection in May, 2007, 22,519 students responded to the survey instrument: 11,630 (51.7% of respondents) were L/L participants and 10,889 (48.3% of respondents) were students who lived in TRH environments. The observed response rate to the 2007 NSLLP, 20.4%, was below the response rate for the 2004 study and lagged the 30% response rate Crawford, Couper, and Lamia (2001) have argued is typical for Web-based surveys. As a result, SSG constructed proportional weights so that the demographic characteristics of each institution's final dataset mirrored that of their original sample. Although available, because these weights were only germane to institution-level analyses, they were not used in this study.

The program dataset. Upon their enrollment in the NSLLP, participating institutions were asked to provide a complete list of L/L programs offered to resident students. In total, institutional contacts identified 611 L/L programs. In April, 2007, the SSG followed up with each institutional contact via electronic mail, sending them a message with instructions on how to access a Web-based survey instrument (described below) developed to collect more detailed L/L program data. Institutional contacts provided additional data on 461 programs, yielding a response rate of 75.5%.

Instrumentation

The datasets used in this study were based upon responses to two separate survey instruments, described below.

The student instrument. All student respondents completed the Residence Environment Survey-Baseline (RES-B) via the Web. The RES-B represents the "third generation" of NSLLP survey instrumentation, replacing previous versions known simply as the Residence Environment Survey (RES). The NSLLP research team initially developed the RES through pilot testing in 2002 and 2003, and finalized the instrument for the 2004 NSLLP administration. Throughout the process of instrument development, the research team took several steps to ensure the validity and reliability of the RES and the measurement scales derived from it.

Construct validity is the extent to which items constituting a measurement scale are a faithful operationalization of the construct they are believed to represent (DeVellis, 2003). To promote construct validity, a panel of 15 L/L program directors and two survey methodologists reviewed RES items and their related constructs during initial pilot testing, offering critical feedback to the research team (Inkelas, Vogt et al., 2006).

Although in construct validity a scale is evaluated relative to the definition of the latent variable it seeks to measure, in criterion-related validity a scale is evaluated relative to an external standard (DeVellis, 2003). DeVellis outlines a number of forms of criterion-related validity, including convergent validity (a measurement scale is correlated with a construct theoretically related to it) and divergent validity (a measurement scale is not correlated with a construct theoretically unrelated to it). Inkelas, Vogt et al. (2006) and Longerbeam (2005) explored inter-scale correlations on the RES using 2004 NSLLP data, and found evidence of both convergent and divergent validity. For example, two scales that measured students' conversations with their peers about academic and social issues evidenced a moderate correlation (r=.60), while scales measuring students' conversations about social issues was found to have a negligible correlation with their alcohol use (r=-.04) (Inkelas, Vogt et al., 2006; Longerbeam, 2005).

Although the RES-B is substantively similar to its predecessor (the RES), it varies in two important dimensions. First, due to a focus of the 2007 NSLLP on the influence of L/L programs on women's success in science, engineering, technology, and mathematics (STEM), the research team added several item sets to assess students' STEM experiences. Second, to minimize respondent fatigue, the research team shortened several RES scales that were to be replicated on the RES-B. To do so, I used SPSS's RELIABILITY procedure and 2004 NSLLP data to identify and delete existing RES items that reduced (or made only a minor contribution toward) their larger scale's Cronbach alpha. As a result, the final RES-B consisted of 65 questions soliciting a range of information, including students' (a) demographic data, (b) pre-college achievement, attitudes, and self-confidence data, (c) college academic and co-curricular experiences,

(d) perceptions of and engagement with the collegiate environment, (e) self-reports of growth on several cognitive and social-cognitive outcomes, and (f) future plans. A copy of the RES-B is provided in Appendix A.

Although the variation between the two instruments was believed to be minor, because the RES-B did differ slightly from the RES, the research team decided to re-evaluate the instrument's factor structure. Therefore, after all data had been collected, the research team used principal axis factor analysis with Varimax rotation to identify which items on the RES-B might represent underlying latent variables of interest and ultimately form measurement scales for future study. Thirty-seven factors emerged from this analysis. The team then conducted Cronbach's alpha reliability analyses on the item sets representing each of the factors. These analyses allowed us to determine which combination of items in the potential measurement scales maximized each scale's internal consistency. Once the optimal combination of variables was identified for each measurement scale, composite scores were computed through simple summation. Appendix B lists each measurement scale used in this study, its constituent items, and the scale's Cronbach alpha; prior item-factor loadings were not published (Inkelas, Brower, et al., 2008). It should be noted that the psychometric data listed in Appendix B was generated in advance of this study, using the full sample of study participants and different factor analytic techniques than were proposed here.

The program instrument. Detailed information about participating institutions' L/L programs was collected via the Living-Learning Programs Survey (LLPS), a survey administered over the Web. Instructions on how to access and complete the survey were sent originally to each institution's administrative contact, who either completed the instrument on his or her own for all programs on the administrator's campus, or forwarded the link to individual program directors. The LLPS consists of 30 questions, covering domains such as programs' (a) goals and objectives, (b) size, (c) student composition and method of selection, (d) staffing, budgets, and reporting lines, (e) faculty roles and course offerings, (f) academic and student affairs staff roles, (g) co-curricular activities, and (h) for programs with a STEM focus, special STEM-related curricular and co-curricular activities. A copy of the LLPS is provided in Appendix C.

Data Analysis

Instrumental research question one. This study's first research question asked whether a team of raters could develop a parsimonious, thematic typology of L/L programs. To do so, it relied upon a particular methodological approach found within the family of techniques known as content analyses. Neuendorf (2002) has argued that all content analyses share similar traits, including a focus on content-laden messages that pass between two or more parties (be they a couple discussing dinner plans, a group negotiating a complex issue, or an individual responding to a researcher through a survey instrument) and attention to the summarization of a large pool of messages into something more readily generalizable. Content analyses may be further distinguished by the purpose for which they are to be undertaken, including (a) descriptive analyses, (b) psychometric analyses, or (c) predictive analyses (Neuendorf, 2002).

This study, with its goal of developing a typology that can subsume numerous programs with a diversity of objectives into a small but useful number of categories, was a *descriptive* content analysis. The messages analyzed were administrative contacts' responses to the LLPS (program) instrument. I describe Neuendorf's (2002) nine steps for completing such an analysis below.

The first three steps in Neuendorf's (2002) process of content analysis are preparatory, taking place well before the researcher begins data analysis. In step one, the researcher conducts a review of relevant literature and settles upon one or more research questions to explore. This is followed by step two, in which the researcher determines which constructs must be assessed to address those questions. Finally, the researcher operationalizes those constructs for the purpose of measurement.

Accordingly, these initial steps were completed at the proposal phase of this research. The salience of L/L theme has been identified by both the theoretical (Lenning & Ebbers, 1999; Shapiro & Levine, 1999) and empirical literature (Clarke, Miser, & Roberts, 1988; Inkelas et al., 2005). This study's first research question emerged from my critique of that literature, which suggested I could benefit researchers and practitioners by revising existing thematic typologies of L/L programs. Discerning a program's theme, the central construct of interest, will be facilitated through the use of three questions on the LLPS (program) instrument, designed by Inkelas et al. (2004) to gather information about features past scholarship has characterized as central to the success of L/L programs (Shapiro & Levine, 1999; Schoem, 2004). Specifically, I relied upon respondents' answers to three LLPS questions: (a) a description of the program's explicit goals or objectives, (b) a respondent's rating of the importance (not at all important to very important) of a list of student outcomes measured by the 2007 NSLLP, and (c) a program's name. Specific outcomes rated by respondents included participants' ability to: (a) explore the meaning of facts when introduced to new ideas, (b) apply something learned in one class to another, (c) demonstrate growth in ability to critically analyze ideas and information, (d) demonstrate growth in developing values and ethical standards, (e) demonstrate openness to views that are different than their own, (f) learn about people from backgrounds other than their own, (g) volunteer for and/or perform community service, (h) report increased confidence in their academic abilities, (i) report increased confidence in interpersonal relationships, (j) feel a sense of belonging to the institution, (k) experience a smoother academic transition to the institution, (l) experience a smoother social transition to the institution, (m) develop a greater enjoyment of challenging intellectual pursuits, (n) develop healthy behaviors around alcohol and other drugs, and (o) develop greater self-awareness.

Neuendorf's (2002) fourth step in content analysis is the development of two tools used in the coding process: (a) an initial codebook that outlines the coding scheme raters use to measure the variables contained in the messages they are analyzing, and (b) a coding form on which raters report their findings. Developing the codebook, she noted, can be a complex task because it must serve two purposes. First, the researcher uses the codebook to provide raters instructions on how to complete each step of the coding process. Second, raters use the codebook as a "key," containing criteria that allow them to code each message in a reliable manner. Writing the codebook is made more difficult because, as each subsequent step of the content analyses unfolds, it is subject to revision as errors are uncovered or more clearly drawn coding criteria are sought. By comparison, the coding form can be a simple tally on which raters note variable characteristics for each message (Neuendorf, 2002).

To develop a codebook, its author must identify the initial criteria by which each message variable will be evaluated (Neuendorf, 2002). For the purposes of the proposed research, I identified an initial system with which raters could code each of a program's three message variables—in this case, those representing its goals and objectives, most important student learning outcomes, and name. Neuendorf offers little guidance for this process, suggesting only that rating criteria be specific so as to maximize inter-rater reliability. However, her description of an *evolving* codebook suggests that, in this case, the criteria must not be so rigid or fixed as to preclude the identification of an emergent theme. Finally, that her description of the content analysis process begins with a literature review suggests a codebook author should be informed by relevant theory (Neuendorf, 2002). As a result, Inkelas et al.'s (2004) original thematic typology of L/L programs was selected as the basis for this study's initial codebook. For the sake of providing specific coding criterion, suggested by Neuendorf, the initial codebook used the 26 secondary thematic types originally identified by Inkelas et al.

For this study, I developed an electronic codebook and coding form, using a Microsoft Access database to combine both functions. This was supplemented by a packet of information that provided a description of each code and instructions for each rater. A copy of the rater instruction packet, which included the initial codebook (e.g., a list of the thematic types identified in 2004) appears in Appendix D. A screen-shot of the coding database appears in Appendix E. As can be seen there, the top half of the screen displayed message information, including a description of each program's goals and objectives, ratings of the relevance of various NSLLP outcomes to that program, and the program's name. The bottom half of the screen contained a drop-down menu of all available codes, from which raters selected the code they were assigning to each program. Raters were also given the option to mark that their preferred code was "not listed" and were provided an empty text box where new thematic types could be suggested.

The instructions provided to raters outlined a three-step process for coding. First, the rater was directed to read a 50-word description of each program's goals and objectives, and compare it to the descriptions of the types of L/L programs identified by Inkelas et al. (2004). If the rater believed that he or she could definitively categorize the program using an existing code, he or she indicated its typological assignment in the coding form and proceed. If the description was missing or was inconclusive, the rater was asked to turn to the learning outcome(s) on which the program focused. If the rater believed the selected outcome(s) allowed for assignment to an existing category, he or she categorized the program and moved to the next program. If the second variable was missing or inconclusive, the rater consulted the program's name. Using this last piece of information, the program was assigned to an existing category, marked as being "unique" (that is, having a description, set of outcomes, or title that suggested a new category should be created to encompass it), or marked as being "uncategorizable" (that is, not having enough information to place the program in a category or suggest a new, encompassing category).

The fifth step in Neuendorf's (2002) process of content analysis is to determine what sample will be drawn from the larger message pool (population). Given the small number of messages (n=611) to be analyzed, a full population sample was possible for this study.

Neuendorf's (2002) sixth step in the process of content analysis is the training of raters and the determination of pilot reliability. I began by familiarizing the raters with the NSLLP project, the specific purposes of my own research, and how to use the electronic codebook and coding form. In that training, raters were given the chance to mock code several actual messages within the overall pool. Both messages that were complete and subjectively well written, as well as those that contained missing data or otherwise might have presented challenges to coders were demonstrated.

After raters reported that they were comfortable with the coding process, pilot coding began. The purpose of pilot coding was to verify the utility of the codebook and calculate preliminary reliability scores using a subsample of the total message pool (Neuendorf, 2002). In this phase, coders worked independently to make their coding decisions. Neuendorf has noted that there are no universally agreed upon standards for determining the number of messages that coders should rate as part of pilot testing, and that although suggestions typically range from 10% to 20% of the total sample, some have called for subsamples approaching 100%.

As suggested by Neuendorf (2002), approximately 20% of the program dataset was randomly selected (n=123) for this initial trial. Coders were then split in to two groups (one consisting of NSLLP-experienced raters and the other consisting of NSLLP-naïve raters), and each group received approximately half of the pilot cases in the Access coding database. Raters had one week to independently complete the coding process after receipt of the Access database containing their trial messages. Following coding, raters returned their databases via electronic mail to the researcher for analysis, most within two business days.

After coders completed pilot coding, I calculated preliminary inter-rater reliability statistics. As suggested by Neuendorf (2002), I used SPSS 15.0 to compute Krippendorff's alpha (Hayes & Krippendorff, 2007). This index reflects the extent of inter-rater agreement while controlling for agreement that might occur by chance, and is appropriate for nominal data coded by more than two raters (Neuendorf). Scores on Krippendorf's alpha range from 0.00 to 1.00, with values greater than .70 being preferred and values as low as .60 being acceptable for exploratory studies (Lombard, Snyder-Duch, & Bracken, 2002). Those results are detailed in Chapter 4. Neuendorf has suggested that, should reliabilities found to be below acceptable levels, a variety of remedies be implemented, including (a) further coder training, (b) clarification of rating criteria, and (c) the "identification of problematic coders" (p. 148). As will be seen in Chapter 4, while the latter step was not required, codebook refinement and further training was indicated after review the results of inter-rater agreement tests. Specific steps taken to improve rater consistency in preparation for final coding are described in more detail there.

The seventh step in Neuendorf's (2002) process of content analysis is final coding. In final coding, all messages are distributed among raters for evaluation. Neuendorf has suggested that at least 10% of messages should "overlap" between coders to calculate final reliability statistics. In this case, the small number of messages in the message pool allowed each rater to code a greater number of messages without undue burden. As described in Chapter 4, each coder ultimately rated at least 45% of the message pool. After raters have coded the message pool, the eighth step in Neuendorf's process is the computation of final reliability statistics. To do so, I re-computed Krippendorff's alpha indices.

The ninth and final stage of Neuendorf's (2002) process of content analysis is the tabulation and reporting of results. The rating of the majority of coders formed the basis for each program's assignment to its final membership in the new typology. After assignment, a listing of each category and its constituent programs was generated, along with the reliability coefficients associated with the analysis. Finally, so that it could be used to identify the learning outcome to be used in further analyses undertaken as part of research question three, the student learning outcome rated "most important" by the largest number of L/L program directors was identified for each category.

Neuendorf (2002) does not explicitly address the issue of identifying qualified raters. In the absence of specific guidance, and based upon feedback provided by the examining committee assembled for this dissertation, the following criteria for identifying raters were established: (a) an investment in the trustworthiness of the final product, (b) familiarity with the general functions of a residential life program, and (c) an interest in learning more about the NSSLP and living-learning programs. No specific knowledge of living-learning programs expected, as, *a priori,* there was no reason to assume specific experience with L/Ls would improve the quality of raters' efforts. A total of six raters were chosen, three of whom were members of the NSLLP Research Team and three others who were aware of the NSLLP's existence but were not connected to the project itself were selected. They included:

- *Kristan Cilente.* Ms. Cilente is a doctoral candidate in the College Student Personnel Administration program and, at the time of her selection, served as Coordinator of Community Service-Learning at the University of Maryland's Stamp Student Union. Prior to her work at the University of Maryland, Ms. Cilente served as Assistant Director of Orientation and Leadership Programs at Georgetown University. Ms. Cilente had no specific, professional knowledge of living-learning programs and was considered new to the project.

- *Marybeth Drechsler*. Ms. Drechsler is a doctoral candidate in the College Student Personnel Administration program and a member of the NSLLP research team. In addition to her work on the research team, Marybeth serves as a living-learning generalist for the University of Maryland's College Park Scholars program. Before coming to the University of Maryland, she was a full-time living-learning practitioner at the University of Missouri.
- *Daniel Ostick*. At the time of the coding procedure, Dr. Ostick was a doctoral candidate in the College Student Personnel Administration program and currently serves as Coordinator of Leadership Curriculum Development and Academic Partnerships at the University of Maryland's Stamp Student Union. Before that, Dr. Ostick was a professional staff member in the residential life systems of the University of Maryland and the University of Texas. Dr. Ostick had no prior connection to the NSLLP team.
- *Graziella Pagliarulo McCarron*. Ms. McCarron is a third-year doctoral student in College Student Personnel Administration and a member of the

NSLLP research team. Ms. McCarron is currently the Director of Student Development for Metropolitan College at The Catholic University of America. While earning her Master's degree at the University of Maryland, Ms. McCarron worked with the Beyond the Classroom Living-Learning Community.

- *José-Luis Riera.* Mr. Riera is a doctoral candidate in the College Student Personnel Administration program and currently serves as the Coordinator of the University of Maryland's Student Honor Council. Prior to coming to Maryland, Mr. Riera held several progressively responsible positions within Drexel's Residential Life Office, culminating in his appointment as Director of Residential Living. Mr. Riera had no prior experience with the NSLLP team.
- *Katalin Szelényi*. Dr. Katalin Szelényi is an Assistant Professor of Higher Education at the University of Massachusetts—Boston. Before her appointment there, she served two years as a postdoctoral research fellow with the NSLLP. She was directly involved in all phases of the 2007 NSLLP, including instrumentation design, data collection, and data analysis. She has extensive training in the use of qualitative research methods, and has taught research methodology courses at the University of California – Los Angeles and at the University of Maryland.

Instrumental research question two. This study's second research question investigated the status of participatory equity among students from different racial/ethnic backgrounds and socioeconomic statuses.

Race/ethnicity was the focus of my first group of equity analyses. The RES-B instrument asked student respondents to indicate their identification with any or all of the following racial/ethnic categories: (a) African American/Black, (b) Asian/Pacific Islander, (c) American Indian or Alaskan Native, (d) Hispanic/Latino, (e) White/Caucasian, or (f) Race/ethnicity not included. For the purpose of this analysis, I considered all students who selected more than one racial category (including "race/ethnicity not included") to be Multiracial. Due to complications in interpretation, I did not compute equity indices (EIs) for students who responded only with "race/ethnicity not included," although these students were not excluded from SES-based analyses. As such, I explored a total of six racial categories.

The second group of equity analyses focused on socioeconomic status. As noted earlier in this chapter, this construct was operationalized as the combination of a measure of annual family income along with the maximum value reported for maternal or paternal educational attainment. Once the new socioeconomic status composite was created, students were be partitioned into low-, medium-, and high-SES groups.

To determine whether inequities existed in the participation rates for students in each of the six racial categories and three socioeconomic statuses, I computed a series of equity indices for each type of L/L program. Recall that each equity index is meant to represent participatory equity by type of program, conditional on L/L participation. As such, the numerator of the EI was be calculated by dividing the number of individuals in a particular group (e.g., Hispanic/Latino students) participating in a particular type (e.g., honors) of L/L program by the total number of students participating in that type of L/L program and the denominator was calculated by dividing the number of individuals in a particular group participating in *all* L/L programs by the total number of students in all L/L programs.

An alternative method for calculating the EI—which would have changed the unit of analysis of the denominator to "resident students," rather than "all L/L participants"—would have generated an estimate of participatory equity by type of program, conditional on on-campus residence. While this might be seen as preferable to the approach employed above, the design of the NSLLP made it impossible. Recall that the NSLLP recruited a full or random sample of L/L participants, which should generate an unbiased measure of demographics at the L/L level, but a *matched* comparison sample of TRH students. Because this latter sample is a biased measure of TRH characteristics, the "resident students" denominator could not be calculated.

Because Bensimon et al. (2003) have characterized EIs of 0.80 or above as being "almost at equity" (p. 14), I considered a subset of type × group combinations that elicited an EI below that cut-off for more in-depth analyses in research question three. As will be described in Chapter 4, complications arising from small sample sizes resulted in the evaluation of three combinations, after being forced to combine types of programs that shared a similar most highly rated learning outcome.

Instrumental research question three. This study's final question focused on two groups of students: (a) L/L participants who were underrepresented in a certain type (or types) of L/L program on the basis of their race or SES, and (b) their same-race/ethnicity or same-SES peers who were living in TRH environments. I sought to contrast L/L participants' scores on a single student learning outcome—the outcome rated "most important" by L/L programs directors—with those of their non-participating peers. Theoretically, that difference represented the gain (or loss) associated with L/L participation, a gain (or loss) that a non-participating student might have accrued had he or she been able to access this particular form of high-impact program.

To do so, I evaluated the model proposed in Figure 3.1 using EQS version 6.1, build 94, to implement a structural equation modeling technique known as structured latent mean modeling (LMM). While traditional structural equation modeling is concerned only with the relationships between manifest or latent constructs, LMM addresses not only the relationships between constructs but also those constructs' means (Byrne, 2006). Inspection of Figure 3.1 reveals that, although its analysis requires the use of latent variable techniques, it is conceptually analogous to the more familiar MANCOVA: measured covariates and a grouping variable are believed to influence multiple dependent variables which, in this case, also are part of a causal structure. This use of LMM to specify this family of models is well-documented in the SEM literature, and includes the work of Sörbom (1978), Bagozzi and Yi (1989), and Kano (2000). The general process used to test this LMM model is set forth by Thompson and Green (2006) and Hancock and Mueller (2009), and is described later in this chapter.

Preparation of data files for analysis. First, I created separate data files for each of the student group × L/L program type(s) combinations identified through research question two. Then for each student group to be analyzed, I generated a comparison file of TRH students from that group, removing students attending institutions not represented in the corresponding L/L file. Cases were ineligible for selection if they met one of the following conditions: (a) one or more indicators of the dependent variable were missing, (b) students identified only with the “other” racial category, (c) students identified as transgender, or (d) students identified as graduate students or belonging to an “other” academic class level.

Treatment of missing data. Although cases with missing data on any indicator of the dependent variable were removed in the data screening process, missing data still remained on other predictor variables in each sample file. After discarding the common approaches of listwise and pairwise deletion on the basis of their inappropriateness due to small starting sample sizes and possible incompatibility with structural equation modeling (see Cool, 2000), two options suggested by Enders (2006) were considered: (a) parameter estimation on the basis of full-information maximum likelihood (FIML) and (c) missing data replacement via multiple imputation (MI).

Substantive differences exist between these two approaches, chief among them whether missing data are actually replaced. In FIML, the hypothesized model and patterns in the observed data are used in tandem to estimate the parameters (e.g., loadings and path coefficients) that would be expected to exist in a complete dataset, but no replacement values are generated (Enders, 2006). In MI, multiple replicates of the source dataset are created, each containing statistically tenable values for the missing data on the basis of observed data, not an *a priori* model (Yoo, 2009). This distinction has both practical and theoretical consequences. From the practical perspective, MI is burdensome, requiring the researcher to analyze multiple datasets and then combine the results. However, MI has the advantage of producing complete datasets that can be used in any future analysis. For its part, using FIML in an analysis is simple (in fact it is automatic in some programs) and produces one set of results that can be reported without further manipulation. However, missing data are never actually replaced, making reanalysis impossible. From the theoretical perspective, the choice of FIML or MI pits confidence in one's model versus confidence in the non-missing data: FIML's model-driven approach depends upon having identified the right relationships at the outset, while MI's data-driven approach depends upon having collected non-missing data that is otherwise "good."

Ultimately, MI was selected to address missing data. This choice was made for two reasons. First, suggestions by Enders (2006) and Graham and Schafer (1999) that MI may be more appropriate for small sample sizes and non-normal data—both the case in this research—made it an attractive option. Second, the choice of the MI approach made it easier to incorporate auxiliary variables in the augmentation procedure. Using auxiliary variables—variables that are hypothesized to influence either missing values or patterns of missingness in the data—can improve the quality of missing data replacement and can address a key assumption of missing data augmentation approaches: that data are either missing completely at random (MCAR) or that possible forces driving patterns of missingness can be accounted for (Enders, 2006; Yoo, 2009). Auxiliary variables can be used in either FIML or MI. However, because FIML is model-driven, auxiliary variables must be added to the model to be analyzed, increasing sample size burden. Because MI occurs before analysis, as many auxiliary variables as may be potentially relevant can be used, irrespective of sample size.

Because EQS 6.1 did not include a facility to conduct multiple imputation, features of two other statistical packages were employed: (a) SPSS 16.0, to evaluate the extent and type of missingness found in the data, and (b) LISREL 8.80, to conduct the multiple imputation. First, I used SPSS to determine whether missing data was missing completely at random (MCAR), or whether missingness was related to other variables in the model. Then, as suggested by Yoo (2009) and Enders (2006), I used LISREL to create 10 replicates of each data file. In addition to the variables already included in the model to be analyzed, more than two dozen auxiliary variables were included in the imputation model on the basis that they might exert an influence on either the extent of missing data or on the missing data itself. They included: (a) proxy-pretests for all NSLLP outcome measures, (b) constructs that past scholars had identified as potentially related to the outcomes of L/L participation, but had not been included in the model for the sake of parsimony (e.g., SAT score or academic major), and (c) indicators of other outcomes of L/L participation that L/L program directors had rated as highly important, but were not the focus of this study. Below, Table 3.2 lists the auxiliary variables used in imputation.

Table 3.2

Auxiliary Variables Used in Multiple Imputation

	Rationale for Inclusion		
Variable	Proxy-pretest	Literature review	Ancillary outcome
SAT score		✓	
Academic major		✓	
Hours/week spent on co-curricular activities		✓	
Pre-college import. of communicating with faculty	✓		
Pre-college import of getting to know others	✓		
Pre-college import of exploring the meaning of facts	✓		
Pre-college import of applying what was learned in class to another	✓		
Pre-college import of critically analyzing ideas	✓		
Pre-college import of developing values	✓		
Pre-college import of openness to opposing views	✓		
Pre-college import of learning about people from different backgrounds	✓[a]		
Pre-college import of volunteering	✓		
Pre-college import of feeling a sense of belonging	✓		
Pre-college import of doing well academically	✓[b]		
Indicators of "Diversity Appreciation" outcome (three items)			✓[a]
Indicators of "Smooth Academic Transition" outcome (three items)			✓[b]
Indicators of "Growth in Liberal Learning" outcome (three items)			✓
Indicators of "Sense of Belonging" outcome (four items)			✓

Notes.

[a] Included in LM model when outcome was diversity appreciation. (White students.)

[b] Included in LM model when outcome was academic transition. (Asian/Pacific Islander and low-SES students.)

Additional details surrounding the MI process can be found in Chapter 6.

Plan for analysis. In essence, this study's third research question asked: For each group × type combination identified in the second research question, were the mean scores of L/L students on that program type's most highly rated outcome equal to that of their TRH peers? Within the more traditional univariate, measured variable realm, this question could be evaluated with a simple t-test. Within the context of model depicted in Figure 3.1, however, this question was more complex. Indeed, this study sought to identify the value of a latent mean—a construct that cannot be measured except by inference—as that mean was influenced by both measured covariates (e.g., gender, high school grades, socio-economic status, race, and a proxy-pretest) and other latent constructs (e.g., key living and learning environments involving peers, faculty, and residence hall climates), and then to simultaneously test whether the value of that latent mean varied across two groups (e.g., L/L and TRH students).

To analyze the model depicted in Figure 3.1, multi-group structural equation modeling techniques identified by Byrne (2006) were combined with LM modeling techniques identified by Thompson and Green (2006) and Hancock and Mueller (2009). What resulted was a seven step process that was repeated across all 10 imputed datasets for each group × type combination identified as part of research question two. Those steps are described below.

First, the factor structure (also known as the measurement model) of the latent constructs in Figure 3.1 (i.e., those representing the outcome factor and the six key living and learning environments) was tested separately for L/L and TRH students, using each of the 10 imputed datasets generated for each group × type combination. In each case, the original factor structure identified by Inkelas, Brower et al. (2008) was used to specify the original model. As suggested by Finney and DiStefano (2006), EQS was instructed to estimate each model using its polychoric correlation matrix and to use robust estimators in the calculation of fit statistics, given the ordinal scaling of each indicator variable.

After each analysis, EQS outputs were reviewed to identify items that appeared to load on more than one factor or exhibited low loading (<.6) on a single factor, unless doing so would reduce the number of factor indicators to less than two. Items that demonstrated problematic behaviors in three or more of the ten imputed datasets were reviewed and, if both statistically and theoretically indicated, removed from the model. Then, model statistics were re-estimated.

This process of testing and respecification was repeated until all potentially problematic items were removed from each model's factor structure. Although not strictly necessary, to ease interpretation items removed from either the L/L or TRH measurement models due to bad fit were trimmed from the other model. Finally, data-model fit statistics were examined to determine whether the measurement model was a tenable representation of the observed data.

Following the guidance of Hu and Bentler (1999), joint fit criteria of a standardized root mean residual (SRMR) of less than .09 and a root mean square error of approximation (RMSEA) of less than .06 were adopted as being indicative of good fit. Because most items in the model were ordinally-scaled, Likert-type variables, robust (Satorra-Bentler) chi-squares were used to rescale RMSEA estimates (Byrne, 2006). However, since chi-square-based fit statistics are not robust to sample size, they were not used to assess data-model fit (Finney & DiStefano, 2006). Then, I proceeded to the next phase of analysis.

In this second step of analysis, I tested whether the measurement model was appropriate to be used with both samples simultaneously. Byrne (2006) has identified two purposes for this step: (a) to set the stage for the next phase of analysis, in which the researcher tests whether parameters in the model are invariant across groups (i.e., that the item-factor loadings would be identical for L/L and TRH students), and (b) to provide a standard against which to judge subsequent model adjustments. As such, I simultaneously evaluated the final L/L and TRH models identified through the previous step, and reviewed data-model fit statistics. Again, Hu and Bentler's (1999) joint fit criteria were used. After having verified that my *a priori* SRMR and RMSEA thresholds had been met for each model, analysis continued to the third step.

In the third step of analysis, I estimated each L/L and TRH model simultaneously, putting in place constraints that allowed me to determine whether item-factor loadings, factor variances, and factor covariances were identical across models (Byrne, 2006; Thompson & Green, 2006). In an iterative process, each model was analyzed using EQS, and output reviewed. Constraints that contributed to substantial data-model misfit were removed and models reanalyzed. As was the case in previous steps, constraints were considered problematic if, in three or more of the 10 multiply imputed datasets, they contributed to poor data-model fit. Once all untenable constraints were removed, data-model fit was reassessed using Hu and Bentler's (1999) joint fit criteria. After having verified that models remained tenable for the observed data, I proceeded to the next phase of analysis: testing the structural connections posited in Figure 3.1.

Paralleling the first three phases of model testing, the fourth step in data analysis involved testing L/L and TRH models separately. The goal of the fourth step was to determine whether the structural equations hypothesized to exist between key living and learning environments and the student learning outcome of interest were tenable, given the data collected. Importantly, it was also at this phase when measured covariates (e.g., gender, high school grades, race/ethnicity, socio-economic status, and the outcome's proxy-pretest) were included in the model. Because EQS does not allow measured, independent variables to be used when polychoric correlations are to be analyzed, covariance matrices were analyzed from that point forth. Robust estimators continued to be employed to attenuate error introduced by the use of Pearson correlations.

To begin the fourth step, L/L and TRH models were estimated separately, across each of the 10 multiply imputed datasets. Figure 3.1 depicts the structural relationships tested in the phase. Additionally, although the literature reviewed in Chapter 2 might have suggested a more nuanced approach, each of the model covariates was explicitly posited to have a direct effect upon each latent factor (e.g., the six key living and learning environments, as well as the outcome factor). Doing so ensured that the totality of possible direct and indirect effects was captured in the model, creating a more accurate picture of each covariate's influence. Finally, although not shown in Figure 3.1, I felt it important to explicitly allow for the prospect that the latent factors representing key living and learning environments might have one or more inter-relationships (e.g., a student's perception of her residence hall's social climate might be related to her perception of its academic climate). As such, the factor disturbances (i.e., the portion of the factor not accounted for by the covariates) of key living and learning environments were allowed to freely covary, taking on whatever relationship was evidenced in the data.

After estimation, EQS outputs were reviewed for indications that the hypothesized structural model was a poor fit to the observed data. Three sets of statistics were considered: (a) overall data-model fit statistics, vis-à-vis Hu and Bentler's (1999) joint fit criteria, (b) LaGrange Multiplier statistics, which identify the absence of potentially important structural relationships, and (c) Wald test statistics, which identify the presence of non-informative structural relationships. Fan and Hancock (2006) have cautioned researchers to use extreme caution at this phase of modeling, arguing that acting upon LaGrange Multiplier tests or Wald tests is tantamount to revising one's theory simply to match the data collected. Ultimately, I concluded that no model modifications would be made unless joint fit statistics indicated the presence of misfit, deciding atheoretical model modification was inappropriate if its only purpose was to marginally improve already acceptable fit measures. As such, after ensuring good data-structural model fit in both the L/L and TRH models, I proceeded to the fifth stage of analysis.

After having established that the L/L and TRH structural models had good fit individually, the fifth step evaluated their fit when separately, but simultaneously, analyzed. The purpose of this step, as was the case for the second step (above), was to prepare for subsequent phases of analysis in which the models would be tested simultaneously under the constraint of invariance (i.e., that the relationships between factors would be identical for L/L and TRH students) and to provide a baseline estimation of data-model fit (Byrne, 2006). As such, L/L and TRH models were evaluated simultaneously across all 10 multiply imputed datasets, and EQS outputs reviewed. After having verified that the data-structural model fit of the models in combination fell within Hu and Bentler's (1999) range of acceptability, I proceeded to the sixth stage of analysis.

In the sixth stage of analysis, the structural L/L and TRH models for each group × type combination were evaluated simultaneously, with constraints that forced covariate-factor and factor-factor paths to be equivalent across groups. After each analysis, EQS output was reviewed to identify which constraints were statistically untenable. Constraints were relaxed in a serial fashion, re-estimating the model at each point and determining which additional constraints, if any, needed to be released. After removing all problematic constraints, I re-evaluated data-structural model fit using Hu and Bentler's (1999) joint fit criteria. After verifying fit was satisfactory across all models, I proceeded to the last stage of analysis.

Finally, latent means were modeled. So that the mean structure of the model could be estimated, a unit constant was added to all latent factors and measured covariates. As suggested by Hancock and Mueller (2009), unit constants associated with latent factors in the TRH model were set to zero while the unit constants associated with latent factors in the L/L model were freely estimated. Models were simultaneously analyzed and, as was the case in previous tests of constrained multi-group models, EQS outputs were reviewed to identify constraints that were statistically untenable. An iterative process of constraint release and reanalysis took place until no there was no further evidence of model misspecification. At that point, data-LM model fit was evaluated a final time using Hu and Bentler's (1999) joint fit criteria. Finally, t-tests were conducted to determine whether the unit constants associated with the L/L group differed from zero (the TRH mean), indicating statistically significantly higher means in the L/L group.

Interpretation of final LM models. This study's third research question asked whether, holding other student characteristics constant, students underrepresented in a given thematic type (or types) of L/L program reported higher mean score son key student learning outcomes when compared to their peers living in the TRH environment. The "weight" associated with the latent outcome factor's unit constant represented the answer to that question. To interpret it, three questions were asked: (a) was the value statistically significant, (b) was the value positive or negative, and (c) what was the resulting magnitude of between-group difference. (Note that although the outcome factor was of primary significance given the research question, the values of all other latent factors—including those associated with key living and learning environments—were similarly interpreted.)

If the weight associated with the latent mean's unit constant was not statistically significant, then the null hypothesis—that the difference was indistinguishable from zero—would fail to be rejected and no further analysis was required. If significant, then the sign of the weight became important. Because the value of the TRH environment's unit constant was constrained to be zero, positive weights indicated that L/L students had comparatively more of the latent construct, while negative weights indicated that L/L students had comparatively less of the latent construct. Finally, to characterize magnitude of the difference between the two groups, the standardized effect size was computed in accordance with the technique identified by both Hancock (2001) and Thompson and Green (2006). That effect size represented the number of standard deviations the L/L mean was above or below the TRH mean. It is described below, and is abbreviated ES in text.

Equation 3.2

Calculating the Standardized Effect Size of Latent Mean Differences.

$$ES_{Diff} = \frac{\kappa_O}{\sqrt{\frac{N_T \psi_{O_T} + N_L \psi_{O_L}}{N_T + N_L}}}$$

Where K_O represents the weight of the unit constant associated with the outcome factor, N_T and N_L represents the number of TRH and LL cases in each analysis, and ψ_{OT} and ψ_{OL} represents the disturbance associated with the outcome factor in the TRH and L/L models, respectively.

A note about multiple imputation, fit indices, point estimates, and confidence intervals. Although MI has a number of strengths as a missing data replacement strategy, it presents two unique challenges to the modeler. First, as noted by Enders (2006), there is no agreed-upon method for the aggregation of SEM fit statistics when multiple imputation is used. In the spirit of full disclosure and providing the reader as much information as possible when making judgments about data-model fit, I present the observed minimums, averages, and maximums of each fit statistic in the results. In the case of the SRMR, these statistics were reported as $SRMR_{min}$, SRMR, and $SRMR_{max}$, respectively. Because EQS automatically computes a 90% confidence interval around its point estimates of the RMSEA, the smallest lower boundary ($CI90_{Lowest}$), the average (RMSEA), and the greatest upper boundary ($CI90_{Highest}$) were reported.

Second, multiple imputation requires the aggregation of point estimates through the use of what are known as "Rubin's Rules," described by Schafer and Olsen (1998) and others. Point estimates are aggregated through simple averaging. The calculation of those estimates' standard errors is more complex, and is a function of variances both within and between imputations. The benefit, however, is that standard errors can be utilized to create confidence intervals around aggregated point estimates, providing the reader additional context with which to judge a particular mean, loading, or covariance. A variety of software packages have been developed to automatically compute point estimates, standard errors, and confidence intervals using Rubin's Rules. For the purposes of this study, an R implementation of Schafer's CAT program was used (Harding & Tussell, 2009). Appendix H provides more information about the particular calculations employed.

Finally, a note about what confidence intervals can—and cannot—tell us seems warranted. Most readers are familiar with the general notion that confidence intervals represent the range of values within which I would expect a point estimate to appear at a given rate (typically 95%). However, as noted by Wolfe and Hanley (2002), many readers confuse the connection between confidence intervals and null hypothesis difference testing. It is true that, within one sample, a confidence interval that "overlaps" zero indicates that one must fail to reject the null hypothesis that the quantity of interest equals zero. When two samples are involved, overlapping confidence intervals are more complex. In the two-sample situation, non-overlapping confidence intervals do signal statistically significant differences between two groups. However, it is *not* that case that overlapping intervals indicate that one must fail to reject the hypothesis that two quantities are equal (Belia, Fidler, Williams, & Cumming, 2005). As such, the reader is urged to remember the context in which confidence intervals are presented in this work, particularly whether they are being highlighted as results of tests—such as tests of model invariance—in which statistically significant difference has already been established.

Chapter Conclusion

This chapter outlined three distinct methodologies for addressing the research questions posed by this study. The first research question, the development of a typology of L/L programs based upon their goals and objectives, relied upon a content analysis in which a team of raters thematized more than 1800 pieces of information—three variables on each of more than 600 programs — into a parsimonious classification system. The second, an analysis of participatory equity in L/L programs, asked how institutions are providing access to programs that some have argued have the chance to positively reinvent undergraduate education (Kuh et al., 2007; National Leadership Council, 2007). Finally, I addressed the third research question by using an advanced quantitative analytical technique that allows for the more accurate — and powerful — modeling of students' residential learning experience, and examines whether L/L participation is more educationally beneficial for underrepresented students than TRH environments. In the chapters that follow, I have presented both an accounting of the research effort and the findings of that work, as well as a discussion of the implications that work holds for all involved in the L/L enterprise.

Chapter 4 : Results of Thematic Typology Development

The First Instrumental Research Question

This study had three purposes. First, it sought to develop a thematic typology of L/L programs, so that programs with like stated goals could be more readily compared. Second, it aimed to extend the work of Soldner et al. (2007) and their study of participatory equity within Honors L/L programs by attempting determine whether students from traditionally non-dominant racial/ethnic and socioeconomic backgrounds demonstrated equitable participation rates across a wider range of L/L program types. Third, for students belonging to groups underrepresented in L/L programs, it sought to determine the benefits associated with participation in L/L programs versus living in a traditional residence hall (TRH) environment. As such, this study's first research question asked: Using L/L programs' goals and objectives as its foundation, could a team of raters develop a thematic typology of L/L programs, allowing more accurate comparisons of equity and outcomes between types of programs?

Addressing The First Instrumental Research Question

As directed by Neuendorf (2002), the process of content analysis began with the development of the tools raters were to use throughout the coding process, including (a) rater instructions, (b) a written codebook that described each code and (c) the Microsoft-Access driven coding database in which raters recorded the code they had assigned to each program. The initial codebook, adapted from Inkelas et al.'s (2004) original thematic typology, appears in Appendix D. No modifications were made to the original codes or their descriptions, although a small number of clarifications were offered in an attempt to minimize confusion in the early rounds of coding.

Once the codebook and rating program were developed, raters were trained on how to use the coding tools. During that training, care was taken to ensure that raters were shown examples of messages that were complete (that is, there was no missing data for a given program) as well as incomplete (that is, programs were missing descriptions or ratings of the importance of possible program goals), and messages that were likely to fit within the existing typology (e.g., a message that described a career or major exploration program) as well as those that were likely to represent thematic types of L/L programs the existing typology did not consider (e.g., a message that described a program focused on environmental sustainability). In this way, raters were familiarized with the range of information they would encounter during the real rating task and were able to pose questions about the coding materials' use to the researcher.

Pilot coding. Following training, raters were divided into two groups for the purposes of pilot coding: (a) group 1, whose members all came from the NSLLP research team, and (b) group 2, whose members had no formal exposure to the NSLLP. As suggested by Neuendorf (2002), approximately 20% of the program dataset was randomly selected (n = 123) for this initial trial, and distributed between the two groups (group 1 n = 62; group 2 n = 61). Raters had one week to independently complete the coding process after receipt of the Access database containing their trial messages. Following coding, raters returned their databases via electronic mail to the researcher for analysis, most within two business days.

Analysis began with the computation of simple agreement statistics, followed by the Krippendorf's alpha indices suggested by Neuendorf (2002). Group 1 raters were in perfect agreement about individual program's thematic type in 34 of 62 cases (55%), and two of three raters agreed in 54 of 62 cases (87%). Using SPSS 15.0, Krippendorf's alpha was computed for Group 1, and found to be .652. Group 2 raters were in perfect agreement about individual program's thematic type in 35 of 61 cases (57%), and two of three raters agreed in 58 of 61 cases (95%). Krippendorf's alpha for Group 2 was found to be .671. While within the .60 to .70 range of acceptability established by Lombard et al. (2002) for Krippendorf's alpha, the results suggested that either the existing typology failed to include all the thematic types found by raters in trial coding, that raters needed additional training, or both.

Reviewing raters' decisions by group identified emergent codes to be used in the first round of full coding and provided insight into the observed Krippendorf's alphas. For example, although Group 1 evidenced perfect agreement about programs' thematic types in 34 cases, five of those instances represented unanimous agreement that the program's type was "not listed." For three programs, the emergent code suggested by each rater was nearly identical, and included: (a) "religion," "faith-based," and "religion/faith-based," (b) "ROTC, "ROTC," and "ROTC," and (c) "cooking," "culinary/nutrition," and "domestic culinary." Because raters were unanimous in their identification of these codes, three emergent types—"religion/faith-based," "ROTC," and "culinary programs"—were added to the codebook to be used in the first round of full coding. Similarly, Group 2 raters also unanimously agreed that five programs in their pilot coding were "not listed," and reached consensus about an emergent type exhibited by two programs: “environmental awareness,” “sustainability,” and “eco-environmental-sustainability” were dubbed “environmentalism and sustainability” and added to codebook. Because of the hierarchical nature of Inkelas et al.’s (2004) previous typology, in which specific types of programs were nested within broader categories, the researcher tentatively grouped religion/faith-based programs with wellness programs, ROTC and environmentalism and sustainability programs with civic engagement programs, and culinary programs with fine arts programs.

A review of each group’s work also revealed that some disagreements between raters were due to confusion about definitions of specific program types. For example, although Inkelas et al.’s (2004) original typology intended multicultural/diversity programs to refer to issues of “domestic diversity” (p. IV-2), several raters applied the type to programs addressing international issues. Although this did not result in an entire group of raters erroneously coding any international/global programs as multicultural/diversity programs, it did preclude perfect agreement in several cases. Similarly, several raters used Inkelas et al.’s multidisciplinary type, described as referring to programs that “are often umbrella organizations which house several smaller communities clustered around a specific theme [such as] a ‘living-learning center’” (p. IV-2), to refer to programs that might more aptly be considered interdisciplinary (i.e., bridging one or more academic disciplines).

To minimize confusion stemming from multicultural/diversity programs, international/global programs, and multidisciplinary programs, two additional changes were made to the codebook. First, a section was added to the codebook instructions clarifying the distinction between domestic multicultural/diversity programs and programs that were truly international/global in scope. Second, a new code was added to the codebook for programs that were truly "interdisciplinary," grouped by the researcher with other disciplinary programs.

Full coding, round one. After pilot coding, 547 of the original 611 programs required coding: 488 that raters had not yet reviewed, 54 that raters did not unanimously code during trial coding, and 5 that raters had unanimously marked as "not listed" but for which no agreed upon code could be identified. Due to the relatively small number of programs to be reviewed, it was determined that Neuendorf's (2002) suggestion that each rater's pool of messages overlap with that of another coder by at least 10% could be substantially exceeded. The 547 programs to be rated were put into blocks of 50, and then assigned to one of six panels with three raters each: (a) Rater A, Rater B, and Rater C, (b) Rater B, Rater C, and Rater D, (c) Rater C, Rater D, and Rater E, (d) Rater D, Rater E, and Rater F, (e) Rater E, Rater F, and Rater A, and (f) Rater F, Rater A, and Rater B. As a result, each individual rater was assigned responsibility for between 45% (*n*=247) and 55% (*n*=300) of the dataset.

As raters were being informally polled to determine the most convenient starting and ending dates for this phase of coding, two of six raters indicated that, due to pre-existing plans, they would be unavailable for the in-person debriefing meeting that had been planned to follow the first attempt at full coding. As a result, one final change was made to coding materials before the process began: a free-text "comments" field was added to the Microsoft Access coding database, and raters were asked to provide a rationale for their coding decisions. In this way, the "voice" of the absent coders could be represented in the debriefing meeting, albeit incompletely. As was the case in pilot coding, raters were sent the codebook and coding database via electronic mail. Raters were given ten days to complete the first round of full coding, at which point all materials were returned to the investigator for analysis.

Following receipt of all coding databases, inter-rater reliability statistics for each rating panel were computed using SPSS 15.0. Krippendorf's alphas for panels A through E were .694, .658, .693, .656, .650, and .701, respectively, in each case above the threshold for acceptability established by Lombard et al. (2002). Across all panels, perfect agreement was reached on 298 out of 547 programs (or 54%), although in 16 instances (3.5%) raters agreed only that the program was "not listed" in the codebook provided for the first round of full coding. Twenty-one possible new codes emerged from this round of coding, although in no case was a program unanimously described by each rater on a panel using an emergent code. Newly suggested thematic types of L/L programs, and the number of times coders used each type, included: (a) academic transition (once), (b) agriculture (six times), (c) arts and sciences (once), (d) aviation (twice), (e) communication (three times), (f) criminal justice (once), (g) engineering/math/computer science (twice), (h) ethnic studies (twice), (i) film (once), (j) first-generation students (once), (k) general theme (once), (l) hospitality/culinary arts (once), (m) mathematics (once), (n) natural resources (twice), (o) outdoor recreation/sports (once), (p) politics (eleven times), (q) science and society (three times), (r) social interaction (once), (s) societal exploration (four times), that is, programs that explored a given locality, typically an urban center, (t) special-population serving (once), and (u) sports (once). It should be noted that some of the codes offered up by raters as new were in fact present in either the original codebook (e.g., outdoor recreation) or the expanded codebook created after trial coding (e.g., culinary arts).

Despite these inconsistencies, review of raters' responses, both as individuals and as members of a panel of raters, indicated two things. First, it appeared that the expanded codebook was still insufficient to describe the variation in thematic type found among L/L programs in the study. The prevalence of certain emergent codes, including those related to politics, agriculture, and societal exploration, hinted where the thematic typology might see expansion in future coding. Second, it seemed that inter-rater agreement was unlikely to improve with coders continuing to work independently. Genuine disagreement (that is, two out of three panelists rating a program in the same manner) was more frequent than apparent confusion (that is, each panelist rating a program differently), with 196 programs in the former category and 53 in the latter. As a result, all available raters were asked to convene for an in-person debriefing to review the coding process and, if possible, type the remaining 249 programs by consensus.

Full coding, round two. In an attempt to identify the thematic type of the remaining 249 programs in the LLPS dataset, the researcher convened a meeting attended by all available raters. As noted above, two of the raters were unavoidably absent. To ensure that his or her voice was included in the final coding meeting, missing coders left detailed notes in the Access coding database about how they reached his or her coding decisions. As such, although six coders participated in the whole of the coding process, only four coders were present at the final coding meeting.

The researcher began the meeting by leading the group through four introductory issues, including (a) a review of the coders' work to date, (b) a history of the evolving codebook, (c) an introduction of 21 possible codes that had emerged from the first round of full coding, and (d) using a laptop and LCD projector, a demonstration of the new Microsoft Access coding database, which now displayed not only program information, but also each panelist's coding of each program during the prior round, as well as any comments individual raters had provided (see Appendix F for a depiction of the screen used in this final version of the database).

Over the course of approximately three hours, the researcher led the four coders who were able to be present for the final coding meeting through each program remaining to be coded in the LLPS dataset. Raters were advised that their goal was to reach consensus as to the thematic type of each program, either by using an existing thematic type or by unanimously concluding that a new type should be established, or to conclude as a group that the program's thematic type could not be determined on the basis of available information. Raters were also encouraged to read the electronic notes left after the first round of full coding, including comments left by the two raters who were unable to attend the final coding meeting, and to consider whether those comments informed their work during the current coding process. Although it proceeded efficiently and without incident, one aspect of the second round of full coding merits comment.

At least two raters had professional knowledge about a small number of programs in the LLPS dataset that rendered them *de facto* experts when it came to determining their appropriate thematic type. In one instance, a rater had worked in a participating institution's housing office, providing her specific knowledge of several programs' purposes. In the other, a rater had visited a participating institution as part of a NSLLP-sponsored qualitative data collection, and had become well-acquainted with the school's programmatic offerings. Rater expertise appears to have had at least one positive consequence for the rating process: a rater's prior knowledge allowed a program with an inscrutable name ("Pangea") and no description to be properly categorized. However, rater expertise may also have introduced an element of bias in the process that had not been anticipated: a judgment about whether a program's stated purpose matched its actual implementation.

For example, a rater unaffiliated with the NSLLP presumably would have read a program description and, having no knowledge of how that program actually operated, assigned the program a thematic type based simply on the information provided. An expert rater, however, might have read a program description and, believing it to be inaccurate for any number of reasons, assigned the program a thematic type based upon his or her personal knowledge. For example, a program entitled "Public Affairs Learning Community" and described as "[attracting] students who are interested in understanding and impacting society at the local, state, national, and global levels" was thought by one naïve rater to be best described as a civic engagement program, and by another naïve rater to be best described as a political interest program. An expert rater, however, who had interviewed the program's director as part of her work with the larger NSLLP project, knew that the program's full name was "The Social Science and Public Affairs Learning Community," and believed it was a social science disciplinary program. Ultimately, the opinion of the expert rater dominated.

Taken together, these two examples highlight the possible benefits—and potential pitfalls—of using expert raters. Gauging the consequences of expert knowledge upon this sort of rating task is complicated by a number of factors, including raters' lack of conscious awareness of their own biases, distinguishing between a biased response and one that reflects a genuine disagreement or belief, and the lack of a technique to weigh the benefits accrued by the use of expert raters versus the costs. More than anything, perhaps, these examples suggest that the process of coding described here was as much art as it was science.

At the conclusion of the second round of coding, raters unanimously agreed upon 10 new thematic types. They included: (a) agriculture and pre-veterinary medicine programs, (b) communication and journalism programs, (c) general leisure programs, (d) graduate student programs, (e) law and criminal justice programs, (f) local community exploration programs, (g) mathematics programs, (h) new student transition programs for diverse populations, (i) political interest programs, and (j) transfer student transition programs.

Despite raters' efforts to identify the thematic type of every program in the LLPS, at the end of coding, 56 programs remained uncategorized due a lack of information. In 18 cases, the only data provided by participating institutions were the programs' names. In some instances, of course, a program's name had been sufficient for raters to determine its purpose: Consider the aptly named "Social Justice" or "Engineering House" programs. Among the uncategorized programs, however, names were less helpful, at some times suggesting multiple possible themes (e.g., "Living in a Free Environment" or "Campus Connection"), and at other times being wholly enigmatic (e.g., "SAIP" or "Newhouse"). In 36 cases, participant institutions provided both program names and descriptions. Unfortunately, 30 of the 36 programs came from the same institution, whose staff provided the same goal statement for each of their programs: "Explore interdisciplinary concepts, develop out-of-classroom connection to faculty, utilize [the local community] as a learning laboratory, examine common interest, guided by faculty while living with students who share the interest." While the raters felt it might be possible for the researcher to gather more information about these programs via the Web and code them independently, both the researcher and raters agreed that such a task fell beyond the approved analysis plan. With all possible programs categorized, the coding process ended, and raters were adjourned.

Final steps in the typology development process. The final thematic typology appears in Appendix G, and consists of 41 specific types of L/L programs. To remain faithful to the style of typology developed by Inkelas et al. (2004), each newly identified specific type of L/L program was subsumed, if appropriate, under a larger broad category. The five new codes that emerged from trial coding and the 10 new codes that emerged from full coding were organized such that: (a) agriculture/pre-veterinary medicine programs, communication and journalism programs, interdisciplinary programs, law and criminal justice programs, and mathematics programs were subsumed under the larger category of disciplinary programs; (b) culinary programs were subsumed under the larger category of fine and creative arts programs; (c) general leisure programs and local community exploration programs were subsumed under a newly created larger category of leisure programs; (d) new student transition programs for diverse populations and transfer student transition programs were subsumed under the larger category of transition programs; (e) spirituality/faith-based programs were subsumed under the larger category of wellness programs; and (f) sustainability and environmentalism programs were subsumed under the larger category of civic engagement programs. Three specific thematic types of L/L programs—graduate student, ROTC, and political interest programs—were not subsumed under any larger category, becoming stand-alone categories like honors programs or general academic programs in Inkelas et al.'s typology.

Finally, the names or descriptions of several thematic types were slightly altered for the sake of clarity. First, to minimize confusion stemming from Inkelas et al.'s (2004) use of the term multidisciplinary to refer to programs that served as umbrella organizations sponsoring a variety of sub-programs (e.g., College Park Scholars), they were renamed, simply, "umbrella programs." Then, outdoor recreation programs, a stand-alone category in Inkelas et al.'s original typology, were subsumed under the broad category of leisure programs. Finally, the broad category of transition programs was redefined to include the experiences of students of any class level, not just first-year students, making it possible to accommodate programs for transfer students.

Determining thematic types' most highly rated outcome. The final task associated with the first research question was to determine the most important student learning outcome associated with participation in each thematic type of L/L program, based upon ratings provided by program administrators. The purpose of doing so was to identify the outcome that would serve as the focal dependent variable for subsequent analyses of program efficacy in research question three. Administrators' ratings (4=very important to 1=not at all important) for each outcome were averaged across programs within a given thematic type, and then examined to find the most highly rated outcome. Table 4.1 lists all program types, along with the average ratings of NSLLP outcomes for each type.

Table 4.1

Ratings of the Importance of NSLLP Outcomes, by Thematic Type

Thematic Type	Outcome 1*. Ability to Explore Ideas	Outcome 2. Ability to Apply Learning	Outcome 3. Growth in Cognitive Skills	Outcome 4. Growth in Pers. Knowledge	Outcome 5. Growth in Liberal Learning	Outcome 6. Diversity Apprec.	Outcome 7. Community Service	Outcome 8. Academic Confidence
Civic Engagement								
Civic Engagement	3.20	3.40	3.40	3.80†	3.80†	3.80†	3.80†	3.20
Environmental	2.45	2.55	3.00	3.09†	3.00	2.91	2.82	2.64
Leadership	3.29	3.00	3.35	3.53	3.71†	3.59	3.41	3.24
Service / Social Justice	2.71	2.43	2.71	3.14	3.43†	3.36	3.36	2.43
Disciplinary								
Agriculture/Vet. Med.	3.33	3.33	3.33	3.17	3.17	3.17	2.83	3.33
Business	2.95	3.11	3.11	3.26	3.53	3.42	2.58	3.42
Comm./Journalism	3.00	2.67	3.00	3.00	3.33	3.33	3.00	3.33
Education & C.S.	3.08	3.17	3.17	3.33	3.58†	3.58†	2.83	3.25
Engineering	2.95	3.10	3.25	3.10	3.30	3.35	2.35	3.45
General Sciences	2.22	2.56	2.67	2.78	3.00	2.78	2.11	3.11
Health Sciences	2.87	2.93	3.07	2.93	3.13	3.13	2.87	3.27
Humanities	3.13	2.75	3.13	3.00	3.63	3.38	2.25	3.13
Interdisciplinary	2.17	2.17	2.50	2.83	3.33†	2.83	2.50	2.50
Law/Criminal Justice	2.50	2.00	2.50	2.50	2.75	2.50	2.50	3.25
Mathematics	3.33	3.67†	3.67†	2.67	3.33	3.00	2.67	3.33
Social Sciences	3.00	3.00	3.43	3.00	3.29	2.86	2.43	3.43
Fine and Creative Arts								
Fine Arts	2.77	2.63	2.86	2.71	3.29	3.37	2.40	2.89
Culinary Arts	2.33	2.00	2.50	3.33	3.67	4.00†	2.33	2.00
General Academic	3.18	3.29	3.29	3.29	3.59†	3.41	2.18	3.41
Graduate Student	3.50	2.50	3.50	3.50	3.50	3.50	1.50	4.00†
Honors	3.29	3.07	3.46	3.24	3.49	3.51†	2.71	3.24

Note. Ratings are from 4=very important to 1=not at all important.

* The full text of each outcome statement rated by administrators is located in Question 2 of the LLPS, reproduced in Appendix C.

† Indicates the highest rated outcome for a given type, including ties. Highest outcome may appear on continuation pages.

Table 4.1

Ratings of the Importance of NSLLP Outcomes, by Thematic Type (continued)

Thematic Type	Outcome 1*. Ability to Explore Ideas	Outcome 2. Ability to Apply Learning	Outcome 3. Growth in Cognitive Skills	Outcome 4. Growth in Pers. Knowledge	Outcome 5. Growth in Liberal Learning	Outcome 6. Diversity Apprec.	Outcome 7. Community Service	Outcome 8. Academic Confidence
International and Global Programs								
International/Global	2.85	2.18	2.72	2.92	3.74	3.77†	2.36	2.38
Language	2.00	2.23	2.23	2.08	3.00	3.15†	1.69	2.54
Multicultural	3.27	2.20	3.20	3.33	3.93†	3.87	2.53	2.60
Leisure								
General Leisure	2.00	3.00†	3.00†	2.50	3.00†	3.00†	2.50	3.00†
Local Exploration	3.00	1.50	2.00	4.00†	4.00†	4.00†	4.00†	2.50
Outdoor Recreation	3.00	2.71	3.00	3.00	3.57	3.57	2.14	3.14
Political Interest	3.00	2.00	3.00	2.29	3.00	2.71	1.86	1.86
Research				(No Response)				
Residential College	3.17	3.17	3.33	3.17	3.83†	3.67	2.50	3.33
ROTC	3.33	3.33	3.67	3.67	3.67	3.33	4.00†	3.33
Transition								
For 1st Year Students	2.88	2.88	3.04	3.08	3.35	3.38	2.58	3.23
Career/Major Explor.	3.11	3.00	3.56	3.33	3.44	3.33	2.22	3.56
For Diverse Pops.	3.00	3.00	3.00	3.00	4.00†	4.00†	2.50	4.00†
For Transfer Students	3.00	3.00	3.00	3.00	3.00	3.00	3.00	3.00
Umbrella	3.00	2.60	3.20	3.60	3.80	3.80	1.80	3.40
Upper Division	2.33	2.00	2.33	2.67	3.33†	3.33†	2.00	3.00
Wellness								
Spirituality	2.67	2.00	2.00	2.67	3.33†	2.33	2.33	2.00
Wellness	2.39	2.28	2.44	3.00	3.00	2.94	2.67	2.78
Women's								
Leadership	3.00	2.50	3.25	3.50	3.50	3.50	2.75	3.25
In STEM	2.85	2.92	3.08	3.08	3.38	3.23	2.31	3.54

Note. Ratings are from 4=very important to 1=not at all important.
* The full text of each outcome statement rated by administrators is located in Question 2 of the LLPS, reproduced in Appendix C.
† Indicates the highest rated outcome for a given type, including ties. Highest outcome may appear on continuation pages.

Table 4.1

Ratings of the Importance of NSLLP Outcomes, by Thematic Type (continued)

Thematic Type	Outcome 9*. Interpersonal Confidence	Outcome 10. Sense of Belonging	Outcome 11. Academic Transition	Outcome 12. Social Transition	Outcome 13. Intellectual Challenge	Outcome 14. Alcohol/Drug Behaviors	Outcome 15. Self-Awareness
Civic Engagement							
Civic Engagement	3.20	3.60	3.40	3.40	3.40	3.00	3.40
Environmental	2.73	3.09†	2.64	2.91	2.73	2.18	2.91
Leadership	3.53	3.65	3.59	3.53	3.29	3.00	3.71†
Service / Social Justice	3.14	3.14	2.79	3.00	3.14	2.36	3.21
Disciplinary							
Agriculture/Vet. Med.	3.17	3.33	3.50†	3.50†	3.17	3.17	3.17
Business	3.42	3.63†	3.58	3.53	3.21	3.11	3.47
Comm./Journalism	3.67†	3.67†	3.67†	3.67†	3.00	3.00	3.33
Education & C.S.	3.25	3.58†	3.58†	3.50	3.25	2.83	3.25
Engineering	3.45	3.65†	3.65†	3.50	3.45	2.70	3.15
General Sciences	2.78	3.78	3.89†	3.78	3.22	2.44	3.11
Health Sciences	3.20	3.33	3.60†	3.40	3.07	2.87	3.00
Humanities	3.00	3.38	3.75†	3.63	2.75	2.25	3.25
Interdisciplinary	3.00	3.00	2.67	3.00	3.17	2.33	2.83
Law/Criminal Justice	3.25	3.50†	3.50†	3.50†	3.00	2.50	2.75
Mathematics	3.00	3.67†	3.33	3.00	3.67†	3.33	3.33
Social Sciences	3.14	3.14	3.71†	3.57	3.14	2.14	2.86
Fine and Creative Arts							
Fine Arts	2.94	3.51†	3.20	3.26	3.34	2.51	3.17
Culinary Arts	3.67	3.67	2.83	3.50	3.33	3.33	3.33
General Academic	2.94	3.47	3.35	3.29	3.24	2.88	3.35
Graduate Student	3.50	4.00†	4.00†	4.00†	4.00†	3.50	4.00†
Honors	3.17	3.46	3.51†	3.39	3.46	2.68	3.29

Note. Ratings are from 4=very important to 1=not at all important.
* The full text of each outcome statement rated by administrators is located in Question 2 of the LLPS, reproduced in Appendix C.
† Indicates the highest rated outcome for a given type, including ties. Highest outcome may appear on continuation pages.

Table 4.1

Ratings of the Importance of NSLLP Outcomes, by Thematic Type (continued)

Thematic Type	Outcome 9*. Interpersonal Confidence	Outcome 10. Sense of Belonging	Outcome 11. Academic Transition	Outcome 12. Social Transition	Outcome 13. Intellectual Challenge	Outcome 14. Alcohol/Drug Behaviors	Outcome 15. Self-Awareness
International and Global Programs							
International/Global	3.10	3.10	2.84	3.18	3.16	2.34	3.28
Language	2.62	2.62	2.31	2.54	2.54	1.77	2.46
Multicultural	3.33	3.53	3.33	3.60	3.13	2.73	3.53
Leisure							
General Leisure	2.00	3.00†	3.00†	3.00†	3.00†	3.00†	2.50
Local Exploration	4.00†	4.00†	3.50	4.00†	3.50	2.50	3.50
Outdoor Recreation	3.29	3.86†	3.71	3.86†	3.00	2.57	3.43
Political Interest	3.00	3.00	2.14	2.86	3.43†	1.86	2.57
Research				(No Response)			
Residential College	3.17	3.33	3.33	3.17	3.33	2.50	3.17
ROTC	3.67	3.67	3.67	3.33	3.33	3.33	3.33
Transition							
For 1st Year Students	3.31	3.81	3.88†	3.62	2.96	2.81	3.54
Career/Major Explor.	3.22	3.89†	3.89†	3.78	3.33	2.89	3.44
For Diverse Pops.	3.00	3.50	4.00†	3.50	3.00	3.00	3.50
For Transfer Students	3.00	4.00†	3.00	3.00	3.00	3.00	3.00
Umbrella	4.00†	3.60	3.40	3.60	3.60	2.60	3.40
Upper Division	3.00	3.33†	3.00	3.00	2.67	3.00	3.00
Wellness							
Spirituality	2.33	3.00	2.67	2.67	3.00	2.33	3.00
Wellness	3.00	3.33	3.22	3.33	2.72	3.67†	3.56
Women's							
Leadership	3.00	3.75†	3.25	3.25	3.25	2.75	3.75†
In STEM	3.38	3.69	3.85†	3.69	3.46	2.92	3.31

Note. Ratings are from 4=very important to 1=not at all important.
* The full text of each outcome statement rated by administrators is located in Question 2 of the LLPS, reproduced in Appendix C.
† Indicates the highest rated outcome for a given type, including ties. Highest outcome may appear on continuation pages.

Although I had expected that one single outcome would be most highly rated for each thematic type of L/L program, such was not the case: Several ties (denoted with a dagger symbol) are evidenced in Table 4.1. The four most frequently cited top outcomes were (a) enhanced sense of belonging and smoother academic transition, each noted in 16 thematic types, (b) increased growth in appreciation for liberal learning, noted in 13 thematic types, and (c) increased growth in appreciation for diverse others, noted in 10 thematic types. An increased ability to explore facts and ideas was not noted as the highest outcome for any thematic type, while students' ability to apply knowledge across contexts, their growth in the ability to critically analyze ideas, and their increasingly positive behavior surrounding alcohol and other drug use were each rated highest by only two thematic types, and students' growth in personal values and ethical standards, their propensity for civic engagement, their academic or interpersonal confidence, and their increased self-awareness were each rated highest by only three thematic types.

From one perspective, these findings complicated planned analyses of programs' efficacy: The assumption had been that a single outcome would rise to the top for each thematic type of program, thus serving as the dependent variable in this study's third research question. Instead, for many program types, it appeared that program administrators felt several outcomes were identically (or nearly identically) as important for their participants to attain. Although the implication of this finding vis-à-vis the third research question was initially unclear, a seemingly positive note was that much of the ambiguity was reasonably related to program types and the outcomes with which they were affiliated: Disciplinary programs, for example, were nearly evenly split between fostering a smooth academic transition for its participants and promoting their sense of belonging, while international and global programs were primarily focused on promoting diversity appreciation and a growth in liberal learning.

From another perspective, however, these findings illuminated the notion that although L/L programs were highly variable in their themes (as determined by their goals and objectives), there was substantially less variation in the broader outcomes they hoped their participants would attain. The practical implications of this finding are discussed later, in Chapter 5.

Indeed, this dichotomy was not unlike that which is found in post-secondary education more generally: Students' content-area learning may vary substantially by major (the goal of which is to promote discipline-specific skills, something the NSLLP does not measure), but the broader outcomes of college attendance (with which the NSLLP is concerned) are relatively invariant irrespective of course of study. In Table 4.2, the four most commonly cited top outcomes for L/L programs—enhanced sense of belonging, smoother academic transition, growth in appreciation for liberal learning, and growth in appreciation for diverse others—are shown along with the thematic types of programs that rated them most highly. Among the 41 thematic types of L/L programs documented in this study, only five—political interest programs, research programs, ROTC programs, umbrella programs, and general wellness programs—did not include one of four commonly rated top outcomes as one of their highest.

Table 4.2

Thematic Types of L/L Programs, By Most Commonly Cited Top Outcomes

Outcome	Thematic Types and Rating[a]
Sense of Belonging	Civic Engagement: Environment (3.09)
	Disciplinary: Business (3.63)
	Disciplinary: Communication/Journalism (3.67)
	Disciplinary: Education (3.58)
	Disciplinary: Engineering (3.65)
	Disciplinary: Law/Criminal Justice (3.50)
	Disciplinary: Mathematics (3.67)
	Fine and Creative Arts: Fine Arts (3.51)
	General Academic (3.47)
	Graduate Student (4.00)
	Leisure: General Leisure (3.00)
	Leisure: Local Community Exploration (4.00)
	Leisure: Outdoor Recreation (3.86)
	Transition: Career and Major Exploration (3.89)
	Transition: Transfer Students (4.00)
	Upper Division (3.33)
	Women's: Leadership (3.75)
Academic Transition	Disciplinary: Agriculture (3.50)
	Disciplinary: Communication/Journalism (3.67)
	Disciplinary: Education (3.58)
	Disciplinary: Engineering (3.65)
	Disciplinary: General Sciences (3.89)
	Disciplinary: Health Sciences (3.60)
	Disciplinary: Humanities (3.75)
	Disciplinary: Law/Criminal Justice (3.50)
	Disciplinary: Social Sciences (3.71)
	Graduate Student (4.00)
	Honors (3.51)
	Leisure: General Leisure (3.00)
	Transition: First Year Student (3.88)
	Transition: Career and Major Exploration (3.89)
	Transition: Diverse Populations (4.00)
	Women's: In STEM (3.85)
Liberal Learning	Civic Engagement: Civic Engagement (3.80)
	Civic Engagement: Leadership (3.71)
	Civic Engagement: Service Learning and Social Justice (3.43)
	Disciplinary: Education (3.58)
	Disciplinary: Interdisciplinary (3.33)
	International/Global: Multicultural Programs (3.93)
	General Academic (3.59)
	Leisure: General Leisure (3.00)
	Leisure: Local Community Exploration (4.00)
	Residential Colleges (3.83)
	Transition: Diverse Populations (4.00)
	Upper Division (3.33)
	Wellness: Spirituality (3.33)

Outcome	Thematic Types and Rating[a]
Diversity Appreciation	Civic Engagement: Civic Engagement (3.80)
	Disciplinary: Education (3.58)
	Fine and Creative Arts: Culinary Arts (4.00)
	General Academic (3.41)
	Honors (3.51)
	International/Global: International/Global (3.77)
	International/Global: Language (3.15)
	Leisure: General Leisure (3.00)
	Leisure: Local Community Exploration (4.00)
	Transition: Diverse Populations (4.00)
	Upper Division (3.33)

Notes.
Thematic types listed twice rated outcomes identically, and those not listed rated a less frequently occurring outcome as their most important.
[a] Ratings, which appear after program types, are from 4=very important to 1=not at all important.

Chapter Conclusion

In this chapter, results pertaining to this study's first research question were detailed. I presented the results of a descriptive content analysis, in which a team of six raters reviewed the name, espoused goals, and brief description of 611 L/L programs with the objective of identifying each program's thematic type. As a result of their efforts, all but 56 of those programs were placed into a new thematic typology of L/L programs, consisting of 41 specific types and 17 broader (aggregate) types. While this expansion may reflect increased diversification of L/L programs by theme, it may also be an artifact of the substantially larger sample of L/L programs in 2007 than in 2004, when the number of programs participating in the study more than doubled in size.

After the typology was developed, I examined the learning outcomes most frequently associated with each programmatic type, based upon ratings provided by programs' directors. Four outcomes of L/L participation were consistently among those most highly rated across the majority of programs: (a) smooth academic transition to college, (b) sense of belonging, (c) appreciation for diversity, and (d) growth in appreciation for liberal learning.

In the next chapter, I describe how the newly developed thematic typology was used as the basis for calculating a series of race/ethnicity and SES-based indices describing programs' participatory equity.

Chapter 5 : Results of Evaluation of Participatory Equity

The Second Instrumental Research Question

This study's second instrumental research question investigated the status of participatory equity among students from different racial/ethnic groups and SES statuses for each thematic type of L/L program. Specifically, it asked: Did students from different racial/ethnic backgrounds or socioeconomic statuses participate in various thematic types of L/L programs at a rate proportional to their presence in the larger L/L community?

Addressing the Second Instrumental Research Question

To address this study's second instrumental research question, I employed Hao's (2002) Equity Index (EI). The EI is calculated by dividing two proportions: the percentage of students in a given thematic type of L/L program who belong to the group of interest and the percentage of students in the L/L population who belong to the group of interest. An EI of 1.0 indicates that a group is participating in a L/L program at a rate that matches their prevalence in the population, achieving perfect equity. Based upon the recommendation of Bensimon et al. (2003), inequity was considered to be present when EIs fell below .80.

To compute EIs by race/ethnicity and SES for each thematic type of L/L program, I added three variables to the full NSLLP student dataset: (a) a variable representing the thematic type of L/L program in which students participated, (b) a variable representing students' race/ethnicity, and (c) a variable representing students' membership in one of three SES terciles.

Identifying students by thematic type of L/L program. To connect each L/L participant to a particular thematic type of program, each thematic type identified in research question one was first assigned a numeric code in the program dataset. Then, program information stored in each L/L participant's student record was matched to the program dataset, copying the code representing thematic type to the student dataset.

Unexpectedly, a large number (*n*=2128, or 18.3%) of L/L participants failed to identify the program to which they belonged. Further investigation revealed that although all institutions demonstrated some degree of missingness, three institutions were responsible for a total of 30.1% of the missing data. Table 5.1 lists each thematic type of L/L program (excluding programs that could not be coded), along with the number and percent of students participating in each. As can be seen there, nearly half of L/L participants came from two broad categories of programs: disciplinary programs (encompassing 12 specific subtypes of programs, comprising 26% of the sample) and honors programs (with no specific subtypes, comprising 24% of the sample).

Table 5.1

Distribution of L/L Participants Among Different Types of Programs

Thematic Type of L/L Program	Valid % of Total L/L Participants
Civic Engagement	
Civic Engagement	.57%
Environmental	.79%
Leadership	2.80%
Service/Social Justice	1.41%
Disciplinary	
Agriculture/Vet. Med.	.89%
Business	3.25%
Comm./Journalism	1.58%
Education	2.33%
Engineering & C.S.	5.18%
General Sciences	5.03%
Health Sciences	3.35%
Humanities	1.44%
Interdisciplinary	.50%
Law/Criminal Justice	.38%
Mathematics	.17%
Social Sciences	1.88%
Fine and Creative Arts	
Culinary Arts	.31%
Fine Arts	4.52%
General Academic	4.52%
Graduate Student	.07%
Honors	23.71%
International and Global Programs	
International/Global	4.76%
Language	.54%
Multicultural	.85%
Leisure	
General Leisure	.12%
Local Exploration	.10%
Outdoor Recreation	.72%

(continues)

Table 5.1

Distribution of L/L Participants Among Different Types of Programs, continued

Political Interest	.14%
Research	.09%
Residential College	4.29%
ROTC	.23%
Transition	
Career/Major Exploration	1.93%
For 1st Year Students	10.01%
For Diverse Populations	.12%
For Transfer Students	.13%
Umbrella	4.51%
Upper Division	.76%
Wellness	
Spirituality & Faith-Based	.01%
Wellness & Healthy Living	2.33%
Women's	
Leadership	1.15%
In STEM	2.53%
Total	100.00%

Identifying students by racial/ethnic identification. As noted earlier, NSLLP respondents were able to identify affiliation with one or more of following six racial/ethnic categories: (a) African American/Black, (b) Asian or Pacific Islander, (c) American Indian or Alaskan Native, (d) Hispanic/Latino, (e) White/Caucasian, and (f) Race/ethnicity not included above. To create a single race/ethnicity variable, I recoded any response from a student who selected only one race/ethnicity category (with the exception of "not included above," due to its lack of interpretability) into his or her self-identified monoracial/monoethnic category. I recoded any response from a student who selected any combination of two or more categories (including "not included," if another valid race/ethnicity was chosen) into a "multiracial/multiethnic" category.

The results of recoding appear below, in Table 5.2. Chi-square goodness-of-fit tests were conducted on valid responses to determine whether the observed distribution of students by race/ethnicity between the L/L and TRH sample was consistent with its expected distribution, and standardized residuals (R) requested to identify where disproportionalities existed, if present. The chi-square was statistically significant, $\chi^2(5, 22182) = 80.877$, $p \leq .000$). Residual analysis suggested that African American students' underrepresentation in the L/L sample (R = -4.7) was a contributor to the chi-square test's statistical significance.

Table 5.2

Racial/Ethnic Composition of L/L and TRH Samples

Race		L/L Sample	TRH Sample
Valid Responses			
	African American/Black	548 (4.8%)	746 (7.0%)
	Asian or Pacific Islander	991 (8.6%)	740 (6.9%)
	American Indian or Alaskan Native	30 (0.3%)	27 (0.3%)
	Hispanic/Latino	425 (3.7%)	443 (4.1%)
	White	8738 (76.2%)	8200 (76.5%)
	Multiracial/multiethnic	730 (6.4%)	564 (5.3%)
	Total	11462 (100%)	10720 (100%)
Invalid Responses			
	User-missing	60 (35.7%)	75 (44.4%)
	Race/ethnicity not included	108 (64.3%)	94 (55.6%)
	Total	168	169

Identifying students by SES. To develop a single composite variable for SES, I drew upon information from three variables: (a) mothers' educational attainment (1=high school to 6=doctorate or professional degree), (b) fathers' educational attainment (1=high school to 6=doctorate or professional degree), and (c) family income (1=less than $25,000 to 9=more than $200,000). To accommodate students from households headed by single parents, I used the higher of mothers' or fathers' educational attainment to form the first half of the composite. Then, so that parental attainment would carry a similar weight as family income when the two were combined, I multiplied educational attainment scores by 1.5. Finally, I added family income to the rescaled educational attainment variable. Because imputing either SES component seemed inappropriate and likely to introduce error into EI calculations, missing values on either half of the component resulted in listwise deletion at this phase.

The results of the construction of the SES composite are detailed in Table 5.3 below. I conducted robust independent samples t-test to determine whether a statistically significant mean difference existed between the mean SES composite of L/L participants when compared to their TRH peers. No such difference was found, $t(20892.93) = 1.64$, $p = .10$ (two-tailed). After computing SES composites for each student, I proceeded to assign students to one of three SES groups, based upon terciles identified by SPSS. While both the L/L and TRH sample shared the same score (9.00) for the 33rd percentile, there was a slight difference in score (12.50 versus 12.00) marking the 67th percentile. Knowing that the practical consequence of the decision would be to shift 710 L/L students (approximately 3% of the sample) from the middle-SES to the highest-SES group by moving their cut-off from 12.50 to 12.00, groups were constructed such that there was: (a) a lowest SES group, $1 \leq x \leq 9$, (b) a middle SES group, $9 < x \leq 12$, and (c) a highest SES group, $12 < x \leq 18$.

Table 5.3

SES Composition of L/L and TRH Samples

Parameter Estimate	L/L	TRH
Valid *n*	10947	10159
Missing *n*	683	730
Minimum $\leq x \leq$ Maximum	$1 \leq x \leq 18$	$1 \leq x \leq 18$
Mean	10.54	10.45
Std. Error of Mean	.04	.04
Mode	10.00	10.00
Std. Deviation	3.70	3.80
33rd Percentile	9.00	9.00
67th Percentile	12.50	12.00
Skewness[a]	-.072	-.089

Note.
[a] Skewness is the extent to which the mass of the distribution is shifted right (negative) or left (positive).

The distribution of SES by race/ethnicity. After creating SES and race/ethnicity groups, the distribution of SES by race/ethnicity was explored.. Table 5.4, below, suggests that certain racial/ethnic groups—most notably African American, American Indian, and Latino—are disproportionately represented in the bottom SES group, while students from other racial/ethnic groups are more normally distributed. To determine whether the observed distribution of SES by race deviated from expected proportions in a statistically significant manner, a chi-square test of independence was calculated. The test was statistically significant, $\chi^2(12, 21064) = 930.577$, $p \leq .001$, and residual analysis suggested that the high frequency of African American (R=252.1) and Latino (R=178.8) students in the lowest SES group were meaningful contributors to the significance of the omnibus test.

Table 5.4

Distribution of SES by Race/Ethnicity

Racial/ethnic group	Lowest Third	Middle Third	Highest Third
African American	64.4%	21.2%	14.4%
Asian Pacific Islander	44.4%	26.5%	29.1%
American Indian	65.4%	25.0%	9.6%
Latino	65.0%	31.4%	15.5%
White	31.9%	32.0%	36.1%
Other	46.1%	23.0%	30.9%
Multiracial/multiethnic	40.0%	29.5%	30.6%

Evaluation program missingness by race/ethnicity and SES. Given the unexpectedly large proportion of missing program identifier data (18.3%), it seemed prudent to explore whether race/ethnicity and SES—key analytic variables of interest—were in some way related to that missingness, even if there were no obvious ways to rectify a potential problem. To do so, chi-square tests of independence were conducted, with an indicator of missingness serving as the dependent variable and race/ethnicity and SES group variables serving as independent variables.

Chi-square tests indicated that missingness was not independent of race/ethnicity, $\chi^2(6, 2107) = 57.12$, $p \leq .001$. Residual analysis indicated that African American L/L participants were disproportionately likely to not report the program in which they were enrolled (R=5.24) while White participants were disproportionately likely to have reported their program (R=-6.85). Similarly, missingness was not found to be independent of SES group, $\chi^2(2, 1927) = 16.75$, $p \leq .001$. Residual analysis indicated that students from the low SES group were disproportionately represented among those students who did not provide information about program membership (R=4.08).

The potential bias introduced by non-random missingness on the program membership variable cannot be quantified. However, because each group was analyzed separately, it is important to note that such missingness is only important if it is non-random *within* a student group. For example, if African American students in civic engagement programs are just as likely to not report their program of enrollment as are African American students in sustainability programs, the EI calculations are unaffected. If there is a connection between program type and non-reporting, EI calculations would be biased. In any event, one thing can be said with certainty: missingness on the program membership variable resulted in elimination of subsequent analyses, reducing statistical power.

The computation of race/ethnicity EIs. I evaluated the status of participatory equity of students from different racial/ethnic backgrounds between different thematic types of L/L programs by computing a series of equity indices, as described by Hao (2002). Before doing so, I had to consider the nature of the bias introduced by the approximately 18% of L/L participants who, for whatever reason, failed to report the L/L program in which they were enrolled. The instinctive response, list-wise deletion, seemed to be a mistake: doing so would have altered the denominator of the EI fraction (the proportion of students from a certain racial/ethnic group in the total L/L population), a known value that was independent of students' failure to provide program information. Were any adjustment to have been made, it would have had to have been to the numerator of the EI, which reflected a group's prevalence in a given thematic type. However, there was no basis in the data with which to impute a students' program membership, short of the participation rates that were already known which, if used, would have only reproduced the existing EIs. Ultimately, I concluded that no adjustment was better than an obviously flawed adjustment, accepting the missing values as a limitation of the study. The resulting analyses appear below, in Table 5.5.

Table 5.5

Equity Indices of Thematic Types of L/L Programs by Race/Ethnicity

Thematic Type	African Amer.	Amer. Indian	Asian/ Pac. Isl.	Latino	White	Multi-racial
Civic Engagement						
Civic Engagement	1.609	.000	**.445**	**.519**	1.085	**.604**
Environmental	1.696	.000	**.469**	1.093	.957	1.697
Leadership	1.297	.000	**.403**	1.359	1.058	**.730**
Service/Social Justice	1.101	.000	1.739	1.622	.888	.944
Disciplinary						
Agriculture/Vet. Med.	.000	4.603	**.139**	.000	1.138	1.703
Business	1.112	.000	**.769**	.806	1.055	**.730**
Comm./Journalism	.842	.000	**.388**	**.724**	1.153	**.316**
Education	**.768**	3.505	**.212**	**.619**	1.161	**.432**
Engineering & C.S.	.923	2.408	.826	1.643	.989	.990
General Sciences	**.404**	.000	**.695**	**.521**	1.129	**.640**
Health Sciences	1.203	.000	1.146	**.689**	1.031	**.502**
Humanities	**.475**	.000	**.789**	1.022	1.014	1.546
Interdisciplinary	.000	8.129	**.492**	**.574**	.949	3.007
Law/Criminal Justice	1.793	.000	.000	**.771**	1.049	1.346
Mathematics	3.922	.000	**.723**	1.686	**.738**	1.963
Social Sciences	**.239**	.000	**.330**	**.462**	1.207	**.359**
Fine and Creative Arts						
Culinary Arts	1.494	.000	1.239	.000	1.031	**.561**
Fine Arts	.944	.908	1.044	1.089	.985	1.119
General Academic	1.239	.000	**.658**	1.150	1.016	1.042
Graduate Student	.000	.000	.000	8.990	**.656**	2.617
Honors	**.410**	**.348**	**.727**	**.651**	1.095	.908
International/Global Programs						
International/Global	.817	.878	2.925	1.302	**.778**	1.011
Language	.820	.000	1.134	2.115	.900	1.539
Multicultural	8.313	.000	1.631	2.766	**.320**	1.812
Leisure						
General Leisure	.000	.000	2.313	.000	1.049	.000
Local Exploration	.000	.000	1.285	2.997	.874	1.745
Outdoor Recreation	**.615**	.000	**.680**	**.397**	1.157	**.231**
Political Interest	.000	.000	.890	.000	1.110	1.208
Research	2.615	.000	1.446	3.371	.820	.000
Residential College	**.417**	.000	1.356	**.336**	1.040	.901
ROTC	.996	18.194	.000	.000	1.124	**.748**

Note. Bolded values are below .800, and represent potential participatory inequity.

Table 5.5

Equity Indices of Thematic Types of L/L Programs by Race/Ethnicity, continued

Thematic Type	African Amer.	Amer. Indian	Asian/ Pac. Isl.	Latino	White	Multi-racial
Transition						
Career/Major Exploration	1.162	2.123	1.221	1.948	.911	1.047
For 1st Year Students	1.408	1.225	1.124	1.239	.928	1.241
For Diverse Populations	.000	243.133	1.051	.000	**.119**	2.855
For Transfer Students	1.743	.000	.000	2.247	1.093	.000
Umbrella	.849	.000	1.132	1.287	.999	.824
Upper Division	.896	.000	1.157	**.385**	1.049	**.673**
Wellness						
Spirituality & Faith-Based	.000	.000	.000	.000	1.312	.000
Wellness/Healthy Living	**.581**	1.769	1.017	**.250**	1.063	.945
Women's						
In STEM	1.341	1.633	1.285	1.268	.908	1.275
Leadership	2.565	3.604	**.436**	**.254**	.990	1.037

Note. Bolded values are below .800, and represent potential participatory inequity.

Before interpreting Table 5.5, a caution: The relatively large number of thematic types, when combined with the relatively small number of students who identified as non-White, meant that the *n* underlying many cells was quite small. Indeed, of the 252 total cells in the analysis, 52 had no participants and 73 represented fewer than five. To minimize the risk of overstatement, I excluded cells with individual *n*s less than five from further discussion. Unfortunately, this precluded meaningful discussion about American Indian students (*n*=30).

As can be seen in Table 5.5, EIs indicating the presence of participatory inequity in different thematic types of L/L programs were found across all racial/ethnic groups. Non-zero cells below .800, Bensimon et al.'s (2003) lower boundary for inequity, are bolded. Students who identified as African American were underrepresented in general science (.404), honors (.410), residential college (.417), general wellness (.581) and education (.768) programs. Students who identified as Asian or Pacific Islander were underrepresented in leadership (.403), general academic (.658), general science (.695), honors (.727), business (.769), and humanities (.789) programs. Students who identified as Latino were underrepresented in general science (.521), honors (.651), and health science (.689) programs. Students who were identified as White were underrepresented in domestic multicultural (.320), mathematics (.738), and international/global (.778) programs. Finally, students who were identified as Multiracial/Multiethnic were underrepresented in education (.432), health science (.502), general science (.630), business (.730), and leadership (.730) programs.

The computation of SES EIs. After computing EIs by race/ethnicity, I moved on to explore participatory equity in different thematic types of L/L programs by students' SES. Table 5.6 lists the EIs for each SES group in each program type. Because L/L participants were disaggregated into only three categories in this analysis, compared to the six-category race/ethnicity-based analysis, fewer cells were blank (*n*=3) or represented EIs based on fewer than five students (*n*=16). Non-zero cells below .800, Bensimon et al.'s (2003) lower boundary for inequity, are bolded. As a reminder, although all EIs are presented in the table below, only those results based on cells with five or more participants are discussed below.

Table 5.6

Equity Indices of Thematic Types of L/L Programs by SES Group

Thematic Type	Lowest SES	Middle SES	Highest SES
Civic Engagement			
Civic Engagement	.936	1.057	1.017
Environmental	1.008	1.070	.927
Leadership	.856	1.079	1.085
Service/Social Justice	1.211	1.004	.766
Disciplinary			
Agriculture/Vet. Med.	1.511	.725	**.693**
Business	1.154	.829	.987
Comm./Journalism	**.779**	.934	1.300
Education	1.448	.987	**.525**
Engineering & C.S.	1.052	1.017	.928
General Sciences	.905	.990	1.112
Health Sciences	1.199	.927	.849
Humanities	**.722**	1.002	1.300
Interdisciplinary	0.857	.881	1.264
Law/Criminal Justice	1.377	**.583**	.968
Mathematics	1.721	1.033	**.187**
Social Sciences	**.787**	.865	1.354
Fine and Creative Arts			
Fine Arts	1.025	.949	1.019
Culinary Arts	**.612**	.857	1.552
General Academic	1.176	.872	.925
Graduate Student	1.836	**.551**	**.499**
Honors	**.724**	1.164	1.152
International/Global Programs			
International/Global	1.058	1.071	.873
Language	1.124	1.079	**.794**
Multicultural	1.716	**.558**	.622
Leisure			
General Leisure	**.751**	**.300**	1.904
Local Exploration	.000	1.101	1.995
Outdoor Recreation	1.085	1.101	.816
Political Interest	.918	**551**	1.496
Research	1.377	1.239	**.374**
Residential College	**.719**	1.035	1.274
ROTC	.826	1.487	.748

Note. Bolded values are below .800, and represent potential participatory inequity.

Table 5.6

Equity Indices of Thematic Types of L/L Programs by SES Group, continued

Thematic Type	Lowest SES	Middle SES	Highest SES
Transition			
Career/Major Exploration	1.024	1.018	.957
For 1st-Year Students	1.216	.912	.846
For Diverse Populations	1.752	.901	**.272**
For Transfer Students	**.688**	1.101	1.247
Umbrella	.962	1.071	.977
Upper Division	1.556	.814	**.564**
Wellness			
Spirituality & Faith-Based	.000	.000	2.993
Wellness/Healthy Living	1.029	1.027	.944
Women's			
In STEM	.943	1.011	1.052
Leadership	1.029	1.173	.811

Note. Bolded values are below .800, and represent potential participatory inequity.

Students in the lowest SES group were underrepresented in culinary arts (.612), residential college (.719), humanities (.722), honors (.724), communication and journalism (.779), and social science programs (.787). Students in the middle SES group were underrepresented in political interest (.551), domestic multicultural (.558), and law and criminal justice (.583) programs. Finally, students in the highest SES group were underrepresented in education (.525), upper division (.564), agriculture (.693), and language (.794) programs.

The identification of thematic types for further analyses. Identifying the group × type combinations for further analysis in this study's third research question was driven by two factors: (a) the extent to which a group of students was underrepresented in a given thematic type, and (b) sample size available for analysis. As noted in Chapter 3, the number of variables in the hypothesized model demanded a minimum sample size of 180 cases per group × type combination. Although the study as a whole involved more than 22,000 participants, sub-setting students by race/ethnicity (or SES) and by specific thematic type of program quickly caused sample sizes to fall to levels well below 180 participants. As a result, an alternative approach that preserved the original intent behind the research question was needed.

Inspiration for a solution came from the discovery, noted earlier, that programs of differing thematic types often shared similar intended outcomes. The rationale for a type-level analysis was the belief that programs with different themes were likely to focus on distinct outcomes, thereby making comparison of dissimilar types suspect. However, if it was possible to identify clusters of program types that shared similar outcomes, then the amalgamation of programs of different themes would less problematic. After reviewing the types of programs in which students from different racial/ethnic or SES backgrounds were underrepresented and commonalities amongst those programs' chief outcomes, three comparison groups were identified, with the hope that, after data cleaning, sample sizes would remain sufficiently robust for analysis. The underrepresented student groups and their associated programs and outcomes included:

- Asian/Pacific Islander students participating in all disciplinary, general academic, honors programs, with a shared outcome of a smooth academic transition;
- White students participating in all international/global programs, with an outcome of diversity appreciation; and,
- Low-SES students participating in honors programs, with an outcome of a smooth academic transition.

Unfortunately, no tenable combination of program types and their accompanying outcomes were possible for African American, Latino, or Multiracial/multiethnic students, or students from middle or upper SES brackets: where students were underrepresented, the number of participants was simply too low. For example, grouping all underrepresented African American students, irrespective of common outcome, would have yielded only 81 students, comprised of: (a) Education programs, with eight participants, (b) General science programs, with nine participants, (c) Humanities programs, with three participants, (d) Social science programs, with two participants, (e) Honors programs, with 43 participants, (f) Outdoor recreation programs, with two participants, (g) Residential Colleges, with eight participants, and (h) Wellness/healthy living programs, with six participants. Ultimately, I attempted to devise a solution that, to the extent possible, balanced the spirit of the research question with the available data.

Description of selected outcome factors and quasi-pretests. The three group × type combinations identified for study involved two student learning outcomes, students' smooth academic transition to college and their appreciation for diversity.

The "smooth academic transition" factor was indicated by three items, and included students' ratings of agreement (six points ranging from "very difficult" to "very easy") with the following questions:

- *Students' smooth academic transition to college (*prior *α=.762)*
    - "Communicating with instructors outside of class." (ACADT1)
    - "Ease with seeking assistance or personal help when needed." (ACADT2)
    - "Ease with forming study groups." (ACADT3)

Importantly, the academic transition outcome was relevant only to student's first year in college. To maintain internal validity, it was critical to ensure that L/L respondents had, indeed, participated in a L/L program during that first year. Because the NSLLP dataset only included current-year living arrangement data, it was necessary to restrict analyses related to academic transition to first-year L/L participants only. For the sake of comparability, TRH samples were similarly constrained.

The "diversity appreciation" factor was indicated by three items, and included students' ratings of agreement (four points ranging from "strongly disagree" to "strongly agree") with the following statements:

- *Students' diversity appreciation (*prior *α=.820)*
    - "Since coming to college, I have learned a great deal about other racial/ethnic groups." (DIVAPPRC1)
    - "Since coming to college, I have become aware of the complexities of inter-group understanding." (DIVAPPRC2)
    - "I have gained a greater commitment to my racial/ethnic identity since coming to college." (DIVAPPRC3)

Because this outcome was not bounded in time as had been the case for students' academic transition, no additional sample constraints were necessary.

Each factor was paired with a quasi-pretest designed to assess students' pre-college motivation to attain the student learning outcome to which it was related. Both were measured on a four point scale (from "not at all important" to "very important"), and used the question stem: "Thinking back to before you started college, please rate how important you imagined these aspects of college would be?" For the academic transition factor, the aspect of college students were asked to consider was "communicating with instructors outside of class," and for the openness to diversity factor, the aspect of college students were asked to consider was "learning about people from backgrounds other than your own." In each case, the aspect to be considered for the quasi-pretest was parallel to each latent factor's first indicator.

Chapter Conclusion

In this chapter, results pertaining to this study's second research question were detailed. Using a thematic typology of L/L programs, I sought to identify whether specific populations of students, defined by race/ethnicity and SES, were potentially underrepresented in certain types of L/L programs. To do so, I computed Equity Indices by race/ethnicity and SES for each thematic type of L/L program. As a result of those calculations, a variety of potential inequities—including those suggesting both potential under- and over-representations—were noted. The low absolute number of students in some group × L/L type combinations placed limitations on subsequent analyses that sought to determine whether under-representation actually inhibited students' attainment of important learning outcomes. Three combinations were identified for further review as part of the study's third research question: (a) Asian/Pacific Islander students participating in honors L/L programs, (b) White students participating in international/global L/L programs, and (c) Low-SES students of all races/ethnicities participating in honors L/L programs.

In the next chapter, I present the results of statistical tests designed to quantify the outcomes attained by students in the three group × type combinations listed above and comparing them to their peers living in the TRH environment. In so doing, the practical consequence of under-representation can be better understood.

Chapter 6 : Results of Evaluating Differences in Student Outcomes

The Third Instrumental Research Question

This study's final instrumental research question investigated whether students underrepresented in particular types of L/L programs accrued differential benefits associated with their participation, when compared to their peers living in traditional residence hall environments. Specifically, it asked: Holding other student characteristics constant, did students underrepresented in a given thematic type (or types) of L/L program accrue differential benefits by participating in L/L programs when compared to their peers living in traditional residence hall environments?

The Model

The model evaluated here is depicted in Figure 3.1. Based on Astin's (1991) IEO model and Pascarella and Terenzini's (1980) Structural Mediation Model, the heart of the model is six key living and learning environments believed to influence student outcomes in residential settings: (a) academic/vocational conversations with peers, (b) socio-cultural conversations with peers, (c) course-related faculty interaction, (d) non-course-related faculty interaction, (e) socially supportive residence hall climates, and (f) academically supportive residence hall climates. These environments are influenced by students' pre-college characteristics, including a proxy-pretest, all of which in turn are believed to be related to students' eventual attainment of important student learning outcomes. Variables included in the model *a priori* are listed in Table 3.1. As described below, some of those variables were removed from the model through the process of confirmatory factor analysis.

Addressing the Third Instrumental Research Question

The first step in analyzing the third instrumental research question was the preparation of data files corresponding to each of the planned comparisons identified in Chapter 5. Each comparison had two data files: one that contained all of the L/L students in a given group × type combination, and another that contained all TRH students from that same group. To control, in part, for the effect of institution of attendance, only those TRH students attending institutions represented in the L/L file were retained for analysis. Then, as detailed in Chapter 3, data were screened to remove cases that had missing indicators of the dependent variable or were otherwise beyond the scope of this study.

As a result, six files were created. For Asian/Pacific Islander students, the L/L data file contained 188 first-year students and the matching TRH data file contained 306 first-year students. For White students, the L/L data file contained 204 students of all class levels and the matching TRH data file contained 3273 students of all class levels. Finally, for low-SES students, the L/L data file contained 332 first-year students, and the TRH data file contained 1535 first-year students.

Treatment of missing data. Before export to LISREL for multiple imputation, SPSS was used to examine unimputed files' patterns of missingness. Three parameters were of particular interest for each file, as they provide context for the extent of the imputation to follow: (a) Little's test of the MCAR assumption, which, if significant, suggests that data are not missing completely at random and indicates a need to model potential sources of missingness through the use of variables not included in Figure 3.1 but that might be related to item non-response (b) the variable in Figure 3.1 with the greatest missingness, and (c) the auxiliary variable with the greatest missingess. Next, data was imported into LISREL for multiple imputation. As described in Chapter 3, ten imputed datasets were created from each original unimputed file. The results of pre-imputation testing and the imputation process, appears below in Table 6.1. As a reminder, model variables are those that are either covariates or are used to measure key living and learning environments or the outcome of interest. Auxiliary variables are those included in the multiple imputation process to improve the quality of estimates for item-level non-response.

Table 6.1

Results of Multiple Imputation

File	Little's MCAR test	Model variable with greatest missingness	Auxiliary variable with greatest missingness	Overall missing
Asian/Pacific Islander Students				
L/L	$\chi^2(2902) = 2863.87$, $p > .05$	SES = 7.0%	SAT = 31.9%	3.62%
TRH	$\chi2(4921) = 5037.71$, $p > .05$	SES = 5.2%	SES = 36.8%	3.29%
White Students				
L/L	$\chi^2(1297) = 1247.29$, $p > .05$	ACADH1 = 2.5%	SAT = 22.1%	1.25%
TRH	$\chi^2(11249) = 11233.71$, $p > .05$	SES = 3.5%	SAT = 18.3%	1.33%
Low-SES Students				
L/L	$\chi^2(3643) = 3965.33$, $p \leq .001$	See list [a] = 1.5%	SAT = 11.3%	1.83%
TRH	$\chi^2(14252) = 14961.52$, $p \leq .001$	See list [b] = 4.6%	SAT = 27.2%	3.57%

Note.
[a] Four indicator variables were missing in 1.5% of cases: SOCP1, SOCP2, ACADH3, and ACADH4.
[b] Two indicator variables were missing in 4.6% of cases: ACADH3 and SOCH2.

As can be seen in Table 6.1, Little's MCAR test was not statistically significant in the Asian/Pacific Islander or White samples, indicating that there was no discernable, systematic pattern of missingness in the data and, technically, imputation could proceed without the use of auxiliary variables. Because I could not reject the null hypothesis that missingness in the low-SES samples was systematic, auxiliary variables were necessary to improve the quality of the imputation process. It should be noted, however, that missingness in all samples was quite low, ranging between 1.5% and 7.0% for model variables. Among auxiliary variables, SAT was the most frequently missing, ranging between 18.3% to 31.9%. Across all variables (model and auxiliary), overall missingness was below 4% in all data files. After imputation, analysis proceeded with the first step in LM Modeling, separate-group confirmatory factor analysis (CFA).

Single group CFA. The goal of confirmatory factor analysis was, for both L/L and TRH samples, to independently confirm the factor structure of the constructs representing each group's outcome variable and the six key living and learning environments hypothesized to influence it. To do so, CFAs were run on each multiply-imputed data set for each group × type combination, for a total of 60 analyses (10 replicates × 3 student groups × L/L or TRH). The process for the Asian Pacific American models is described below. The process of White and low-SES students was substantively identical, but its discussion is omitted here for the sake of brevity. Details on CFA modeling for all groups is outlined in Tables 6.2 through 6.6.

L/L and TRH models for Asian Pacific American students went through five cycles of estimation and re-specification. The initial factor structure suggested by Inkelas, Brower et al. (2008), hereafter referred to as the "baseline" model, exhibited good data-measurement model fit for the L/L model, with an average SRMR of .073 ($SRMR_{min}$ = .072, $SRMR_{max}$ = .074) and an average RMSEA of .045 ($CI90_{Lowest}$ = .033, $CI90_{Highest}$ = .056), both below Hu and Bentler's (1999) joint fit criteria of .09 and .06, respectively. So, too, did the baseline TRH model, which had an average SRMR of .072 ($SRMR_{min}$

.071, $SRMR_{max}$ = .074) and an average RMSEA of .051 ($CI90_{Lowest}$ = .044, $CI90_{Highest}$ = .057). Table 6.2 lists the baseline fit statistics for each of the three models I analyzed.

A review of EQS-suggested model modifications, however, identified variables in both the L/L and TRH models that did not cleanly load on the factors originally posited by Inkelas, Brower et al. (2008). As described in Chapter 3, each model was separately fitted, and variables that exhibited either cross-loading or single-factor loadings below 0.6 were considered for removal, unless doing so would reduce the number of indicators on a factor to two or fewer. Once both models appeared both statistically (i.e., by virtue of modification indices) and theoretically (i.e., by virtue of face validity) consistent, variables present in one model but absent in the other were removed so that subsequent comparisons were based upon models that were as similar as possible.

As a result of that process, a total of four factor indicator variables were removed from the L/L and TRH models of Asian Pacific American students, all due to cross-loading. Removed items included: (a) SOCP5 ("discussions with students whose personal values were different") which also loaded on the academic-vocational peer conversations factor (average MI in L/L sample = 40.59, average MI in TRH sample = 15.46), (b) CRSEF4 ("worked on a research project with instructor") which also loaded on the non-course-related faculty mentorship factor (MI_{LL} = 32.08, MI_{TRH} = 104.09), (c) SOCH3 ("intellectually stimulating environment"), which also loaded on the academically supportive residence hall climate factor (MI_{LL} = 14.00, MI_{TRH} = 18.88) and (d) SOCH7 ("peer academic support"), which also loaded on the academically supportive residence hall climate factor (MI_{LL} = 72.50, MI_{TRH} = 15.19). Table 6.3 documents the items removed from the Asian/Pacific Islander model, while Tables 6.4 and 6.5 document items removed from the White and Low-SES models, respectively.

At the conclusion of modeling, the L/L model for Asian Pacific American students had an average SRMR of 0.065 ($SRMR_{min}$ = .064, $SRMR_{max}$ = .066), an improvement of .008, and an average RMSEA of .042 ($CI90_{Lowest}$ = .027, $CI90_{Highest}$ = .055), also an improvement of .003. In the TRH model, the average SRMR was .061 ($SRMR_{min}$ = .060, $SRMR_{max}$ = .062), an improvement of .011, and the average RMSEA was .040 ($CI90_{Lowest}$ = .029, $CI90_{Highest}$ = .050), an improvement of .011. As a result, both were judged to evidence good data-model fit, based upon Hu and Bentler's (1999) joint fit criteria. Table 6.6 documents the final data-measurement model fit of all three models analyzed.

Table 6.2

Baseline Fit Statistics for Single-Group CFAs

Model	Satorra-Bentler χ^2 [a]	Average SRMR [b]	Average RMSEA [c]	90% RMSEA Confidence Interval [d]
Asian/Pacific Islander Students				
L/L	$\chi^2(356, 188) = 490.665$	.073	.045	.033, .056
TRH	$\chi^2(356, 306) = 640.950$	.072	.051	.044, .057
White Students				
L/L	$\chi^2(356. 204) = 596.953$	.080	.058	.048, .067
TRH	$\chi^2(356, 3272) = 3331.010$	.051	.051	.049, .052
Low-SES Students				
L/L	$\chi^2(356, 332) = 572.549$	.070	.043	.035, .051
TRH	$\chi^2(356, 1535) = 1605.161$	.053	.048	.045, .051

Notes.

[a] All chi-squares were significant at the p < .05 level, however, due to sample size, were not taken as indicators of poor fit. Value represents the average chi-square across all multiply imputed datasets.

[b] Value represents the average SRMR across all multiply imputed datasets.

[c] Value represents the average RMSEA across all multiply imputed datasets.

[d] Confidence interval constructed by taking the lowest lower bound and the highest upper bound across all multiply imputed datasets.

Table 6.3

Items Eliminated in Single-Group CFAs, Asian/Pacific Islander Model

			Average Modification Indices	
Model	Item Number: Text	Original × Cross-Loading Factor	L/L	TRH
Asian/Pacific Islander Students	CRSEF4: “worked on a research project with instructor”	Course-related × Non-course-related faculty interaction	32.08	104.09
	SOCH7: “peer academic support”	Socially × academically supportive residence hall	72.50	15.19
	SOCH3: “intellectually stimulating environment”	Socially × academically supportive residence hall	14.00	18.88
	SOCP5: “discussions where personal values were different”	Socio-cultural × academic peer conversations	40.59	15.46

Table 6.4

Items Eliminated in Single-Group CFAs, White Model

Model	Item Number: Text	Original × Cross-Loading Factor	Average Modification Indices L/L	Average Modification Indices TRH
White students	CRSEF4: "worked on research project with instructor"	Course-related × Non-course-related faculty interaction	18.75	541.99
	SOCH7: "peer academic support"	Socially × academically supportive residence hall	63.42	530.56
	SOCH3: "intellectually stimulating environment"	Socially × academically supportive residence hall	15.84	514.21
	SOCP5: "discussion where personal values were different"	Socio-cultural × academic peer conversations	35.09	497.32
	SOCH5: "would recommend this residence hall"	Socially × academically supportive residence hall	19.93	321.73

Table 6.5

Items Eliminated in Single-Group CFAs, Low-SES Model

Model	Item Text	Original × Cross-Loading Factor	Average Modification Indices	
			L/L	TRH
Low-SES students	CRSEF4: "worked on research project with instructor"	Course-related × Non-course-related faculty interaction	30.79	249.46
	SOCP5: "discussion where personal values were different"	Socio-cultural × academic peer conversations	16.90	169.03
	SOCH7: "peer academic support"	Socially × academically supportive residence hall	17.21	192.34
	SOCH3: "intellectually stimulating environment"	Socially × academically supportive residence hall	20.50	184.44

Table 6.6

Final Fit Statistics for Single-Group CFAs

Model	Average Satorra-Bentler χ^2 [a]	Average SRMR [b]	Δ SRMR	Average RMSEA [c]	Δ RMSEA	90% RMSEA Confidence Interval [d]
Asian/Pacific Islander Students						
L/L	$\chi^2(254, 188) = 337.80$	.065	-.008	.042	-.003	.027, .055
TRH	$\chi^2(254, 306) = 377.02$	.061	-.011	.040	-.011	.029, .050
White Students						
L/L	$\chi^2(231, 204) = 357.458$	.072	-.008	.052	-.006	.039, .064
TRH	$\chi^2(231, 3272) = 1379.008$	.043	-.008	.039	-.012	.037, .041
Low-SES Students						
L/L	$\chi^2(254, 332) = 378.424$	.064	-.006	.038	-.005	.028, .049
TRH	$\chi^2(254, 1535) = 821.042$	.045	-.008	.038	-.010	.035, .041

Notes.

[a] All chi-squares were significant at the $p < .05$ level. Due to the number of parameters to be estimated and sample size, significant chi-squares were not taken as indicators of poor fit. Value represents the average chi-square across all multiply imputed datasets.

[b] Value represents the average SRMR across all multiply imputed datasets.

[c] Value represents the average RMSEA across all multiply imputed datasets.

[d] Confidence interval constructed by taking the lowest lower bound and the highest upper bound across all multiply imputed datasets.

As noted in Tables 6.2 through 6.6, the CFA process was quite similar across all three models. Initially, baseline data-model fit for all three models was acceptable even before modification, using the standards established by Hu and Bentler (1999). Despite that, modification indices revealed several items—four of which were common across all three models—that did not load cleanly on a single factor.

A review of the cross-loading items for face validity suggested possible reasons for their lack of clarity. Two items appeared potentially misclassified from the start, at least for this sample. For example, although Inkelas, Brower et al. (2008) had associated items related to "peer academic support" and an "intellectually stimulating environment" with a *socially* supportive residence hall, their affinity for the latent factor representing an *academically* supportive hall suggested that such an association may have been inappropriate. Other items may simply have not been able to discriminate sufficiently between two nuanced contexts. "[Having] worked on a research project with a professor" may have been seen by respondents as course-related, if the research project they envisioned had been a class assignment. However, had the respondent read the question as asking whether he or she had worked on a professor's scholarly research, the association would have been non-course-related. Similarly, "discussions where personal values were different" might have occurred for some students within the context of socio-cultural issues, for others within the context of academic-vocational issues, and for still others both.

After removing the problematic items identified in Tables 6.3 through 6.5, the models were estimated a final time. Each model evidenced good fit, using Hu and Bentler's (1999) joint fit criteria. A slight improvement in data-model fit was seen in all models as a result of model modification, with declines in both SRMR and RMSEA averaging -.008 (see Table 6.6). Table 6.7 summarizes the items chosen to represent each key living and learning environment. Those items that have been struck through were removed during the CFA process. Now confident that each model was a tenable representation of the observed data, I prepared for multi-group analysis by simultaneously analyzing unconstrained L/L and TRH models.

Multiple group CFA (without constraints). As described in Chapter 3, the next phase of analysis involved analyzing each student group's L/L and TRH models simultaneously, but without the constraint that each item-factor loading in the models was identical. This step verifies that subsequent constrained multi-group analyses, necessary to identify where constraints are inappropriate and groups evidence statistically significant differences, are worthwhile: if simultaneously analyzed models do not evidence satisfactory data-model fit before the imposition of constraints, it is unlikely they would do so once constraints were put in place.

Using EQS, each group's now-modified L/L and TRH models were analyzed across all multiply imputed datasets. Simultaneous analysis of the L/L and TRH models for Asian Pacific American students yielded good data-model fit across all multiply imputed datasets, with an average SRMR of .063 ($SRMR_{min}$ = .062, $SRMR_{max}$ = .064) and an average RMSEA of .041 ($CI90_{Lowest}$ = .032, $CI90_{Highest}$ = .049). Table 6.8 summarizes results for all groups. Because unconstrained models evidenced good data-model fit for Asian Pacific American, White, and Low-SES students, I proceeded to the next phase of analysis.

Table 6.7

Summary of Key Living and Learning Environments

Latent Factor Representing Environment Variable
Peers: Socio-cultural conversations
Discussions with students whose political opinions were very different
Discussions with students whose religious beliefs were very different
Discussed major social issues, such as peace, human rights, and social justice
Discussed your views about multiculturalism and diversity
~~Discussions with students whose personal values were very different~~
Peers: Academic-vocational
Shared your concerns about classes or assignments
Discussed something learned in class
Talked about current news events
Faculty: Course-related
Visited informally [with a faculty member] before or after class
Made an appointment to meet in [a faculty member's] office
Asked [a faculty member] for information related to a course you were taking
~~Worked on a research project [with a faculty member]~~
Faculty: Non-Course-related
Discussed personal problems or concerns [with a faculty member]
Discussed academic problems or concerns [with a faculty member]
Visited informally [with a faculty member] during a social occasion
Hall: Academic Climate
My residence hall clearly supports my academic achievement
Most students in my residence environment (RE) study a lot
It's easy for students to form study groups in my RE
Staff in my RE spend a great deal of time helping students succeed academically
Hall: Social Climate
Students in my RE are concerned with helping and supporting one another
I find that students in my RE have an appreciation for people from different religious backgrounds
~~Life in my RE is intellectually stimulating~~
I find that students in my RE have an appreciation for people from different races or ethnic groups
I would recommend this RE to a friend[1]
I see students with different backgrounds having a lot of interaction in my residence hall
~~I have enough peer support in my RE to succeed academically~~

Notes.
Struck-through items removed due to modeling process.
[1] *Item not included in model for White students.*

Table 6.8

Fit Statistics for Multiple Group, Unconstrained CFA Model

Model	Satorra-Bentler χ^2 [a]	Average SRMR [b]	Average RMSEA [c]	90% RMSEA Confidence Interval [d]
Asian/Pacific Islander Students	$\chi^2(508, 494) = 716.947$	.063	.041	.032, .049
White Students	$\chi^2(462, 3477) = 1894.26$	.060	.042	.040, .044
Low-SES Students	$\chi^2(508, 1867) = 1232.672$	.056	.039	.036, .042

Notes.
[a] All chi-squares were significant at the $p < .05$ level, however, due to sample size, were not taken as indicators of poor fit. Value represents the average chi-square across all multiply imputed datasets.
[b] Value represents the average SRMR across all multiply imputed datasets.
[c] Value represents the average RMSEA across all multiply imputed datasets.
[d] Confidence interval constructed by taking the lowest lower bound and the highest upper bound across all multiply imputed datasets.

Multiple group CFA (with constraints). In the final phase of testing each group's measurement models, the L/L and TRH models were estimated with the constraint that each item-factor loading, factor variance, and factor covariance was equal across models. This process, establishing the measurement model's invariance for both the L/L and TRH sample, was iterative. First, fully constrained models were estimated. Then, non-invariant constraints were identified and released in a serial fashion until LaGrange Multiplier tests indicated those constraints that remained were tenable. After the last non-invariant constraint was released, data-model fit statistics were checked to ensure good fit, as defined by Hu and Bentler (1999).

Baseline fit for the fully constrained models for Asian/Pacific Islander students was good, as described by Hu and Bentler (1999), but the imposition of constraints did cause both SRMR and RMSEA to rise. As noted in Table 6.9, the average SRMR across all multiply imputed datasets was .079 ($SRMR_{min}$ = .078, $SRMR_{max}$ = .081), and the average RMSEA was .040 ($CI90_{Lowest}$ = .031, $CI90_{Highest}$ = .048).

Through an iterative process of testing and relaxing constraints, three problematic between-model constraints were identified (see Table 6.10). Two were factor covariances and the third was an item-factor loading. By reviewing EQS output, it was possible to determine the true parameter values associated with each relationship, once they were allowed to freely vary between the L/L and TRH models. To do so, the observed covariances and loadings were combined across each of the multiply imputed data sets, yielding single point estimates. Then, 95% confidence intervals were constructed around each point estimate to gain a sense of its likely variability in the population.

The first non-invariant parameter was the relationship between students' perception that their residence hall was academically supportive and the frequency with which they had conversations with their peers that focused on academic and vocational issues. As noted in Table 6.10, the true covariance of these constructs (often denoted ϕ, or phi) in the TRH population was estimated to fall between .035 and .236 with 95% confidence. Its combined point estimate was found to be $\phi = .136$. However, in the L/L population, the 95% confidence interval ranged from -.066 to .126. Because this confidence interval included zero, it was impossible to conclude with 95% surety that a relationship between an academically supportive hall and academic/vocational conversations existed. As such, although a point estimate of $\phi = .030$ is noted, the reader should assume this relationship is not statistically significant. As such, it is marked by a dagger (†).

The second non-invariant parameter noted through the process of testing constraints was the relationship between students' perceptions that their residence hall was academically supportive and their course-related faculty interaction. In this case, however, the 95% confidence interval for both the L/L and the TRH model included zero (see Table 6.10). As such, although point estimates of $\phi = .017$ and $\phi = .108$ were noted for the L/L and TRH populations, respectively, the reader should assume no statistically significant relationship between these two latent constructs.

The third non-invariant parameter uncovered through the process of testing constraints was the item-factor loading of ACADT3 ("ease with forming study groups"). As noted in Table 6.10, the point estimate for the unstandardized loading (often denoted λ, or lambda) in the L/L population was $\lambda = .726$ and, in the TRH population, $\lambda = .964$.

The item's standardized loadings in each group are also presented in Table 6.10, along with a 95% confidence interval in the standardized metric. The statistically significant difference between the two loadings suggests that, at least in this sample, ACADT3 exhibited differential item functioning in the L/L and TRH models.

After releasing all non-invariant constraints in the models for Asian/Pacific Islander students, the process above was repeated for the models for White students and Low-SES students. Table 6.9 contains information related to baseline data-model fit for all three models, before any non-invariant constraints were released. Tables 6.10, 6.11, and 6.12 summarize the process of releasing non-invariant constraints for the Asian/Pacific Islander, White, and Low-SES models, respectively. Finally, Table 6.13 summarizes data-model fit for all three models at the conclusion of this stage of analysis.

Table 6.9

Baseline Fit Statistics for Multiple Group, Constrained CFA Models

Model	Satorra-Bentler χ^2 [a]	Average SRMR [b]	Average RMSEA [c]	90% RMSEA Confidence Interval [d]
Asian/Pacific Islander Students	$\chi^2(554, 494) = 769.605$	.079	.040	.031, .048
White Students	$\chi^2(507, 3477) = 1949.752$	.089	.040	.038, .043
Low-SES Students	$\chi^2(554, 1867) = 1277.401$	.067	.037	.034, .041

Notes.
[a] All chi-squares were significant at the $p < .05$ level, however, due to sample size, were not taken as indicators of poor fit. Value represents the average chi-square across all multiply imputed datasets.
[b] Value represents the average SRMR across all multiply imputed datasets.
[c] Value represents the average RMSEA across all multiply imputed datasets.
[d] Confidence interval constructed by taking the lowest lower bound and the highest upper bound across all multiply imputed datasets.

Table 6.10

Tests of Constraints, Asian/Pacific Islander Student CFA Models

	TRH Model				L/L Model			
Constraint	Estimate	SE	95% CI		Estimate	SE	95% CI	
Academically supportive residence hall with academic-vocational conversations with peers.	.136	.051	.035, .236		.030	.049	-.066, .126	†
Academically supportive residence hall with course-related faculty interaction.	.017	.052	-.083, .119	†	.108	.060	-.010, .225	†
Loading of ACADT3 on ACADTRAN	.726[a]	.100	.530, .922		.964	.100	.768, 1.160	
	.518[b]		.379, .658		.688		.548, .828	

Notes.
† Relationship is not statistically significant in this group; 95% confidence interval includes zero.
[a] All loadings on this row are unstandardized.
[b]All loadings on this row are standardized, $\lambda\phi$, where $\phi = .510$ for both groups.

Table 6.11

Tests of Constraints, White Student CFA Model

	TRH Model			L/L Model		
Constraint	Estimate	SE	95% CI	Estimate	SE	95% CI
Course-related faculty interaction with academic-vocational conversations with peers	.264	.016	.233, .295	.198	.046	.108, .288
Socially supportive residence hall with frequency of socio-cultural conversations with peers	.050	.015	.020, .080	.203	.041	.123, .283
Socially supportive residence hall with frequency of course-related faculty interaction	.040	.015	.010, .070	.132	.044	.046, .219
Academically supportive residence hall with frequency of socio-cultural conversations with peers	.052	.016	.021, .084	.180	.050	.083, .277
Variance of frequency of socio-cultural conversations with peers	.727	.014	.700, .755	.665	.037	.592, .738

(continues)

Table 6.11

Tests of Constraints, White Student CFA Model (continued)

	TRH Model			L/L Model		
Constraint	Estimate	SE	95% CI	Estimate	SE	95% CI
Loading of DIV3 on	.908[a]	.023	.862, .953	.665	.077	.514, .816
DIVAPPRC	.774[b]		.740, .819	.571		.442, .701
Loading of MENTF3 on	.942[a]	.028	.877, .998	.744	.081	.586, .903
MENTF	.773[c]		.720, .819	.611		.481, .741
Loading of SOCH6 on	.980[a]	.026	.929, 1.031	1.101	.060	.982, 1.218
SOCH	.755[d]		.716, .795	.849		.757, .934
Loading of SOCP2 on	.976[a]	.013	.951, 1.00	1.088	.040	1.009, 1.167
SOCP	.832[e]		.811, .852	.887		.823, .952

Notes.

[a] All loadings on this row are unstandardized.

[b] All loadings on this row are standardized, $\lambda\phi$, where $\phi = .738$ for both groups.

[c] All loadings on this row are standardized, $\lambda\phi$, where $\phi = .674$ for both groups.

[d] All loadings on this row are standardized, $\lambda\phi$, where $\phi = .594$ for both groups.

[e] All loadings on this row are standardized, $\lambda\phi$, where $\phi = .727$ for TRH and $\phi = .665$ for L/L.

Table 6.12

Tests of Constraints, Low-SES Student CFA Model

	TRH Model			L/L Model		
Constraint	Estimate	SE	95% CI	Estimate	SE	95% CI
Course-related faculty interaction with academic-vocational peer conversations	.313	.022	.269, .357	.163	.040	.086, .240
Smooth academic transition with academic-vocational peer conversations	.201	.024	.154, .248	.096	.046	.006, .187
Smooth academic transition with socially supportive residence hall	.162	.023	.118, .207	.283	.038	.208, .357
Loading of SOCH5 on SOCH	.884[a]	.031	.822, .945	.978	.043	.894, 1.062
	.723[b]		.672, .773	.800		.731, .869
Loading of ACADT3 on ACADT	.784[a]	.032	.721, .847	.625	.068	.491, .758
	.649[c]		.597, .702	.518		.407, .628

Notes.

[a] All loadings on this row are unstandardized.

[b] All loadings on this row are standardized, $\lambda\phi$, where $\phi = .669$ for both groups.

[c] All loadings on this row are standardized, $\lambda\phi$, where $\phi = .686$ for both groups.

Table 6.13

Final Fit Statistics, Multi-Group Constrained CFA Models

Model	Satorra-Bentler χ^2 [a]	Average SRMR [b]	Average RMSEA [c]	90% RMSEA Confidence Interval [d]
Asian/Pacific Islander Students	χ^2(551, 494) = 754.535	.080	.039	.030, .047
White Students	χ^2(498, 3477) = 1944.759	.069	.041	.039, .043
Low-SES Students	χ^2(549, 1867) = 1251.288	.065	.037	.034, .041

Notes.

[a] All chi-squares were significant at the p < .05 level, however, due to sample size, were not taken as indicators of poor fit. Value represents the average chi-square across all multiply imputed datasets.

[b] Value represents the average SRMR across all multiply imputed datasets.

[c] Value represents the average RMSEA across all multiply imputed datasets.

[d] Confidence interval constructed by taking the lowest lower bound and the highest upper bound across all multiply imputed datasets.

Single group SEM. After verifying the factor structure of each student group's L/L and TRH models, I turned my attention to evaluating the structural relationships between those factors. The process was identical to that previously employed to evaluate each group's measurement models, and included: (a) analyzing each group's L/L and TRH model separately, (b) analyzing them simultaneously, without constraints, and (c) analyzing fully constrained models.

As noted in Chapter 3, the structural models to be analyzed were identical to that depicted in Figure 3.1, with two minor changes. First, all key living and learning environments were allowed to freely covary with each other so that any relationship between these environments could be modeled. Second, each covariate was associated with each latent factor, so that the totality of covariates' effects could be calculated.

I began the SEM process by separately analyzing the L/L and TRH models for Asian/Pacific Islander students. Using the standards set by Hu and Bentler (1999), both models evidenced good data-model fit (see Table 6.14). Specifically, the L/L model had an average SRMR of .052 ($SRMR_{min}$ = .051, $SRMR_{max}$ = .053) and an average RMSEA of .058 ($CI90_{Lowest}$ = .046, $CI90_{Highest}$ = .067). The TRH model had an average SRMR of .049 ($SRMR_{min}$ = .048, $SRMR_{max}$= .050) and an average RMSEA of .043 ($CI90_{Lowest}$ = .034, $CI90_{Highest}$ = .052).

Next, I reviewed model modification indices. LaGrange Multiplier tests, which in this context would have suggested the need to add a relationship between two previously unrelated constructs, indicated no additions were needed. However, Wald tests—which would have indicated the presence of relationships with little explanatory power—yielded any number of covariate-factor and covariate-covariate relationships that, in a statistical sense, added little to the model. Although dropping these relationships would have slightly improved data-model fit, doing so could also have introduced error into the calculation of covariates' total effect in the model. As a result, no model parameters were altered at this stage of testing.

Given well-fitting models for Asian/Pacific Islander students, I proceeded to analyze the models for White and Low-SES students. Table 6.14 summarizes the results of this phase of modeling. As was the case for the L/L and TRH models for Asian/Pacific Islander students, model testing for White and Low-SES yielded well-fitting models that did not require additional model modification at this stage.

Table 6.14

Baseline Fit Statistics for Single Group SEM Models

Model	Satorra-Bentler χ^2 [a]	Average SRMR [b]	Average RMSEA [c]	90% RMSEA Confidence Interval [d]
Asian/Pacific Islander Students				
L/L	$\chi^2(326, 188) = 528.647$	.052	.058	.046, .067
TRH	$\chi^2(326, 306) = 509.100$	.049	.043	.034, .052
White Students				
L/L	$\chi^2(316, 204) = 528.438$	.057	.058	.048, .067
TRH	$\chi^2(316, 3273) = 2177.229$	.032	.042	.041, .044
Low-SES Students				
L/L	$\chi^2(326, 332) = 531.913$	.051	.044	.035, .052
TRH	$\chi^2(326, 1535) = 1156.396$	.034	.041	.038, .044

Notes.

[a] All chi-squares were significant at the p < .05 level, however, due to sample size, were not taken as indicators of poor fit. Value represents the average chi-square across all multiply imputed datasets.

[b] Value represents the average SRMR across all multiply imputed datasets.

[c] Value represents the average RMSEA across all multiply imputed datasets.

[d] Confidence interval constructed by taking the lowest lower bound and the highest upper bound across all multiply imputed datasets.

Multiple group SEM (without constraints). The next phase of modeling involved simultaneously testing each student group's L/L and TRH models, but without the imposition of any constraints. As noted in Chapter 3, the purpose of doing so is to ensure that both models are sufficiently similar to merit further investigation.

Simultaneous tests of the unconstrained L/L and TRH models for Asian/Pacific Islander students evidenced good data-model fit, as defined by Hu and Bentler (1999). Specifically, the model had an average SRMR = .051 ($SRMR_{min}$ = .050, $SRMR_{max}$ = .052) and an average RMSEA = .049 ($CI90_{Lowest}$ = .042, $CI90_{Highest}$ = .055). As was the case when the models were separately analyzed, modification indices yielded no paths to add to the model, and I concluded removing paths was unnecessary given otherwise good fit. This same pattern was evidenced by the models for White students and Low-SES students. The results of these analyses are summarized in Table 6.15.

Table 6.15

Final Fit Statistics, Multi-Group Unconstrained SEM Models

Model	Satorra-Bentler χ^2 [a]	Average SRMR [b]	Average RMSEA [c]	90% RMSEA Confidence Interval [d]
Asian/Pacific Islander Students	$\chi^2(652, 494) = 1037.884$	.051	.049	.042, .055
White Students	$\chi^2(632, 3477) = 2722.309$	.046	.044	.042, .045
Low-SES Students	$\chi^2(652, 1867) = 1685.659$	.043	.041	.038, .044

Notes.

[a] All chi-squares were significant at the $p < .05$ level, however, due to sample size, were not taken as indicators of poor fit. Value represents the average chi-square across all multiply imputed datasets.

[b] Value represents the average SRMR across all multiply imputed datasets.

[c] Value represents the average RMSEA across all multiply imputed datasets.

[d] Confidence interval constructed by taking the lowest lower bound and the highest upper bound across all multiply imputed datasets.

Multiple group SEM (with constraints). In the final phase of SEM modeling, each group's L/L and TRH models were simultaneously analyzed with constraints. Specifically, cross-model constraints included: (a) item-factor constraints identified through CFA testing of the measurement model, (b) constraints holding constant the relationship between covariates and latent factors, (c) constraints holding constant the relationship among latent factors, (e) constraints holding constant the relationships between covariates, (f) constraints holding constant each factor's disturbance (i.e., its variance unexplained by its predictors), and (g) constraints holding constant the variance of each indicator's error (i.e., its variance unexplained by its latent factor).

In my previous analyses of constrained models, the analytic process was relatively simple: models were estimated, untenable constraints were released, and modified models were re-estimated. Due to a technical limitation of the LaGrange Multiplier test as implemented by EQS, all the constraints in my model could not be evaluated in this manner (K. Kim, personal communication, March 23, 2009). As a result, the process of constraint analysis was bifurcated. After using EQS to evaluate all testable constraints, I computed Satorra-Bentler chi-square difference tests for those constraints that could not be evaluated manually (see Appendix I for more details on the formulas used).

Simultaneous tests of the constrained L/L and TRH models for Asian/Pacific Islander students evidenced good data-model fit, although it was evident that the imposition of constraints had caused some fit indices to rise. The average SRMR = .074 ($SRMR_{min}$ = .073, $SRMR_{max}$ = .075) represented an increase of .023 units. The average RMSEA = .049 ($CI90_{Lowest}$ = .043, $CI90_{Highest}$ = .055) was unchanged. Baseline data-model fit statistics for this model, as well as for the models for White and Low-SES students, are presented in Table 6.16. After estimating each group's baseline fit statistics, cross-model constraints were tested to determine the extent of invariance between each group's L/L and TRH models. Due to the number of non-invariant constraints identified, specific results of these tests are not discussed in text, but are summarized in Tables 6.17, 6.18, and 6.19.

Table 6.16

Baseline Fit Statistics for Multi Group Constrained SEM Models

Model	Satorra-Bentler χ^2 [a]	Average SRMR [b]	Average RMSEA [c]	90% RMSEA Confidence Interval [d]
Asian/Pacific Islander Students	$\chi^2(758, 494) = 1203.742$	.074	.049	.043, .055
White Students	$\chi^2(742, 3477) = 2750.444$	.069	.039	.038, .041
Low-SES Students	$\chi^2(758, 1867) = 1884.532$	.064	.040	.037, .043

Notes.
[a] All chi-squares were significant at the $p < .05$ level, however, due to sample size, were not taken as indicators of poor fit. Value represents the average chi-square across all multiply imputed datasets.
[b] Value represents the average SRMR across all multiply imputed datasets.
[c] Value represents the average RMSEA across all multiply imputed datasets.
[d] Confidence interval constructed by taking the lowest lower bound and the highest upper bound across all multiply imputed datasets.

Table 6.17

Tests of Constraints, Asian/Pacific Islander SEM Model

	TRH Model			L/L Model		
Constraint	Estimate	SE	95% CI	Estimate	SE	95% CI
High school grades with SES	.377	.210	-.035, .788 †	-.377	.256	-.877, .124 †
Pre-test with Female	-.011	.025	-.059, .037 †	.131	.029	.074, .187
SES → Academically supportive residence hall	-.014	.009	-.032, .004 †	.015	.011	-.007, .037 †
High school grades → Socially supportive residence hall	.075	.035	.006, .143	-.049	.048	-.143, .045 †
Variance of SOCH1 error	.202	.023	.157, .247	.139	.022	.096, .183
Variance of SOCH2 error	.203	.027	.150, .255	.121	.021	.080, .161
Variance of SOCH5 error	.392	.044	.305, .478	.261	.034	.194, .327
Variance of ACADT3 error	1.100	.105	.894, 1.306	.861	.136	.595, 1.128
Variance of SES	18.071	1.104	15.908, 20.235	15.769	1.568	12.675, 18.864
Disturbance of academic and vocational conversations with peers	.498	.053	.395, .601	.390	.049	.293, .486

Note.† Relationship is not statistically significant in this group; 95% confidence interval includes zero.

Table 6.18

Tests of Constraints, White SEM Model

	TRH Model				L/L Model			
Constraint	Estimate	SE	95% CI		Estimate	SE	95% CI	
Loading of ACADP2 on ACADP	1.032[a] .786[b]	.023	.987, 1.078		.947 .759	.056	.837, 1.056	
Female with academic class level	-.001	.007	-.015, .013	†	-.094	.028	-.149, -.039	
Academic class level → course-related faculty interaction	.156	.015	.127, .186		-.002	.055	-.109, .106	†
Female → socially-supportive residence hall	.009	.020	-.031, .048	†	.253	.073	.110, .395	
Variance of MENTF3 error	.206	.010	.187, .226		.410	.057	.299, .522	
Variance of SOCH6 error	.297	.012	.274, .321		.196	.044	.109, .282	
Variance of DIV1 error	.227	.014	.199, .254		.161	.030	.102, .219	
Variance of DIV3 error	.330	.014	.302, .357		.614	.063	.490, .737	

Notes.
† Relationship is not statistically significant in this group; 95% confidence interval includes zero.
[a] All loadings on this row are unstandardized.
[b] Loadings on this row are averages of standardized coefficients across all imputed datasets.

Table 6.19

Tests of Constraints, Low-SES SEM Model

	TRH Model			L/L Model		
Constraint	Estimate	SE	95% CI	Estimate	SE	95% CI
Loading of MENTF2 on MENTF	1.129[a] .759[b]	.052	1.027, 1.230	.961 .704	.084	.796, 1.125
High school grades → course-related faculty mentorship	.012	.017	-.022, .046 †	.116	.046	.026, .206
Female → course-related faculty mentorship	-.167	.033	-.232, -.101	-.058	.055	-.167, .051 †
Course-related faculty mentorship → Smooth academic transition	.275	.127	.026, .524	.034	.152	-.263, .331 †
Variance of White	.210	.005	.200, .220	.183	.012	.159, .206
Variance of high school grades	.736	.028	.681, .791	.368	.038	.295, .442
Variance of MENTF1 error	.235	.016	.203, .267	.161	.028	.105, .216
Variance of ACADH4 error	.368	.016	.336, .399	.452	.043	.368, .536
Variance of SOCH5 error	.367	.021	.325, .409	.290	.033	.226, .354

Notes.

† Relationship is not statistically significant in this group; 95% confidence interval includes zero.

[a] All loadings on this row are unstandardized.

[b] Loadings on this row are averages of standardized coefficients across all imputed datasets.

After all non-invariant constraints had been released across each group's L/L and TRH models, I recomputed data-model fit statistics. The Asian/Pacific Islander model evidenced good data-model fit, with an average SRMR = .072 ($SRMR_{min}$ = .071, $SRMR_{max}$ = .073) and an average RMSEA = .046 ($CI90_{Lowest}$ = .039, $CI90_{Highest}$ = .051). Table 6.20 summarizes data-model fit statistics for the Asian/Pacific Islander student model, as well as for the models of White and Low-SES students.

Table 6.20

Final Fit Statistics for Multi Group Constrained SEM Models

Model	Satorra-Bentler χ^2 [a]	Average SRMR [b]	Average RMSEA [c]	90% RMSEA Confidence Interval [d]
Asian/Pacific Islander Students	$\chi^2(748, 494) = 1127.939$	.072	.046	.039, .051
White Students	$\chi^2(734, 3477) = 2733.997$	.064	.040	.038, .041
Low-SES Students	$\chi^2(749, 1867) = 1755.165$	.055	.038	.035, .041

Notes.

[a] All chi-squares were significant at the $p < .05$ level, however, due to sample size, were not taken as indicators of poor fit. Value represents the average chi-square across all multiply imputed datasets.

[b] Value represents the average SRMR across all multiply imputed datasets.

[c] Value represents the average RMSEA across all multiply imputed datasets.

[d] Confidence interval constructed by taking the lowest lower bound and the highest upper bound across all multiply imputed datasets.

LM modeling. After fully testing each student groups' L/L and TRH measurement and structural models and determining the extent of non-invariance between models developed for L/L and TRH students, I was ready to model the latent means that would, ultimately, answer this study's third research question. To do so, mean structures were added to the final models developed in the previous stage of analysis, as described in Chapter 3. First, the unit constant was added to measurement equation of each indicator variable and manifest covariate. Then, unit constants were added to each latent factor. The value of that constant was set to zero in the TRH model and allowing it to be freely estimated in the L/L model, making the TRH model the reference group (Byrne, 2006; Thompson & Green, 2006). Finally, new cross-model constraints of the unit constant for indicator variables and manifest covariates were imposed (Byrne, Thompson & Green).

To begin, I estimated the baseline data-model fit for each LM model, including their new unit constants and constraints. Testing indicated that the baseline data-model fit for the Asian/Pacific Islander students model was good, with an average SRMR = .072 ($SRMR_{Lowest}$ = .071, $SRMR_{Highest}$ = .074) and an average RMSEA = .046 ($CI90_{Lowest}$ = .040, $CI90_{Highest}$ = .052). Table 6.21 summarizes data-model fit statistics for this baseline model, as well as for the baseline models for White and Low-SES students.

Table 6.21

Baseline Fit Statistics for LM Models

Model	Satorra-Bentler χ^2 [a]	Average SRMR [b]	Average RMSEA [c]	90% RMSEA Confidence Interval [d]
Asian /Pacific Islander Students	$\chi^2(770, 494) = 1288.777$	.072	.046	.040, .052
White Students	$\chi^2(756, 3477) = 2817.838$	.064	.040	.038, .042
Low-SES Students	$\chi^2(770, 1867) = 2026.637$	.058	.039	.037, .042

Notes.

[a] All chi-squares were significant at the p < .05 level, however, due to sample size, were not taken as indicators of poor fit. Value represents the average chi-square across all multiply imputed datasets.

[b] Value represents the average SRMR across all multiply imputed datasets.

[c] Value represents the average RMSEA across all multiply imputed datasets, for the covariance matrix only.

[d] Confidence interval constructed by taking the lowest lower bound and the highest upper bound across all multiply imputed datasets.

After testing each baseline model, I reviewed EQS output for evidence of improperly specified equality constraints. A total of five equality constraints were released across the L/L and TRH models for Asian/Pacific Islander students, two related to covariate means and three others related to factor indicator means. First, recalling that the item is reverse coded on the RES-B, L/L participants reported higher mean high school grades κ (kappa) = 1.667 (CI_{95}: 1.547 ≤ κ ≤ 1.786), than their TRH peers, κ = 1.971 (CI_{95}: 1.876 ≤ κ ≤ 2.065). Similarly, L/L participants reported higher mean SES, κ = 10.573 (CI_{95}: 9.989 ≤ κ ≤ 11.157) than did non-participants, κ = 9.115 (CI_{95}: 8.624 ≤ κ ≤ 9.605).

Mean differences on factor indicators were more mixed. Respondents who participated in L/L programs reported higher mean scores on SOCH5 ("I would recommend this residence hall"), κ = 2.662 (CI_{95}: 2.321 ≤ κ ≤ 3.003) than did their peers in the TRH environment, κ = 2.496 (CI_{95}: 2.147 ≤ κ ≤ 2.884). However, students in the TRH environment reported higher scores on ACADH4 ("staff helps with academics"), κ = 2.482 (CI_{95}: 2.089 ≤ κ ≤ 2.766) and on MENTF1 ("discussed personal problems or concerns with an instructor), κ = 1.275 (CI_{95}: 0.921 ≤ κ ≤ 1.629) than did their L/L peers, κ = 2,289 (CI_{95}: 1.949 ≤ κ ≤ 2.630) and κ = 1.162 (CI_{95}: 0.808 ≤ κ ≤ 1.515, respectively).

Table 6.22

Tests of Constraints, Asian/Pacific Islander LM Model

	TRH Model			L/L Model		
Constraint	Estimate	SE	95% CI	Estimate	SE	95% CI
Mean of high school grades	1.971	.048	1.876, 2.065	1.667	.061	1.547, 1.786
Mean of SES	9.115	.250	8.624, 9.605	10.573	.298	9.989, 11.157
Mean of SOCH5	2.496	.178	2.147, 2.844	2.662	.174	2.321, 3.003
Mean of ACADH4	2.428	.173	2.089, 2.766	2.289	.174	1.949, 2.630
Mean of MENTF1	1.275	.181	.921, 1.629	1.162	.180	.808, 1.515

Table 6.23

Tests of Constraints, White LM Model

	TRH Model			L/L Model		
Constraint	Estimate	SE	95% CI	Estimate	SE	95% CI
Loading of SOCH4 on SOCH	.966[a] .772[b]	.026	.915, 1.018	1.155 .832	.057	1.043, 1.267
Mean of Female	.648	.008	.633, .664	.727	.032	.664, .790
Mean of Pre-test	3.007	.015	2.978, 3.036	3.235	.059	3.119, 3.351
Mean of SOCP4	1.217	.079	1.063, 1.371	1.379	.094	1.196, 1.562
Mean of SOCH6	2.625	.059	2.510,2.741	2.820	.102	2.619, 3.021
Mean of DIV1	2.193	.071	2.053, 2.332	2.289	.086	2.120, 2.458
Mean of MENTF3	1.093	.041	1.012, 1.173	1.238	.056	1.129, 1.347

Notes.
[a] All loadings on this row are unstandardized.
[b] Loadings on this row are averages of standardized coefficients across all imputed datasets.

Table 6.24

Tests of Constraints, Low-SES LM Model

	TRH Model			L/L Model		
Constraint	Estimate	SE	95% CI	Estimate	SE	95% CI
Mean of high school grades	1.886	.022	1.843, 1.929	1.354	.033	1.289, 1.418
Mean of White	.701	.012	.677, .725	.762	.023	.717, .807
Mean of ACADH2	2.188	.083	2.026, 2.350	2.322	.089	2.147, 2.498
Mean of ACADH4	2.195	.074	2.049, 2.340	2.017	.084	1.852, 2.182
Mean of SOCP3	2.045	.113	1.824, 2.266	1.918	.118	1.687, 2.149
Mean of SOCP4	1.967	.109	1.754, 2.181	1.786	.115	1.561, 2.011
Mean of CRSEF3	1.941	.094	1.756, 2.126	1.848	.103	1.645, 2.051
Mean of MENTF3	1.255	.053	1.151, 1.358	1.325	.062	1.203, 1.446
Mean of SOCH1	2.401	.084	2.236, 2.567	2.505	.089	2.330, 2.680
Mean of SOCH5	2.470	.090	2.293, 2.646	2.652	.109	2.438, 2.866
Mean of ACADT3	2.669	.126	2.453, 2.946	2.963	.127	2.715, 3.212
Academic and vocational conversations with peers → Smooth academic transition	.192	.065	.064, .319	-.034	.105	-.238, .172 †

Notes.
† Relationship is not statistically significant in this group; 95% confidence interval includes zero.

After releasing all non-invariant constraints, final data-model fit statistics were computed in preparation for interpreting the latent means of the outcome factor. Data-model fit was good for the Asian/Pacific Islander L/L and TRH models, with an average SRMR = .072 ($SRMR_{min}$ = .071, $SRMR_{max}$ = .074), and an average RMSEA = .046 ($CI90_{Lowest}$ = .039, $CI90_{Highest}$ = .052). Fit statistics indicated that the model accounted for an average of 17.4% of variance in TRH participants' scores and 15.1% of the variance in L/L participants' scores, averaged across all multiply imputed datasets. Fit statistics for this model, as well as the models for White and Low-SES students, are summarized in Table 6.25.

Table 6.25

Final Fit Statistics for Multi Group LM Models

Model	Satorra-Bentler χ^2 [a]	Average SRMR [b]	Average RMSEA [c]	90% RMSEA Confidence Interval [d]	Percent Variance Explained in DV
Asian/Pacific Islander Students	χ^2 (765, 494) = 1241.969	.072	.046	.039, .052	TRH = 17.4% LL = 15.1%
White Students	χ^2 (749, 3477) = 2765.018	.063	.040	.038, .042	TRH = 20.1% LL = 21.8%
Low-SES Students	χ^2 (758, 1876) = 1762.575	.056	.038	.035, .041	TRH = 24.1% LL = 26.3%

Notes.
Dependent variable for Asian/Pacific Islander and Low-SES students was smooth academic transition.
Dependent variable for White students was diversity appreciation.
[a] All chi-squares were significant at the $p < .05$ level, however, due to sample size, were not taken as indicators of poor fit. Value represents the average chi-square across all multiply imputed datasets.
[b] Value represents the average SRMR across all multiply imputed datasets.
[c] Value represents the average RMSEA across all multiply imputed datasets, for the covariance matrix only.
[d] Confidence interval constructed by taking the lowest lower bound and the highest upper bound across all multiply imputed datasets.

Asian/Pacific Islander students in academic living-learning programs and their academic transition. As noted in Table 6.25, the final LM model for Asian/Pacific Islander students demonstrated acceptable fit, based upon the standards articulated by Hu and Bentler (1999), and explained 17.4% and 15.1% of the variance in TRH and L/L students' scores, respectively, on the NSLLP's measure of smooth academic transition. Unless explicitly noted, all relationships were invariant across models: because standardized values are presented in text to assist in comparing the relative predictive strength of an effect, apparent differences in parameter estimates for the TRH and L/L groups can be misleading. All results discussed below were statistically significant at $p \leq .05$ or better. A pictorial summary of statistically significant paths, along with standardized regression coefficients for each relationship, is depicted below in Figure 2.

My interpretation of modeling results began with an examination of the latent means associated with the Asian/Pacific Islander student models. As noted in Table 6.25, no statistically significant latent mean differences were noted between the L/L and TRH populations for three of the six key living and learning environments, holding covariates constant: (a) course-related faculty interaction, (b) non-course-related faculty mentorship, and (c) perception of an academically supportive residence hall environment.

However, statistically significant mean differences between Asian/Pacific Islander students who were L/L participants versus those who were living in the TRH environment were found on the three remaining key living and learning environment, adjusting for the influence of covariates. First, using TRH participants' scores as a reference point of zero, L/L participants reported more frequent conversations with their

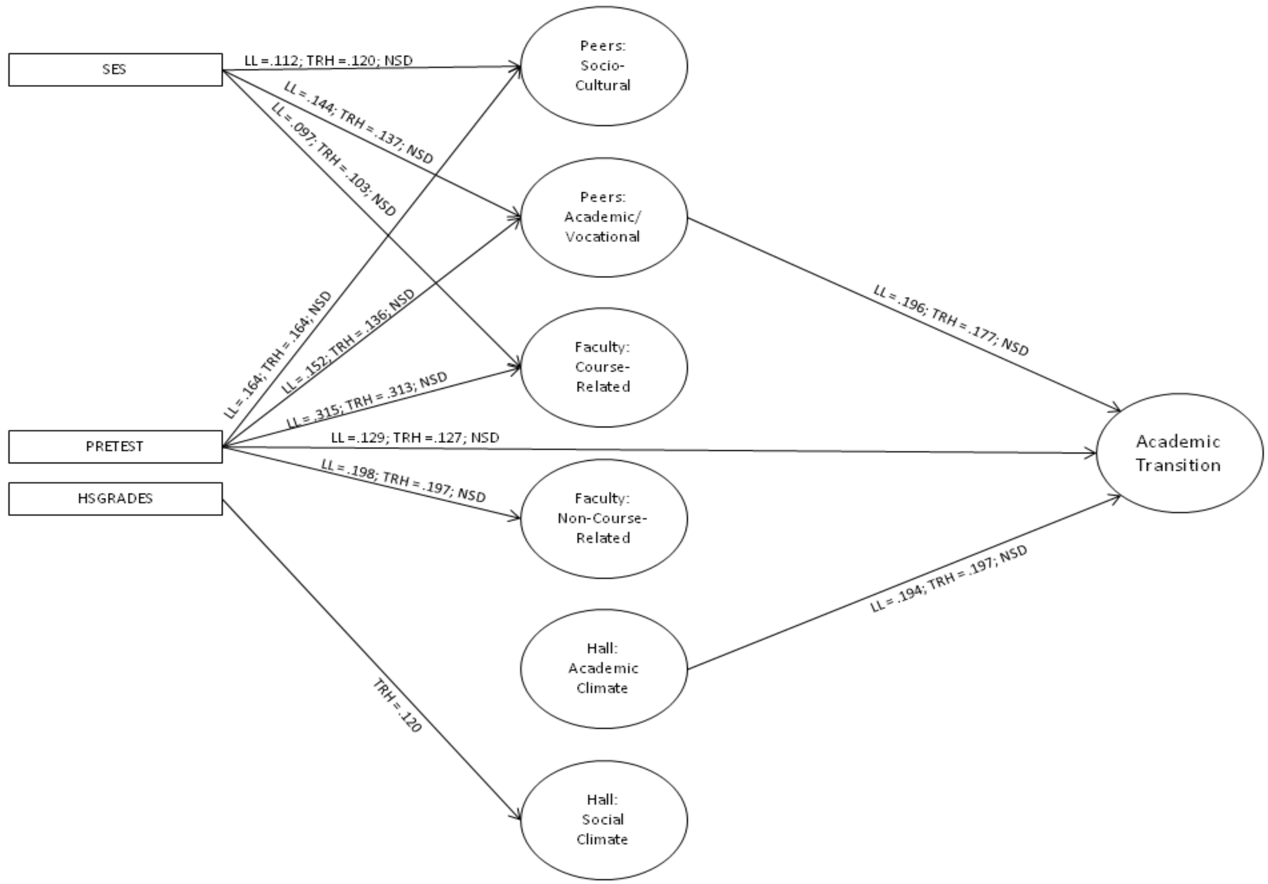


Figure 6.1

Structural Model, Asian/Pacific Islander Students

Notes.
Statistically significant indirect relationships between SES and the dependent variable for L/L students and between the pretest and the dependent variable for all students not depicted.
Statistically significant unanalyzed relationships not depicted.

peers about academic and vocational issues, κ=.220 (CI_{95}: .078 $\leq \kappa \leq$ .361). Translating kappa into a standardized effect size (ES) made it possible to give context to the difference between L/L and TRH students: using the formula offered by Thompson and Greene (2006, see Appendix J), these results suggested a .32 standard deviation difference between L/L students and their non-participating peers, a generally small effect (Cohen, 1988). Differences were also noted in the frequency with which students spoke with their peers about socio-cultural issues, with L/L students reporting a higher mean than their TRH peers, κ=.307 (CI_{95}: .148 $\leq \kappa \leq$ .467), although the effect was small (*ES* = .39). Finally, L/L participants reported that that their residence hall was more socially supportive than their peers living in the TRH environment, κ=.330 (CI_{95}: .134 $\leq \kappa \leq$ .525). Here, the effect size was medium (*ES* = .65).

After inspecting the latent means associated with the six key living and learning environments believed to affect Asian/Pacific Islander students' attainment of a smooth academic transition, the latent mean associated with that outcome was evaluated. As can be seen in Table 6.26, the difference in latent means was not statistically significant after having accounted for the influence of covariates and key living and learning environments. I conclude that, net of other elements in the model, there was no difference between the self-reported academic transitions of Asian/Pacific Islander students who participated in honors L/L programs and those of their TRH peers.

As Table 6.26 notes, however, there are statistically significant differences in the pre-college characteristics of "average" L/L student when compared to the "average" TRH student. Specifically, L/L students report higher SES status, $\Delta\kappa$ = 1.458 (*ES* = .35, or small), and better high school grades, $\Delta\kappa$ = -.304 (*ES* = .36, or small), than their TRH peers.

Other between-group differences were noted in other facets of the LM model for Asian/Pacific Islander students, most notably its measurement and structural components. Table 6.27 presents the final measurement model for each of the model's latent factors. A review of the table's first column, which shows the unstandardized loadings for each indicator variable, suggests that the factor structure that lies at the core of this analysis is identical between L/L and TRH students, save one item: ACADT3 ("ease with forming study groups") was less strongly related to a smooth academic transition for TRH students (λ = .723) than it was for L/L students (λ = 1.031). Indeed, ACADT3's loading in the TRH population, when standardized, was only λ = .528, below the widely held standard of $\lambda \geq$ .600. Why that might have been is not known—it may well be related to Arum & Roksa's (2011) recent finding that group study is unrelated to important student learning outcomes—but relevant to the present discussion, this difference meant that the latent factor representing smooth academic transition was being measured slightly differently in the TRH and L/L groups. While Byrne (2006) has argued that this "partial invariance" is not problematic so long as at least two indicators remain constrained across groups, this finding is worthy of further consideration.

Table 6.28 outlines the effect of each covariate on the latent factors representing six key living and learning environments, as well as the outcome of interest (smooth academic transition). As depicted in Figure 3.1, covariates are believed to have direct effects upon each latent factor, as well as indirect effects on the dependent variable through their action on the six key living and learning environments that are used to predict it. As such, there are two entries for smooth academic transition in the first column: the first entry, followed by a "(D)," represents covariates' direct effects on smooth academic transition, while the second, followed by an "(I)," represents their indirect effects.

Two covariates—gender (i.e., being female) and high school grades—were largely inconsequential in the model. Indeed, being female had no predictive relationship either with students' scores on any of the six key living and learning environments or their scores on the outcome of interest. High school grades were related only to TRH students' perception that their residence hall was socially supportive: each one standard deviation change in students' ratings of their high school GPAs—which, because the item is reverse scaled, represents a *decline*—was associated with a .120 standard deviation increase in their ratings of perceived social support.

In contrast, students' SES evidenced statistically significant relationships with several key model components. As noted in Table 6.28, SES was positively related to both L/L and TRH students' reported frequency of : (a) academic and vocational conversations with their peers (standardized γ = .137 and .144 in the TRH and L/L groups, respectively), (b) socio-cultural conversations with peers (standardized γ = .120 and .112 in the TRH and L/L groups, respectively), and (c) course-related faculty interaction (standardized γ = .103 and .097 in the TRH and L/L groups, respectively). It should be noted that although the standardized gammas (γ) above exhibit slight variation, the unstandardized gammas (see Table 6.28) do not. This serves as a reminder that the covariate relationship between SES and the three latent factors above was invariant across the L/L and TRH models.

Furthermore, recall that because standardized gammas (γ) are presented above, coefficients can be directly compared *within* the TRH or L/L model (although not across the two models). In both the TRH and L/L models, the causal path from SES to academic and vocational peer conversations is stronger than that from SES to socio-cultural conversations, which is, in turn, stronger than that from SES to course-related faculty interaction. Finally, while SES was found to have no direct effect upon students' smooth academic transition, it did evidence an indirect effect in the L/L model, albeit a small one in comparison to its direct effects elsewhere, $\gamma = .051$.

Students' scores on the quasi-pre-test item had statistically significant, direct effects on four of the six key living and learning environments under study, as well as statistically significant direct and indirect effects on the latent factor representing students' smooth academic transition to college (see Table 6.28

). A review of unstandardized gammas indicates that each of these relationships was invariant across the L/L and TRH models. One standard deviation increases in the quasi-pretest were associated with: (a) .136 and .152 standard deviations increases in the frequency of TRH and L/L students' frequency of academic and vocational conversations with peers, respectively, (b) .164 standard deviation increases in the frequency of TRH and L/L students' frequency of socio-cultural conversations with peers, (c) .313 and .315 standard deviation increases in TRH and L/L students' frequency of course-related faculty interaction, respectively, and (d) .197 and .198 standard deviation increases in TRH and L/L students' frequency of non-course-related faculty mentorship, respectively. Not surprisingly, the quasi-pre-test also demonstrated statistically significant direct and indirect relationships with the latent variable representing students' smooth academic transition. A one standard deviation increase in students' quasi-pre-test score was directly associated with a .127 and .129 standard deviation increase in TRH and L/L students' smooth academic transition, respectively, and indirectly related to an increase in .055 standard deviations in both groups.

Although Figure 1 posited that all six key living and learning environments would evidence a statistically significant causal relationship to students' smooth academic transition, Table 6.29 demonstrates that this was not the case for the Asian/Pacific Islander students in my sample. Indeed, only two living and learning environments were statistically significant in this analysis: (a) the frequency with which students had academic and vocational conversations with their peers, standardized $\beta = .196$ for TRH students and $\beta = .177$ for L/L students, and (b) students' perception that their hall was academically supportive, standardized $\beta = .194$ for TRH students and $\beta = .197$ for L/L students. Although slight differences existed in causal paths' standardized values, a review of Table 6.29 and unstandardized parameter estimates remind the reader that the causal paths between those environments and the outcome of interest were invariant across the L/L and TRH models. The implication of this finding for future research and practice are discussed in more detail in Chapter 5.

Finally, Table 6.30 documents the so-called "unanalyzed" relationships in the final structural model for the Asian/Pacific Islander students in my sample. This nomenclature, employed by Hancock and Mueller (2009), is meant to suggest that no causal relationship between two variables is hypothesized. As such, standardized estimates (phis, or ϕ), can be thought of as simple correlations. Initially, statistically significant relationships were noted between three pairs of covariates: (a) high school grades and being female (ϕ = -.142 in both the L/L and TRH models), suggesting that women were more likely to report better high school grades than their male counterparts, (b) high school grades and the quasi-pre-test (ϕ = -.142 in both the L/L and TRH models), suggesting that those with better high school academic performance were more likely to report high quasi-pre-test scores, and (c) among L/L participants, being female and the quasi-pretest (ϕ = .302).

Evidence of statistically significant relationships between several of the key living and learning environments were also noted (see Table 6.30). As a reminder, because these latent factors were hypothesized to be partially predicted by measured covariates, the covariances/correlations presented in Table 6.30 are technically between the disturbances of these factors (i.e., the portion of the factor remaining after removing the influence of covariates), not the latent factors themselves. They are represented by ψ, or psi.

The frequency with which students reported having academically or vocationally focused conversations with their peers was correlated with several latent factors (see Table 6.30). The relationship was invariant between the L/L and TRH models for three factors, including: (a) the frequency of socio-cultural conversations with peers, standardized ψ = .578 and .653 for TRH and L/L models, respectively, (b) the frequency of course-related faculty interaction, standardized ψ = .334 and .377 for the TRH and L/L models, respectively, and (c) the frequency of non-course-related faculty mentorship, standardized ψ = .124 and .140 for the TRH and L/L models, respectively. For the TRH model only, the frequency of students' academic and vocational conversations with their peers was related to their perception of their residence hall's academic climate, standardized ψ = .196.

In addition to its relationship with frequency of academically or vocationally oriented peer conversations, the latent factor representing students' frequency of socio-cultural conversations was correlated with four other key living and learning environments (see Table 6.30). Each was invariant between the L/L and TRH models, and included: (a) frequency of course-related faculty interaction, standardized ψ = .366 for both models, (b) frequency of non-course-related faculty mentorship, standardized ψ = .344 for both models, (c) perception that the residence hall was academically supportive, standardized ψ = .190 for both models, and (d) perception that the residence hall was socially supportive, standardized ψ = .126 for both models.

Finally, two other statistically significant latent factor correlations were of note (see Table 6.30). First, the frequency with which students reported course-related faculty interaction was correlated with the frequency with which they reported non-course-related faculty mentorship, standardized ψ = .738 in both models. Second, students' perceptions of their residence hall as academically and socially supportive were correlated, standardized ψ =.693 in both models.

Table 6.26

Final Mean Structure Model, Asian/Pacific Islander Students

	Estimates		Std. Errors		Estimates/Std. Errors		For Differences	
Effect	TRH	L/L	TRH	L/L	TRH	L/L	95% C.I.[a]	ES[b]
Unit Constant on								
ACADP	.000	.220	n/a	.071	n/d	3.092 *	.078, .361	.32
SOCP	.000	.307	n/a	.081	n/d	3.800 *	.148, .467	.39
CRSEF	.000	.123	n/a	.071	n/d	1.757	-.016, .261 †	n/a
MENTF	.000	.079	n/a	.068	n/d	1.176	-.054, .211 †	n/a
ACADH	.000	.019	n/a	.145	n/d	.140	-.265, .304 †	n/a
SOCH	.000	.330	n/a	.100	n/d	3.307 *	.134, .525	.65
ACADT	.000	.107	n/a	.106	n/d	1.012	-.101, .315 †	n/a
SES	9.115	10.573	.244	.290	37.431 *	36.458 *	.715, 2.220	.35
HSGRADES	1.971	1.667	.048	.061	41.052 *	27.320 *	-.456, -.152	.36
FEMALE	.600	.600	.022	.022	27.273 *	27.273 *	n/a	n/a
PRETEST	2.756	2.756	.039	.039	70.667 *	70.667 *	n/a	n/a

Notes.

[a] Confidence intervals for latent means are derived as described in Appendix H. For covariates, confidence intervals were created around the difference of the listed values (L/L – TRH), using the Satterthwaite approximation of the standard errors.

[b] Standardized effect size of mean difference, calculated as described in Appendix J.

† Confidence interval includes zero; not statistically significant at $\alpha = .05$ level.

Table 6.27

Final Measurement Model, Asian/Pacific Islander Students

	Estimates		Std. Errors		Estimates/Std. Errors				Standardized	
Effect	TRH	L/L	TRH	L/L	TRH		L/L		TRH	L/L
ACADP on										
ACADP1	1.000	1.000	n/a	n/a	n/d		n/d		0.794	0.757
ACADP2	1.036	1.036	0.054	0.054	19.283	*	19.283	*	0.798	0.762
ACADP3	1.082	1.082	0.056	0.056	19.389	*	19.389	*	0.792	0.755
SOCP on										
SOCP1	1.000	1.000	n/a	n/a	n/d		n/d		0.807	0.807
SOCP2	1.003	1.003	0.047	0.047	21.421	*	21.421	*	0.806	0.805
SOCP3	0.968	0.968	0.044	0.044	21.939	*	21.939	*	0.793	0.792
SOCP4	0.941	0.941	0.043	0.043	21.881	*	21.881	*	0.811	0.810
CRSEF on										
CRSEF1	1.000	1.000	n/a	n/a	n/d		n/d		0.854	0.853
CRSEF2	0.897	0.897	0.042	0.042	21.194	*	21.194	*	0.785	0.784
CRSEF3	0.820	0.820	0.045	0.045	18.267	*	18.267	*	0.657	0.655
MENTF on										
MENTF1	1.000	1.000	n/a	n/a	n/d		n/d		0.807	0.805
MENTF2	0.911	0.911	0.058	0.058	15.731	*	15.731	*	0.752	0.751
MENTF3	0.755	0.755	0.065	0.065	11.684	*	11.684	*	0.712	0.710
ACADH on										
ACADH1	1.000	1.000	n/a	n/a	n/d		n/d		0.821	0.821
ACADH2	0.961	0.961	0.049	0.049	19.814	*	19.814	*	0.758	0.759
ACADH3	0.917	0.917	0.056	0.056	16.428	*	16.428	*	0.718	0.718
ACADH4	0.828	0.828	0.061	0.061	13.485	*	13.485	*	0.670	0.671

(continues)

Table 6.27

Final Measurement Model, Asian/Pacific Islander Students (continued)

	Estimates		Std. Errors		Estimates/Std. Errors				Standardized	
Effect	TRH	L/L	TRH	L/L	TRH		L/L		TRH	L/L
SOCH on										
SOCH1	1.000	1.000	n/a	n/a	n/d		n/d		0.753	0.808
SOCH2	0.980	0.980	0.056	0.056	17.472	*	17.472	*	0.747	0.823
SOCH4	0.976	0.976	0.050	0.050	19.446	*	19.446	*	0.748	0.746
SOCH5	1.093	1.093	0.066	0.066	16.611	*	16.611	*	0.669	0.739
SOCH6	1.171	1.171	0.065	0.065	18.011	*	18.011	*	0.770	0.768
ACADT on										
ACADT1	1.000	1.000	n/a	n/a	n/d		n/d		0.731	0.726
ACADT2	0.942	0.942	0.101	0.101	9.314	*	9.314	*	0.685	0.680
ACADT3	0.723	1.031	0.102	0.112	7.089	*	9.193	*	0.528	0.703

Table 6.28

Final Structural Model—Covariates, Asian/Pacific Islander Students

	Estimates		Std. Errors		Estimates/Std. Errors				Standardized	
Effect	TRH	L/L	TRH	L/L	TRH		L/L		TRH	L/L
SES on										
ACADP	0.023	0.023	0.008	0.008	2.913	*	2.913	*	0.137	0.144
SOCP	0.023	0.023	0.009	0.009	2.522	*	2.522	*	0.120	0.112
CRSEF	0.017	0.017	0.008	0.008	2.175	*	2.175	*	0.103	0.097
MENTF	-0.002	-0.002	0.007	0.007	-0.329		-0.329		-0.017	-0.015
ACADH	-0.014	0.014	0.009	0.010	-1.522		1.416		-0.094	0.094
SOCH	-0.001	-0.001	0.006	0.006	-0.167		-0.167		-0.008	-0.008
ACADT (D)	-0.003	-0.003	0.011	0.011	-0.236		-0.236		-0.013	-0.012
ACADT (I)	0.003	0.011	0.005	0.005	0.694		2.280	*	0.016	0.051
HSGRADES on										
ACADP	0.038	0.038	0.041	0.041	0.915		0.915		0.043	0.049
SOCP	-0.002	-0.002	0.046	0.046	-0.043		-0.043		-0.002	-0.002
CRSEF	-0.001	-0.001	0.042	0.042	-0.024		-0.024		-0.001	-0.001
MENTF	-0.002	-0.002	0.037	0.037	-0.051		-0.051		-0.003	-0.003
ACADH	0.029	0.029	0.038	0.038	0.752		0.752		0.038	0.038
SOCH	0.074	-0.049	0.035	0.048	2.129	*	-1.021		0.120	-0.080
ACADT (D)	-0.044	-0.044	0.053	0.053	-0.838		-0.838		-0.041	-0.042
ACADT (I)	0.025	0.012	0.021	0.041	1.204		0.535		0.024	0.011
FEMALE on										
ACADP	0.046	0.046	0.069	0.069	0.663		0.663		0.031	0.035
SOCP	-0.047	-0.047	0.077	0.077	-0.612		-0.612		-0.029	-0.029
CRSEF	-0.056	-0.056	0.070	0.070	-0.806		-0.806		-0.039	-0.039
MENTF	-0.097	-0.097	0.063	0.063	-1.534		-1.534		-0.077	-0.077
ACADH	0.013	0.013	0.063	0.063	0.204		0.204		0.010	0.010
SOCH	-0.045	-0.045	0.052	0.052	-0.875		-0.875		-0.043	-0.044
ACADT (D)	-0.152	-0.152	0.088	0.088	-1.728		-1.728		-0.083	-0.084
ACADT (I)	0.001	0.001	0.034	0.069	0.018		0.018		0.001	0.001

(continues)

Table 6.28

Final Structural Model—Covariates, Asian/Pacific Islander Students (continued)

Effect	Estimates		Std. Errors		Estimates/Std. Errors				Standardized	
	TRH	L/L	TRH	L/L	TRH		L/L		TRH	L/L
PRETEST on										
ACADP	0.112	0.112	0.037	0.037	3.016	*	3.016	*	0.136	0.152
SOCP	0.150	0.150	0.043	0.043	3.488	*	3.488	*	0.164	0.164
CRSEF	0.253	0.253	0.040	0.040	6.315	*	6.315	*	0.313	0.315
MENTF	0.139	0.139	0.035	0.035	3.949	*	3.949	*	0.197	0.198
ACADH	-0.017	-0.017	0.037	0.037	-0.461		-0.461		-0.024	-0.024
SOCH	0.034	0.034	0.031	0.031	1.103		1.103		0.058	0.059
ACADT (D)	0.131	0.131	0.056	0.056	2.346	*	2.346	*	0.127	0.129
ACADT (I)	0.056	0.056	0.028	0.037	2.014	*	2.014	*	0.055	0.055

Table 6.29

Final Structural Model—Latent Factors, Asian/Pacific Islander Students

	Estimates		Std. Errors		Estimates/Std. Errors				Standardized	
Effect	TRH	L/L	TRH	L/L	TRH		L/L		TRH	L/L
ACADP on										
ACADT	.244	.244	.115	.115	2.126	*	2.126	*	.196	.177
SOCP on										
ACADT	.006	.006	.093	.093	.063		.063		.005	.005
CRSEF on										
ACADT	.095	.095	.150	.150	.637		.637		.075	.076
MENTF on										
ACADT	.033	.033	.166	.166	.201		.201		.023	.023
ACADH on										
ACADT	.282	.282	.133	.133	2.116	*	2.116	*	.194	.197
SOCH on										
ACADT	.114	.114	.155	.155	.734		.734		.065	.066

Table 6.30

Final Structural Model—Unanalyzed Relationships, Asian/Pacific Islander Students

	Estimates		Std. Errors		Estimates/Std. Errors				Standardized	
Effect	TRH	L/L	TRH	L/L	TRH		L/L		TRH	L/L
SES with										
HSGRADES	0.377	-0.376	0.209	0.253	1.803		-1.482		0.106	-0.113
FEMALE	-0.088	-0.088	0.090	0.090	-0.977		-0.977		-0.042	-0.045
PRETEST	0.002	0.002	0.161	0.161	0.011		0.011		0.000	0.001
HSGRADES with										
FEMALE	-0.058	-0.058	0.018	0.018	-3.239	*	-3.239	*	-0.142	-0.142
PRETEST	-0.070	-0.070	0.034	0.034	-2.071	*	-2.071	*	-0.096	-0.096
FEMALE with										
PRETEST	-0.011	0.131	0.025	0.029	-0.453		4.500	*	-0.025	0.302
ACADP with										
SOCP	0.323	0.323	0.031	0.031	10.373	*	10.373	*	0.578	0.653
CRSEF	0.159	0.159	0.024	0.024	6.577	*	6.577	*	0.334	0.377
MENTF	0.053	0.053	0.022	0.022	2.381	*	2.381	*	0.124	0.140
ACADH	0.086	-0.007	0.030	0.023	2.905	*	-0.306		0.196	-0.018
SOCH	0.039	0.039	0.021	0.021	1.865		1.865		0.106	0.120
SOCP with										
CRSEF	0.192	0.192	0.033	0.033	5.899	*	5.899	*	0.366	0.366
MENTF	0.164	0.164	0.030	0.030	5.414	*	5.414	*	0.344	0.344
ACADH	0.092	0.092	0.027	0.027	3.370	*	3.370	*	0.190	0.190
SOCH	0.051	0.051	0.023	0.023	2.190	*	2.190	*	0.126	0.126

(continues)

Table 6.30

Final Structural Model—Unanalyzed Relationships, Asian/Pacific Islander Students (continued)

	Estimates		Std. Errors		Estimates/Std. Errors		Standardized	
Effect	TRH	L/L	TRH	L/L	TRH	L/L	TRH	L/L
CRSEF with								
MENTF	0.298	0.298	0.032	0.032	9.194 *	9.194 *	0.738	0.738
ACADH	-0.005	0.049	0.027	0.032	-0.191	1.539	-0.012	0.119
SOCH	-0.024	-0.024	0.021	0.021	-1.124	-1.124	-0.069	-0.069
MENTF with								
ACADH	0.036	0.036	0.022	0.022	1.626	1.626	0.096	0.096
SOCH	-0.032	-0.032	0.019	0.019	-1.686	-1.686	-0.103	-0.103
ACADH with								
SOCH	0.218	0.218	0.025	0.025	8.712 *	8.712 *	0.693	0.693

Notes.
Because latent factors ACADP, SOCP, CRSEF, MENTF, ACADH, and SOCH were technically endogenous due to the influence of model covariates, unanalyzed relationships were between each latent factors' disturbances.

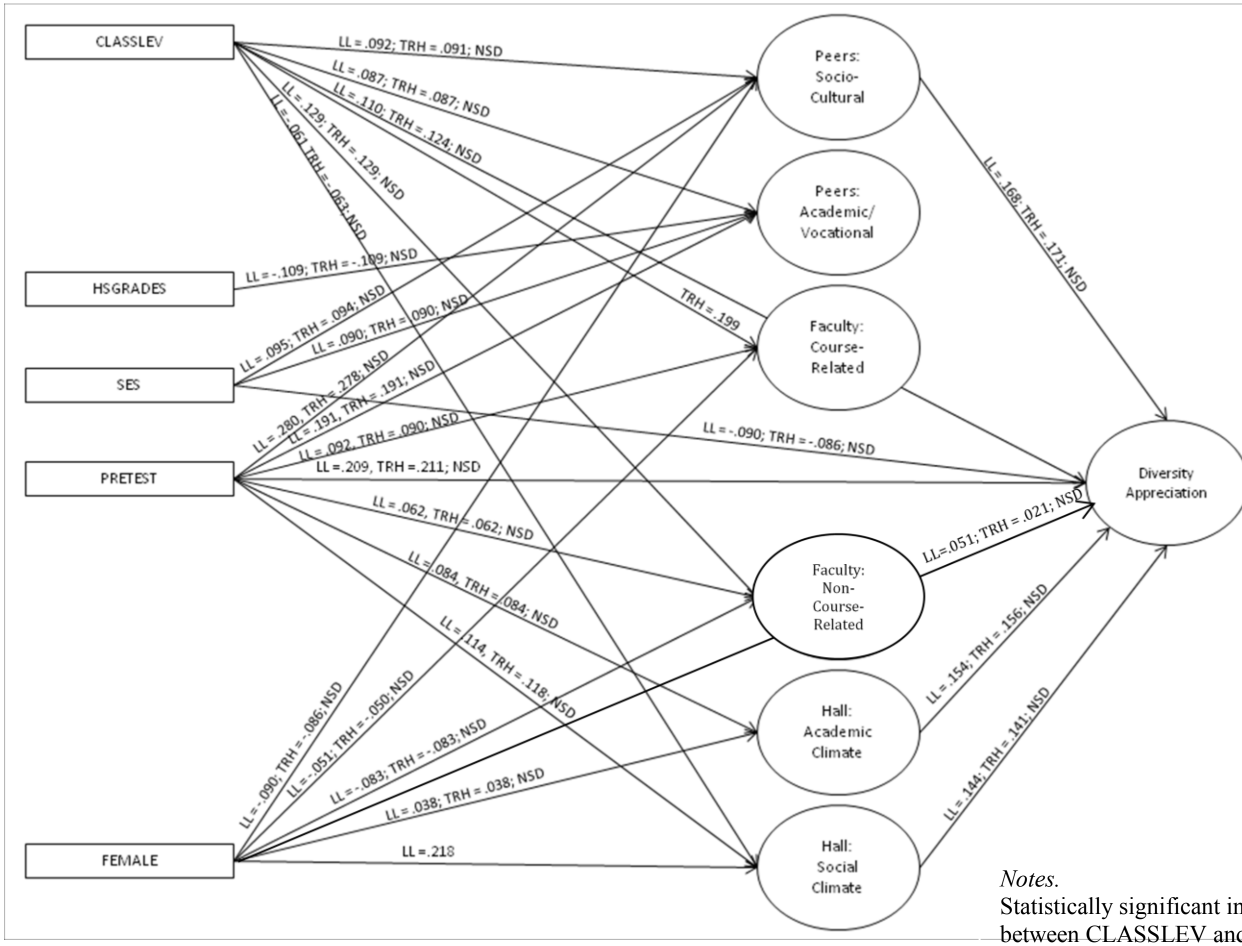
White students in international/global living-learning programs and their appreciation of diversity. As noted in Table 6.31, the final LM model for White students demonstrated acceptable fit, based upon the standards articulated by Hu and Bentler (1999), and explained 20.1% and 21.8% of the variance in TRH and L/L students' scores, respectively, on the NSLLP diversity

appreciation measure. Unless explicitly noted, all relationships were invariant across models: because standardized values are presented in text to assist in comparing the relative predictive strength of an effect, apparent differences in parameter estimates for the TRH and L/L groups can be misleading. All results discussed below were statistically significant at $p \leq .05$ or better. A pictorial summary of statistically significant paths, along with standardized regression coefficients for each relationship, is depicted below in Figure 6.2.

My interpretation of modeling results began with an examination of the latent means associated with the White student models. Statistically significant mean differences were noted on four of the six key living and learning environments, adjusting for the influence of covariates (see Table 6.31). They included: (a) the frequency of students' socio-cultural peer conversations, $\kappa = .309$ (CI_{95}: $.189 \leq \kappa \leq .4310$, with a small effect size (*ES* = .41), (b) the frequency of students' course-related faculty interactions, $\kappa = .386$ (CI_{95}: $.183 \leq \kappa \leq .589$), with a medium effect size (*ES* = .58), (c) students' perceptions of their residence halls as academically supportive, $\kappa = .242$ (CI_{95}: $.148 \leq \kappa \leq .336$), with a small effect size (*ES* = .41), and (d) students' perceptions of their residence halls as academically supportive, $\kappa = .170$ (CI_{95}: $.029 \leq \kappa \leq .310$), with a small effect size (*ES* = .34). No statistically significant mean differences were noted for the frequency of students' academic and vocational conversations with their peers, or their non-course related faculty mentorship.

Next, I turned my attention to the dependent variable for the White students' models, diversity appreciation. Adjusting for the influence of predictor variables, no statistically significant difference was noted between L/L and TRH students' mean scores on the outcome of interest (see Table 6.31). As such, I concluded that, on average, the White students in my sample accrued no benefit vis-à-vis diversity appreciation by virtue of their L/L participation, at least above and beyond that experienced by their TRH peers. Invariance testing of the L/L and TRH models for White students suggested important ways in which these models may have differed, ultimately contributing to the results noted above. The final measurement model for White students is summarized in Table 6.32. As can be seen there, five indicator variables exhibited differential functioning between the TRH and L/L models. They included: (a) ACADP2 ("discussed something learned in class"), which was found to have unstandardized loadings of $\lambda = 1.028$ and 1.019 in the TRH and L/L models, respectively, (b) SOCP2 ("held discussions with those with different religious beliefs"), which was found to have unstandardized loadings of $\lambda = .982$ and 1.017 in the TRH and L/L models, respectively, (c) SOCH4 ("appreciate different race/ethnicities"), which was found to have unstandardized loadings of $\lambda = .966$ and 1.155 in the TRH and L/L models, respectively, (d) MENTF3 ("visited informally with instructor on social occasion"), which was found to have unstandardized loadings of $\lambda = .730$ and .696 in the TRH and L/L models, respectively, and (e) DIV3 ("greater commitment to own identity"), which was found to have unstandardized loadings of $\lambda = .878$ and .659 in the TRH and L/L models, respectively. While all statistically significant, between-model differences in the latter two indicators may have had additional practical significance: when standardized, the loadings for MENTF3 and DIV3 in the L/L group were $\lambda = .466$ and .469, both well below the conventional .600 threshold identified by DeVellis (2003) as the minimally acceptable loading for a latent factor. The consequences for this finding are discussed in more detail in Chapter 7.

The influence of covariates on key living and learning environments and the dependent variable is summarized in Table 6.33. Each was consequential, albeit to varying degrees. Students' high school grades (reverse coded) were associated with the fewest number of latent factors, only bearing a statistically significant relationship with L/L and TRH students' frequency of academic and vocational conversations with their peers, standardized $\gamma = -0.109$ in both groups. Students' SES was found to be related to two key living and learning environments, along with the dependent variable. While it was positively related to both TRH and L/L students' frequency of academic/vocational and socio-cultural conversations with peers, standardized $\gamma = .090$ and .090 and standardized $\gamma = .094$ and .095, respectively, it had a direct *negative* relationship on TRH and L/L students' appreciation for diversity, standardized $\gamma = -.086$ and -.090, respectively. (Note that this direct effect was partially, but not completely, offset in the TRH and L/L groups by SES's indirect action through other key living and learning environments, standardized $\gamma = 0.012$ in both groups.)



Notes.
Statistically significant indirect relationships between CLASSLEV and DIVAPPRC for TRH students, SES to DIVAPPRC for both groups, and FEMALE to DIVAPPRC for TRH students not depicted.
Statistically significant unanalyzed relationships not depicted.

Figure 6.2

Structural Model, White Students

Academic class level, included in the TRH and L/L models to account for possible maturational effects, was statistically significantly related to both key living and learning environments and the dependent variable (see Table 6.33). Academic class level was positively and directly related to TRH and L/L students': (a) frequency of academically and vocationally-focused peer conversations, standardized $\gamma = .087$ in both groups, (b) frequency of socio-cultural peer conversations, standardized $\gamma = .091$ and .092 in the TRH and L/L groups, respectively, (c) non-course-related faculty mentorship, standardized $\gamma = .129$ in both groups, and (d) diversity appreciation, standardized $\gamma = .124$ and .110 in the TRH and L/L groups, respectively. Class level's relationship to course-related faculty interaction was found to vary between models: while it exhibited a strongly positive, direct effect in the TRH model (standardized $\gamma = .199$), its effect in the L/L model was not statistically significant. This may explain why SES exhibited a statistically significant indirect effect on diversity appreciation in the TRH model (standardized $\gamma = .022$), but no indirect effect in the L/L model. Finally, academic class level was negatively and directly related to students' perception that their hall was socially supportive, standardized $\gamma = -.063$ and -.061 in the TRH and L/L models, respectively.

The relationships between being female and the latent factors in the TRH and L/L models were mixed (see Table 6.33). Specifically, being female was associated with negative, direct relationships with students': (a) frequency of socio-cultural peer conversations, standardized $\gamma = -.089$ and -.090 in the TRH and L/L groups, respectively, (b) frequency of course-related faculty interaction, standardized $\gamma = -.050$ and -.051 in the TRH and L/L groups, respectively, and (c) frequency of non-course-related faculty mentorship, standardized γ = -.083, in both groups. However, being female was associated with positive, direct relationships with other factors, including students': (a) perception that their residence hall was academically supportive, standardized γ = .038 in both groups, and (b) diversity appreciation, standardized γ = .021 and .051 in the TRH and L/L groups, respectively. Being female's relationship to students' perception that their hall was socially supportive was found to vary between the TRH and L/L models: while there was no statistically significant relationship among TRH students, in L/L students the relationship was strongly positive, standardized γ = .218. Finally, in the TRH environment only, being female was negatively associated with students' diversity appreciation, standardized γ = -.018.

Finally, the quasi-pretest was directly associated with all key living and learning environments in the TRH and L/L models, as well as both directly and indirectly with the dependent variables in both models (see Table 6.33). All relationships were positive, and included students': (a) reported frequency of academic/vocational peer conversations, standardized γ = .191 in both models, (b) reported frequency of socio-cultural peer conversations, standardized γ = .278 and .280 in the TRH and L/L models, respectively, (c) course-related faculty interaction, standardized γ = .090 and .092 in the TRH and L/L models, respectively, (d) non-course-related faculty mentorship, standardized γ = .062 in both groups, (e) perceptions of an academically-supportive residence hall, standardized γ = .084 in both groups, (f) perceptions of a socially-supportive residence hall, standardized γ = .118 and .114 in the TRH and L/L models, respectively, and (g) diversity appreciation, standardized γ = .211 and .209 in the TRH and L/L models, respectively.

Not surprisingly, the quasi-pretest also evidenced positive, indirect relationships to diversity appreciation, standardized γ =.082 in both groups.

Although all six key living and learning environments had been hypothesized to bear upon White students' diversity appreciation, a review of Table 6.34 indicates that such was not the case. In fact, only three of six environments were found to be statistically significantly related to the dependent latent factor, including students': (a) frequency of socio-cultural conversations with their peers, standardized β = .171 and .168 in the TRH and L/L groups, respectively, (b) perception that their residence hall was academically supportive, standardized β = .156 and .154 in the TRH and L/L groups, respectively, and (c) perception that their residence hall was socially supportive, standardized β = .141 and .144 in the TRH and L/L groups, respectively.

Non-causal relationships between covariates and latent factors are summarized in Table 6.35. Being female was associated with: (a) having a lower academic class level among L/L participants only, standardized ϕ = -.228, (b) reporting better high school grades (reverse coded), standardized ϕ = -.126 in both groups, and (c) reporting higher scores on the quasi-pre-test, standardized ϕ = .181 in both groups. Higher SES and better high school grades were similarly associated with better quasi-pretest scores, ϕ = .036 in both groups and ϕ = -.045 in both groups, respectively.

Latent factors representing all key living and learning environments evidenced statistically significant covariances among their disturbances, save the relationship between students' perception that their residence hall was socially supportive and their reported level of non-course-related faculty mentorship (see Table 6.35). Students' reported frequency of academic and vocational peer conversations were positively associated with their: (a) frequency of socio-cultural peer conversations, standardized ψ = .580 and .585 in the TRH and L/L group, respectively, (b) course-related faculty interaction, standardized ψ = .335 in the TRH model versus ψ = .278 in the L/L model, (c) non-course-related faculty mentorship, standardized ψ = .121 in both groups, (d) perception of an academically supportive residence hall climate, standardized ψ = .094 in both groups, and (e) perception of a socially supportive residence hall climate, standardized ψ = .137 in both groups.

In addition to its relationship with its academic/vocational analog, students' frequency of socio-cultural peer conversations was positively related to the remaining key living and learning environments (see Table 6.35). This included: (a) students' course-related faculty interaction, standardized ψ = .337 and .340 in the TRH and L/L groups, respectively, (b) students' non-course-related faculty mentorship, standardized ψ = .295 and .298 in the TRH and L/L groups, respectively, (c) students' perception that their residence hall was academic supportive, standardized ψ = .051 in the TRH environment versus ψ = .236 in the L/L environment, and (d) students' perception that their residence hall was socially supportive, standardized ψ = .043 in the TRH environment versus ψ = .318 in the L/L environment.

Beyond those relationships noted above, students' course-related faculty interaction was positively associated with three other key living and learning environments (see Table 6.35). They included: (a) the frequency with which students' reported non-course-related faculty mentorship, standardized ψ = .739 in both groups, (b) students' perception that their residence hall was academically supportive, standardized ψ = .101 in both groups, and (c) students' perception that their residence hall was socially supportive, standardized ψ = .071 in the TRH environment versus standardized ψ = .183 in the TRH environment.

Two other relationships were of note (see Table 6.35). First, students' non-course-related faculty mentorship was positively associated with the perception that their residence hall was academically supportive, standardized ψ = .110 in both groups. Finally, students' perception that their residence hall was academically supportive was positively associated with the perception that their residence hall was socially supportive, standardized ψ = .648 in both groups.

Table 6.31

Final Mean Structure Model, White Students

	Estimates		Std. Errors		Estimates/Std. Errors		For Differences	
Effect	TRH	L/L	TRH	L/L	TRH	L/L	95% C.I.[a]	ES[b]
Unit Constant on								
ACADP	.000	.037	n/a	.054	n/d	.689	-.069, .143 †	n/a
SOCP	.000	.309	n/a	.062	n/d	4.979 *	.189, .431	.41
CRSEF	.000	.386	n/a	.103	n/d	3.743 *	.183, .589	.58
MENTF	.000	.079	n/a	.042	n/d	1.880	-.003, .161 †	n/a
ACADH	.000	.242	n/a	.048	n/d	5.038 *	.148, .336	.41
SOCH	.000	.170	n/a	.072	n/d	2.372 *	.029, .310	.34
DIVAPPRC	.000	.088	n/a	.059	n/d	1.493	-.028, .204 †	n/a
CLASSLEV	1.634	1.634	.015	.015	108.933 *	108.933 *	n/a	n/a
SES	11.115	11.115	.061	.061	182.507 *	182.507 *	n/a	n/a
HSGRADES	1.796	1.796	.015	.015	119.733 *	119.733 *	n/a	n/a
FEMALE	.648	.727	.008	.032	81.050 *	22.713 *	.014, .144	.17
PRETEST	3.007	3.235	.015	.059	200.467 *	54.831 *	.109, .347	.27

Notes.

[a] Confidence intervals for latent means are derived as described in Appendix H. For covariates, confidence intervals were created around the difference of the listed values (L/L – TRH), using the Satterthwaite approximation of the standard errors.

[b] Standardized effect size of mean difference, calculated as described in Appendix J.

† Confidence interval includes zero; not statistically significant at $\alpha = .05$ level.

Table 6.32

Final Measurement Model, White Students

	Estimates		Std. Errors		Estimates/Std. Errors				Standardized	
Effect	TRH	L/L	TRH	L/L	TRH		L/L		TRH	L/L
ACADP on										
ACADP1	1.000	1.000	n/a	n/a	n/d		n/d		0.808	0.808
ACADP2	1.028	1.019	0.022	0.042	46.736	*	24.262	*	0.785	0.782
ACADP3	1.005	1.005	0.024	0.024	42.050	*	42.050	*	0.745	0.745
SOCP on										
SOCP1	1.000	1.000	n/a	n/a	n/d		n/d		0.805	0.803
SOCP2	0.982	1.017	0.017	0.031	57.788	*	33.124	*	0.792	0.800
SOCP3	1.015	1.015	0.017	0.017	59.712	*	59.712	*	0.825	0.823
SOCP4	0.988	0.988	0.017	0.017	58.112	*	58.112	*	0.816	0.814
CRSEF on										
CRSEF1	1.000	1.000	n/a	n/a	n/d		n/d		0.783	0.777
CRSEF2	0.768	0.768	0.022	0.022	34.909	*	34.909	*	0.683	0.676
CRSEF3	0.881	0.881	0.021	0.021	41.948	*	41.948	*	0.691	0.684
MENTF on										
MENTF1	1.000	1.000	n/a	n/a	n/d		n/d		0.714	0.715
MENTF2	1.126	1.126	0.037	0.037	30.844	*	30.844	*	0.743	0.744
MENTF3	0.730	0.696	0.035	0.131	20.849	*	5.315	*	0.613	0.466
ACADH on										
ACADH1	1.000	1.000	n/a	n/a	n/d		n/d		0.812	0.812
ACADH2	0.894	0.894	0.022	0.022	40.623	*	40.623	*	0.708	0.708
ACADH3	0.853	0.853	0.023	0.023	37.104	*	37.104	*	0.669	0.669
ACADH4	0.825	0.825	0.025	0.025	33.282	*	33.282	*	0.609	0.609

(continues)

Table 6.32

Final Measurement Model, White Students (continued)

	Estimates		Std. Errors		Estimates/Std. Errors				Standardized	
Effect	TRH	L/L	TRH	L/L	TRH		L/L		TRH	L/L
SOCH on										
SOCH1	1.000	1.000	n/a	n/a	n/d		n/d		0.723	0.735
SOCH2	1.034	1.034	0.025	0.025	41.352	*	41.352	*	0.817	0.825
SOCH4	0.966	1.155	0.026	0.057	37.162	*	20.263	*	0.772	0.832
SOCH6	1.055	1.091	0.029	0.115	36.372	*	9.515	*	0.700	0.789
DIV on										
DIV1	1.000	1.000	n/a	n/a	n/d		n/d		0.819	0.867
DIV2	0.926	0.926	0.023	0.023	40.265	*	40.265	*	0.797	0.800
DIV3	0.878	0.659	0.023	0.063	38.191	*	10.394	*	0.721	0.496

Table 6.33

Final Structural Model—Covariates, White Students

	Estimates		Std. Errors		Estimates/Std. Errors				Standardized	
Effect	TRH	L/L	TRH	L/L	TRH		L/L		TRH	L/L
CLASSLEV on										
ACADP	0.067	0.067	0.014	0.014	4.786	*	4.786	*	0.087	0.087
SOCP	0.084	0.084	0.015	0.015	5.627	*	5.627	*	0.091	0.092
CRSEF	0.157	-0.002	0.015	0.055	10.447	*	-0.038		0.199	-0.003
MENTF	0.072	0.072	0.011	0.011	6.545	*	6.545	*	0.129	0.129
ACADH	-0.024	-0.024	0.013	0.013	-1.815		-1.815		-0.035	-0.035
SOCH	-0.037	-0.037	0.011	0.011	-3.318	*	-3.318	*	-0.063	-0.061
DIVAPPRC (D)	0.080	0.080	0.014	0.014	5.721	*	5.721	*	0.124	0.110
DIVAPPRC (I)	0.017	0.007	0.006	0.008	2.850	*	0.850		0.022	0.009
SES on										
ACADP	0.017	0.017	0.003	0.003	5.387	*	5.387	*	0.090	0.090
SOCP	0.021	0.021	0.004	0.004	5.250	*	5.250	*	0.094	0.095
CRSEF	0.007	0.007	0.004	0.004	1.675		1.675		0.036	0.037
MENTF	0.001	0.001	0.003	0.003	0.433		0.433		0.010	0.010
ACADH	-0.002	-0.002	0.003	0.003	-0.700		-0.700		-0.013	-0.013
SOCH	-0.003	-0.003	0.003	0.003	-0.967		-0.967		-0.020	-0.019
DIVAPPRC (D)	-0.019	-0.019	0.003	0.003	-6.200	*	-6.200	*	-0.086	-0.090
DIVAPPRC (I)	0.002	0.002	0.001	0.001	2.100	*	2.100	*	0.012	0.012
HSGRADES on										
ACADP	-0.084	-0.084	0.014	0.014	-6.007	*	-6.007	*	-0.109	-0.109
SOCP	-0.016	-0.016	0.016	0.016	-1.019		-1.019		-0.018	-0.018
CRSEF	-0.004	-0.004	0.015	0.015	-0.233		-0.233		-0.004	-0.005
MENTF	0.014	0.014	0.011	0.011	1.282		1.282		0.026	0.025
ACADH	-0.012	-0.012	0.013	0.013	-0.931		-0.931		-0.018	-0.018
SOCH	-0.005	-0.005	0.011	0.011	-0.427		-0.427		-0.008	-0.008
DIVAPPRC (D)	0.002	0.002	0.014	0.014	0.122		0.122		0.000	0.000
DIVAPPRC (I)	-0.002	-0.002	0.006	0.006	-0.300		-0.300		-0.002	-0.002

(continues)

Table 6.33

Final Structural Model—Covariates, White Students (continued)

	Estimates		Std. Errors		Estimates/Std. Errors				Standardized	
Effect	TRH	L/L	TRH	L/L	TRH		L/L		TRH	L/L
FEMALE on										
ACADP	0.021	0.021	0.027	0.027	0.763		0.763		0.015	0.015
SOCP	-0.150	-0.150	0.030	0.030	-5.003	*	-5.003	*	-0.089	-0.090
CRSEF	-0.072	-0.072	0.028	0.028	-2.557	*	-2.557	*	-0.050	-0.051
MENTF	-0.084	-0.084	0.021	0.021	-4.019	*	-4.019	*	-0.083	-0.083
ACADH	0.047	0.047	0.024	0.024	1.975	*	1.975	*	0.038	0.038
SOCH	0.010	0.239	0.020	0.068	0.505		3.517	*	0.009	0.218
DIVAPPRC (D)	0.056	0.056	0.026	0.026	2.138	*	2.138	*	0.021	0.051
DIVAPPRC (I)	-0.026	0.018	0.011	0.019	-2.336	*	0.932		-0.018	0.012
PRETEST on										
ACADP	0.153	0.153	0.015	0.015	10.180	*	10.180	*	0.191	0.191
SOCP	0.267	0.267	0.017	0.017	15.694	*	15.694	*	0.278	0.280
CRSEF	0.073	0.073	0.016	0.016	4.556	*	4.556	*	0.090	0.092
MENTF	0.036	0.036	0.012	0.012	2.992	*	2.992	*	0.062	0.062
ACADH	0.060	0.060	0.014	0.014	4.286	*	4.286	*	0.084	0.084
SOCH	0.071	0.071	0.012	0.012	5.925	*	5.925	*	0.118	0.114
DIVAPPRC (D)	0.105	0.105	0.016	0.016	6.544	*	6.544	*	0.211	0.209
DIVAPPRC (I)	0.067	0.067	0.007	0.007	9.437	*	9.437	*	0.082	0.082

Table 6.34

Final Structural Model—Latent Factors, White Students

	Estimates		Std. Errors		Estimates/Std. Errors		Standardized	
Effect	TRH	L/L	TRH	L/L	TRH	L/L	TRH	L/L
ACADP on								
DIVAPPRC	-0.028	-0.028	0.030	0.030	-0.943	-0.943	-0.028	-0.028
SOCP on								
DIVAPPRC	0.145	0.145	0.024	0.024	6.021 *	6.021 *	0.171	0.168
CRSEF on								
DIVAPPRC	0.064	0.064	0.042	0.042	1.507	1.507	0.064	0.062
MENTF on								
DIVAPPRC	0.109	0.109	0.062	0.062	1.771	1.771	0.078	0.077
ACADH on								
DIVAPPRC	0.178	0.178	0.035	0.035	5.086 *	5.086 *	0.156	0.154
SOCH on								
DIVAPPRC	0.190	0.190	0.042	0.042	4.512 *	4.512 *	0.141	0.144

Table 6.35

Final Structural Model—Unanalyzed Relationships, White Students

	Estimates		Std. Errors		Estimates/Std. Errors				Standardized	
Effect	TRH	L/L	TRH	L/L	TRH		L/L		TRH	L/L
CLASSLEV with										
SES	-0.100	-0.100	0.052	0.052	-1.915		-1.915		-0.032	-0.032
HSGRADES	-0.014	-0.014	0.015	0.015	-0.933		-0.933		-0.019	-0.019
FEMALE	-0.001	-0.094	0.007	0.028	-0.143		-3.357	*	-0.002	-0.228
PRETEST	-0.012	-0.012	0.013	0.013	-0.915		-0.915		-0.017	-0.017
SES with										
HSGRADES	-0.001	-0.001	0.055	0.055	-0.009		-0.009		0.000	0.000
FEMALE	-0.026	-0.026	0.029	0.029	-0.883		-0.883		-0.015	-0.015
PRETEST	0.107	0.107	0.051	0.051	2.098	*	2.098	*	0.036	0.036
HSGRADES with										
FEMALE	-0.052	-0.052	0.007	0.007	-7.429	*	-7.429	*	-0.126	-0.126
PRETEST	-0.033	-0.033	0.013	0.013	-2.500	*	-2.500	*	-0.045	-0.045
FEMALE with										
PRETEST	0.072	0.072	0.007	0.007	10.286	*	10.286	*	0.181	0.181
ACADP with										
SOCP	0.286	0.286	0.011	0.011	26.018	*	26.018	*	0.580	0.585
CRSEF	0.144	0.120	0.010	0.026	14.360	*	4.600	*	0.335	0.278
MENTF	0.037	0.037	0.007	0.007	5.343	*	5.343	*	0.121	0.121
ACADH	0.036	0.036	0.008	0.008	4.432	*	4.432	*	0.094	0.094
SOCH	0.045	0.045	0.007	0.007	6.357	*	6.357	*	0.137	0.137

(continues)

Table 6.35

Final Structural Model—Unanalyzed Relationships, White Students (continued)

	Estimates		Std. Errors		Estimates/Std. Errors				Standardized	
Effect	TRH	L/L	TRH	L/L	TRH		L/L		TRH	L/L
SOCP with										
CRSEF	0.171	0.171	0.011	0.011	15.518	*	15.518	*	0.337	0.340
MENTF	0.108	0.108	0.009	0.009	11.944	*	11.944	*	0.295	0.298
ACADH	0.023	0.106	0.010	0.034	2.320	*	3.130	*	0.051	0.236
SOCH	0.016	0.120	0.008	0.026	2.013	*	4.609	*	0.043	0.318
CRSEF with										
MENTF	0.234	0.234	0.011	0.011	21.255	*	21.255	*	0.739	0.739
ACADH	0.040	0.040	0.009	0.009	4.389	*	4.389	*	0.101	0.101
SOCH	0.024	0.061	0.008	0.026	2.950	*	2.342	*	0.071	0.183
MENTF with										
ACADH	0.031	0.031	0.007	0.007	4.414	*	4.414	*	0.110	0.110
SOCH	0.011	0.011	0.006	0.006	1.783		1.783		0.045	0.045
ACADH with										
SOCH	0.193	0.193	0.010	0.010	19.250	*	19.250	*	0.648	0.648

Notes.
Because latent factors ACADP, SOCP, CRSEF, MENTF, ACADH, and SOCH were technically endogenous due to the influence of model covariates, unanalyzed relationships were between each latent factors' disturbances.

Low-SES students in honors programs and their academic transition. As noted in Table 6.36, the final LM model for Low-SES students demonstrated acceptable fit, based upon the standards articulated by Hu and Bentler (1999), and explained 24.1% and 26.3% of the variance in TRH and L/L students' scores, respectively, on the NSLLP measure of smooth academic transition. Unless explicitly noted, all relationships were invariant across models: because standardized values are presented in text to assist in comparing the relative predictive strength of an effect, apparent differences in parameter estimates for the TRH and L/L groups can be misleading. All results discussed below were statistically significant at $p \leq .05$ or better. A pictorial summary of statistically significant paths, along with standardized regression coefficients for each relationship, is depicted below in Figure 4.

My interpretation of modeling results began with an examination of the latent means associated with the Low-SES student models. As summarized in Table 6.36, statistically significant mean differences were noted on five of six key living and learning environments, including students': (a) frequency of academic/vocational peer conversations, $\kappa = .101$ (CI_{95}: $.006 \leq \kappa \leq .196$), with a small effect size ($ES = .15$), (b) frequency of socio-cultural peer conversations, $\kappa = .290$ (CI_{95}: $.129 \leq \kappa \leq .400$), with a small effect size ($ES = .37$), (c) frequency of non-course-related faculty mentorship, $\kappa = .290$ (CI_{95}: $.129 \leq \kappa \leq .450$), with a medium effect size ($ES = .60$), (d) perception of an academically supportive residence hall, $\kappa = .382$ (CI_{95}: $.296 \leq \kappa \leq .478$), with a medium effect size ($ES = .65$) and (e) perception of a socially supportive residence hall, $\kappa = .124$ (CI_{95}: $.050 \leq \kappa \leq .199$), with a small effect size ($ES = .23$). No mean difference was noted for students' frequency of course-related faculty interaction.

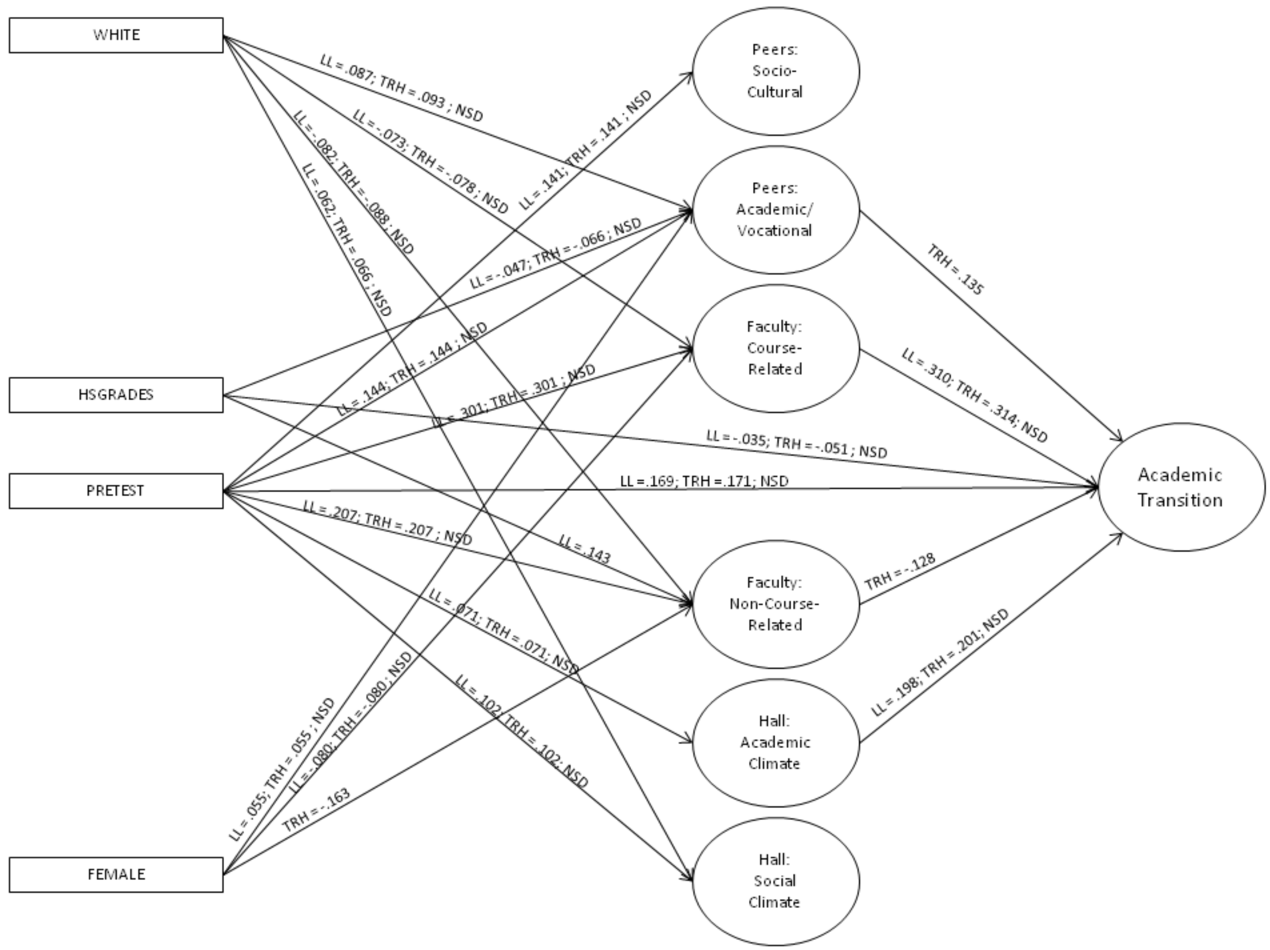


Figure 6.3

Structural model, Low-SES Students

Notes.
Statistically significant indirect relationships between PRETEST and ACADT for both groups not depicted.
Statistically significant unanalyzed relationships not depicted.

Next, I examined the mean difference between Low-SES students living in L/L and TRH environments on the outcome of interest, smooth academic transition (see Table 6.36). No statistically significant difference was noted. As such, I concluded that, on average, the Low-SES students in my sample accrued no benefit vis-à-vis smooth academic transition by virtue of their L/L participation, at least above and beyond that experienced by their TRH peers.

Invariance testing yielded a set of parameters on which the TRH and L/L models for Low-SES students differed, which may have contributed to this study's findings. First, a review of Table 6.37 indicated three factor indicators with statistically significantly different loadings across models. These included: (a) MENTF2 ("discussed career plans and ambitions with instructor"), unstandardized $\lambda = 1.127$ in the TRH model versus $\lambda = .950$ in the L/L model, (b) SOCH5 ("would recommend this hall"), unstandardized $\lambda = 1.065$ in the TRH model versus $\lambda = 1.113$ in the L/L model, (c) ACADT3 ("ease with forming study groups"), unstandardized $\lambda = .784$ in the TRH model versus $\lambda = .647$ in the L/L model.

A review of the models' structural components, and differences between the TRH and L/L groups, also suggested potentially important differences in the relationships between important covariates, key living and learning environments, and the latent factor representing smooth academic transition. The influence of student-level covariates on the latent factors depicted in Figure 6.3 is summarized in Table 6.38. Although it was wholly unrelated to perceptions of a smooth academic transition, being White was directly related to TRH and L/L students': (a) frequency of academic/vocational peer conversations, standardized $\gamma = .093$ and .087, respectively, (b) frequency of course-related faculty interaction, standardized $\gamma = -.078$ and -.073, respectively, (c) frequency of non-course related faculty mentorship, standardized $\gamma = -.088$ and -.083, respectively, and (d) perception of a socially-supportive residence hall, standardized $\gamma = .066$ and .062, respectively.

Students' high school grades had mixed effects on key living and learning environments, as well as the outcome of interest (see Table 6.38). Because the item was reversed coded, its negative relationship to students' frequency of academic/vocational peer conversations, standardized $\gamma = -.066$ and -.047 in the TRH and L/L models, respectively, suggested that better grades corresponded to more frequent interactions around these issues. Similarly, for both groups, a direct negative relationship to smooth academic transition, standardized $\gamma = -0.051$ and -0.035 in the TRH and L/L models, respectively, suggested better high school academic performance resulted in a smoother transition to college. Interestingly, better high school grades were also associated with less frequent non-course-related faculty mentorship in the L/L population, standardized $\gamma = .143$.

My findings vis-à-vis gender were similarly mixed (see Table 6.38). Being female was positively associated with the frequency of students' academic/vocational peer conversations, standardized $\gamma = .055$ in both groups. However, it was negatively associated with the frequency of students' course-related faculty interactions, standardized $\gamma = -.080$ in both groups. In the TRH model only, being female was negatively associated with non-course-related faculty interaction, standardized $\gamma = -.163$.

Finally, the quasi-pre-test was positively and directly related to each key living and learning environment, as well as directly and indirectly to the latent factor representing smooth academic transition. Specifically, the pre-test was related to students': (a) frequency of academic/vocational peer conversations, standardized $\gamma = .144$ for both groups, (b) frequency of socio-cultural peer conversations, standardized $\gamma = .141$ for both groups, (c) frequency of course-related faculty interaction, standardized $\gamma = .301$ for both groups, (d) frequency of non-course-related faculty mentorship, standardized $\gamma = .207$ for both groups, (e) perception that one's residence hall was academically supportive, standardized $\gamma = .071$ for both groups, and (f) perception that one's residence hall was socially supportive, standardized $\gamma = .102$ for both groups. The pre-test was directly, standardized $\gamma = .171$ in the TRH model and .169 in the L/L model, as well as indirectly, standardized $\gamma = .010$ in the TRH model versus .115 in the L/L model, related to the latent factor representing students' smooth academic transition to college.

Contrary to my hypothesis, not all six key living and learning environments were related to students' smooth academic transition to college (see Table 6.39). Four environments were found to be statistically significant for TRH students: (a) students' frequency of academically/vocationally focused peer conversations, standardized $\beta = .135$, (b) students' frequency of course-related faculty interaction, standardized $\beta = .314$, (c) students' frequency of non-course-related faculty mentorship, standardized $\beta = -.128$, and (d) students' perception that their residence hall was academically supportive, standardized $\beta = .201$. For L/L students, only two of the above relationships were significant. Statistically invariant in the L/L and TRH models, they included: (a) students' frequency of course-related faculty interaction, standardized $\beta = .310$, and (b) students' perception that their residence hall was academically supportive, standardized $\beta = .198$.

Exploring relationships among explicitly non-causal variables also provides insight into the phenomena shaping the experiences of low-SES students (see Table 6.40). Covariate-covariate relationships speak to typical patterns about students' pre-college characteristics. In both the TRH and L/L groups, being White was associated with: (a) better high school grades (reverse coded), standardized $\phi = -.080$ and $-.120$, respectively, and (b) lower pre-test scores, standardized $\phi = -.110$ and $-.118$, respectively. Compared to their male peers, women in both the TRH and L/L environments were: (a) more likely to report higher pre-test scores, standardized $\phi = .107$ in both groups, and (b) more likely to report better high school grades (reverse coded), standardized $\phi = -.071$ and $\phi = -.100$, respectively. Finally, better high school grades (reverse coded) were associated with higher pre-test scores in both TRH and L/L groups, standardized $\phi = -.059$ and $-.083$, respectively.

As noted in Table 6.40, statistically significant covariances were noted for most pairs of disturbances even after the influences of gender, race/ethnicity, high school grades, and pre-test scores were accounted for. TRH and L/L students' frequency of academic/vocational peer conversations were positively associated with: (a) the frequency of their socio-cultural peer conversations, standardized $\psi = .622$ in both groups, (b) the frequency of their course-related faculty interaction, standardized $\psi = .399$ in the TRH group versus standardized $\psi = .191$ in the L/L group, (c) the frequency of their non-course-related faculty mentorship, standardized $\psi = .171$ in both groups, (d) a perception of an academically supportive residence hall, standardized $\psi = .145$ in both groups, and (E) a perception of a socially supportive residence hall, standardized ψ = .182 in both groups.

Socio-cultural peer conversations also evidenced statistically significant covariances with all other key environments in the model (see Table 6.40). This included: (a) frequency of course-related faculty interaction, standardized ψ = .374 in both models, (b) frequency of non-course-related faculty mentorship, standardized ψ = .340 in both models, (c) perceptions of an academically supportive residence hall, standardized ψ = .100 in both models, and (d) perceptions of a socially supportive residence hall, standardized ψ = .113 in both models.

Not surprisingly, students' course-related faculty interaction was found to covary with their non-course-related faculty mentorship, standardized ψ = .741 in both groups. Students' frequency of course-related faculty interaction and non-course-related faculty mentorship were also found to be related to the extent to which they found their residence hall climate to be supportive of academics, standardized ψ = .126 in both groups and standardized ψ = .161 in both groups, respectively. However, faculty interaction measures had no relationship with students' perceptions of their residence hall as being socially supportive.

Finally, a statistically significant covariance was found between students' perception that their residence hall was academically supportive and their perception that it was socially supportive, standardized ψ = .737 in both groups.

Table 6.36

Final Mean Structure Model, Low-SES Students

	Estimates		Std. Errors		Estimates/Std. Errors				For Latent Means		
Effect	TRH	L/L	TRH	L/L	TRH		L/L		95% C.I.[a]		ES[b]
Unit Constant on											
ACADP	.000	.101	n/a	.048	n/d		2.100	*	.006, .196		.15
SOCP	.000	.290	n/a	.056	n/d		5.177	*	.179, .400		.37
CRSEF	.000	.006	n/a	.047	n/d		.131		-.088, .100	†	n/a
MENTF	.000	.290	n/a	.080	n/d		3.623	*	.129, .450		.60
ACADH	.000	.382	n/a	.044	n/d		8.739	*	.296, .478		.65
SOCH	.000	.124	n/a	.038	n/d		3.274	*	.050, .199		.23
ACADT	.000	.193	n/a	.103	n/d		1.871		-.010, .396	†	n/a
WHITE	.701	.762	.012	.023	58.417	*	33.126	*	.010, .112		.13
HSGRADES	1.886	1.354	.022	.033	85.736	*	41.015	*	-.610, -.454		.65
FEMALE	.678	.678	.011	.011	61.636	*	61.636	*	n/a		n/a
PRETEST	2.559	2.559	.020	.020	127.950	*	127.950	*	n/a		n/a

Notes.

[a] Confidence intervals for latent means are derived as described in Appendix H. For covariates, confidence intervals were created around the difference of the listed values (L/L – TRH), using the Satterthwaite approximation of the standard errors.

[b] Standardized effect size of mean difference, calculated as described in Appendix J.

† Confidence interval includes zero; not statistically significant at $\alpha = .05$ level.

Table 6.37

Final Measurement Model, Low-SES Students

	Estimates		Std. Errors		Estimates/Std. Errors				Standardized	
Effect	TRH	L/L	TRH	L/L	TRH		L/L		TRH	L/L
ACADP on										
ACADP1	1.000	1.000	n/a	n/a	n/d		n/d		0.821	0.820
ACADP2	0.993	0.993	0.027	0.027	36.223	*	36.223	*	0.785	0.785
ACADP3	0.969	0.969	0.031	0.031	31.471	*	31.471	*	0.728	0.727
SOCP on										
SOCP1	1.000	1.000	n/a	n/a	n/d		n/d		0.802	0.802
SOCP2	0.970	0.970	0.023	0.023	42.183	*	42.183	*	0.775	0.775
SOCP3	0.979	0.979	0.023	0.023	42.583	*	42.583	*	0.807	0.807
SOCP4	0.950	0.950	0.024	0.024	39.899	*	39.899	*	0.789	0.789
CRSEF on										
CRSEF1	1.000	1.000	n/a	n/a	n/d		n/d		0.796	0.796
CRSEF2	0.768	0.768	0.028	0.028	27.708	*	27.708	*	0.678	0.678
CRSEF3	0.899	0.899	0.029	0.029	31.338	*	31.338	*	0.695	0.695
MENTF on										
MENTF1	1.000	1.000	n/a	n/a	n/d		n/d		0.719	0.784
MENTF2	1.127	0.950	0.051	0.073	22.281	*	12.940	*	0.759	0.700
MENTF3	0.671	0.671	0.047	0.047	14.426	*	14.426	*	0.589	0.589
ACADH on										
ACADH1	1.000	1.000	n/a	n/a	n/d		n/d		0.802	0.802
ACADH2	0.900	0.900	0.027	0.027	33.315	*	33.315	*	0.708	0.708
ACADH3	0.878	0.878	0.029	0.029	30.259	*	30.259	*	0.700	0.699
ACADH4	0.782	0.782	0.032	0.032	24.520	*	24.520	*	0.608	0.569

(continues)

Table 6.37

Measurement Model, Low-SES Students (continued)

	Estimates		Std. Errors		Estimates/Std. Errors				Standardized	
Effect	TRH	L/L	TRH	L/L	TRH		L/L		TRH	L/L
SOCH on										
SOCH1	1.000	1.000	n/a	n/a	n/d		n/d		0.770	0.770
SOCH2	0.923	0.923	0.031	0.031	29.666	*	29.666	*	0.771	0.771
SOCH4	0.869	0.869	0.029	0.029	29.845	*	29.845	*	0.754	0.753
SOCH5	1.065	1.113	0.038	0.064	28.013	*	17.498	*	0.689	0.746
SOCH6	1.015	1.015	0.033	0.033	30.656	*	30.656	*	0.734	0.734
ACADT on										
ACADT1	1.000	1.000	n/a	n/a	n/d		n/d		0.814	0.818
ACADT2	0.941	0.941	0.035	0.035	26.871	*	26.871	*	0.757	0.762
ACADT3	0.784	0.647	0.036	0.064	21.781	*	10.087	*	0.625	0.557

Table 6.38

Final Structural Model—Covariates, Low-SES Students

	Estimates		Std. Errors		Estimates/Std. Errors				Standardized	
Effect	TRH	L/L	TRH	L/L	TRH		L/L		TRH	L/L
WHITE on										
ACADP	0.143	0.143	0.041	0.041	3.505	*	3.505	*	0.093	0.087
SOCP	0.001	0.001	0.045	0.045	0.018		0.018		0.001	0.001
CRSEF	-0.116	-0.116	0.039	0.039	-2.939	*	-2.939	*	-0.078	-0.073
MENTF	-0.096	-0.096	0.031	0.031	-3.106	*	-3.106	*	-0.088	-0.082
ACADH	-0.011	-0.011	0.034	0.034	-0.326		-0.326		-0.009	-0.008
SOCH	0.078	0.078	0.031	0.031	2.513	*	2.513	*	0.066	0.062
ACADT (D)	0.083	0.083	0.057	0.057	1.462		1.462		0.038	0.034
ACADT (I)	0.005	-0.063	0.028	0.034	0.176		-1.875		0.002	-0.026
HSGRADES on										
ACADP	-0.054	-0.054	0.022	0.022	-2.415	*	-2.415	*	-0.066	-0.047
SOCP	-0.009	-0.009	0.024	0.024	-0.358		-0.358		-0.009	-0.007
CRSEF	0.001	0.001	0.021	0.021	0.029		0.029		0.001	0.001
MENTF	0.009	0.118	0.017	0.045	0.527		2.615	*	0.015	0.143
ACADH	0.023	0.023	0.019	0.019	1.226		1.226		0.034	0.024
SOCH	0.015	0.015	0.017	0.017	0.894		0.894		0.024	0.017
ACADT (D)	-0.060	-0.060	0.030	0.030	-1.990	*	-1.990	*	-0.051	-0.035
ACADT (I)	-0.002	0.026	0.014	0.024	-0.123		1.082		-0.001	0.016
FEMALE on										
ACADP	0.083	0.083	0.040	0.040	2.080	*	2.080	*	0.055	0.055
SOCP	-0.075	-0.075	0.043	0.043	-1.758		-1.758		-0.044	-0.044
CRSEF	-0.116	-0.116	0.038	0.038	-3.056	*	-3.056	*	-0.080	-0.080
MENTF	-0.175	-0.061	0.033	0.055	-5.316	*	-1.115		-0.163	-0.057
ACADH	0.055	0.055	0.033	0.033	1.668		1.668		0.044	0.044
SOCH	0.032	0.032	0.030	0.030	1.060		1.060		0.027	0.027
ACADT (D)	-0.018	-0.018	0.052	0.052	-0.344		-0.344		-0.009	-0.008
ACADT (I)	0.036	-0.036	0.028	0.030	1.285		-1.199		0.017	-0.016

(continues)

Table 6.38

Final Structural Model—Covariates, Low-SES Students (continued)

	Estimates		Std. Errors		Estimates/Std. Errors				Standardized	
Effect	TRH	L/L	TRH	L/L	TRH		L/L		TRH	L/L
PRETEST on										
ACADP	0.121	0.121	0.021	0.021	5.645	*	5.645	*	0.144	0.144
SOCP	0.132	0.132	0.024	0.024	5.517	*	5.517	*	0.141	0.141
CRSEF	0.241	0.241	0.023	0.023	10.496	*	10.496	*	0.301	0.301
MENTF	0.123	0.123	0.018	0.018	7.017	*	7.017	*	0.207	0.207
ACADH	0.050	0.050	0.019	0.019	2.616	*	2.616	*	0.071	0.071
SOCH	0.065	0.065	0.018	0.018	3.669	*	3.669	*	0.102	0.102
ACADT (D)	0.204	0.204	0.034	0.034	5.997	*	5.997	*	0.171	0.169
ACADT (I)	0.121	0.140	0.018	0.024	6.700	*	5.741	*	0.101	0.115

Table 6.39

Final Structural Model—Latent Factors, Low-SES Students

	Estimates		Std. Errors		Estimates/Std. Errors				Standardized	
Effect	TRH	L/L	TRH	L/L	TRH		L/L		TRH	L/L
ACADP on										
ACADT	0.192	-0.034	0.066	0.104	2.915	*	-0.324		0.135	-0.023
SOCP on										
ACADT	-0.074	-0.074	0.050	0.050	-1.492		-1.492		-0.059	-0.058
CRSEF on										
ACADT	0.467	0.467	0.095	0.095	4.937	*	4.937	*	0.314	0.310
MENTF on										
ACADT	-0.257	0.119	0.126	0.154	-2.047	*	0.772		-0.128	0.058
ACADH on										
ACADT	0.341	0.341	0.089	0.089	3.822	*	3.822	*	0.201	0.198
SOCH on										
ACADT	0.139	0.139	0.094	0.094	1.491		1.491		0.075	0.074

Table 6.40

Final Structural Model—Unanalyzed Relationships, Low-SES Students

	Estimates		Std. Errors		Estimates/Std. Errors				Standardized	
Effect	TRH	L/L	TRH	L/L	TRH		L/L		TRH	L/L
WHITE with										
HSGRADES	-0.031	-0.031	0.009	0.009	-3.467	*	-3.467	*	-0.080	-0.120
FEMALE	-0.007	-0.007	0.005	0.005	-1.400		-1.400		-0.033	-0.036
PRETEST	-0.043	-0.043	0.009	0.009	-4.744	*	-4.744	*	-0.110	-0.118
HSGRADES with										
FEMALE	-0.028	-0.028	0.009	0.009	-3.144	*	-3.144	*	-0.071	-0.100
PRETEST	-0.043	-0.043	0.016	0.016	-2.663	*	-2.663	*	-0.059	-0.083
FEMALE with										
PRETEST	0.042	0.042	0.009	0.009	4.667	*	4.667	*	0.107	0.107
ACADP with										
SOCP	0.340	0.340	0.016	0.016	21.269	*	21.269	*	0.622	0.622
CRSEF	0.178	0.085	0.014	0.021	12.679	*	4.048	*	0.399	0.191
MENTF	0.057	0.057	0.010	0.010	5.730	*	5.730	*	0.171	0.171
ACADH	0.059	0.059	0.012	0.012	4.925	*	4.925	*	0.145	0.145
SOCH	0.068	0.068	0.011	0.011	6.155	*	6.155	*	0.182	0.182
SOCP with										
CRSEF	0.189	0.189	0.016	0.016	11.825	*	11.825	*	0.374	0.374
MENTF	0.130	0.130	0.013	0.013	9.977	*	9.977	*	0.340	0.340
ACADH	0.046	0.046	0.014	0.014	3.314	*	3.314	*	0.100	0.100
SOCH	0.048	0.048	0.012	0.012	3.943	*	3.943	*	0.113	0.113
CRSEF with										
MENTF	0.230	0.230	0.016	0.016	14.356	*	14.356	*	0.741	0.741
ACADH	0.048	0.048	0.012	0.012	3.975	*	3.975	*	0.126	0.126
SOCH	0.012	0.012	0.011	0.011	1.064		1.064		0.034	0.034

(continues)

Table 6.40

Final Structural Model—Unanalyzed Relationships, Low-SES Students (continued)

	Estimates		Std. Errors		Estimates/Std. Errors				Standardized	
Effect	TRH	L/L	TRH	L/L	TRH		L/L		TRH	L/L
MENTF with										
ACADH	0.046	0.046	0.009	0.009	5.111	*	5.111	*	0.161	0.161
SOCH	0.010	0.010	0.008	0.008	1.238		1.238		0.038	0.038
ACADH with										
SOCH	0.234	0.234	0.014	0.014	16.686	*	16.686	*	0.737	0.737

Notes.
Because latent factors ACADP, SOCP, CRSEF, MENTF, ACADH, and SOCH were technically endogenous due to the influence of model covariates, unanalyzed relationships were between each latent factors' disturbances.

Chapter Conclusion

In this chapter, results pertaining to this study's final research question were detailed. Specifically, I sought to understand whether participatory inequities identified in Chapter 5 were likely to have resulted in poorer student outcomes. Three groups of L/L participants were compared to their same-race/same-SES peers living in the TRH environment, including: (a) Asian/Pacific Islander students participating in disciplinary, general academic, and honors L/L programs, (b) White students participating in international/global L/L programs, and (c) Low-SES students participating in honors L/L programs. The specific outcome explored for Asian/Pacific Islander and Low-SES students was their smooth academic transition to college, while, for White students, diversity appreciation was investigated. For each group of students, separate LM models were developed for the TRH and L/L environments.

Using those models, it was possible not only to calculate the means associated with key environmental and outcome measures, but also to determine whether potentially causal relationships among those measures varied between groups. Although the LM models varied within—and across—student groups, one trend was consistent: although TRH and L/L students generally reported statistically significantly different mean scores on key facets of their living and learning environments, those discrepancies did not translate into statistically significant differences in student outcomes. In the next chapter, I consider the implications of this work.

Chapter 7: Discussion

This study explored *whether students from various racial/ethnic and SES groups who were underrepresented in a particular thematic type (or types) of L/L programs reported better outcomes than their peers living in the TRH environment.* It demonstrated that, while differences on NSLLP outcome measures were not detected between L/L and TRH participants, underrepresented students in L/L programs often reported stronger living and learning environments than did their TRH peers. As a consequence, it provides preliminary evidence that certain student subgroups are, for whatever reason, not accessing a potentially high-impact program from which at least some benefit may accrue. Given the challenges faced by many students in postsecondary education today (n.b., Kinzie, Palmer, Hayek, Hossler, Jacob, & Cummings, 2004; Knapp, Kelly-Reid, and Ginder, 2009; Leslie & Brinkman, 1988), misalignment between student needs and student and institutional behaviors is detrimental not only to those directly involved, but to our nation as a whole.

The study answered three instrumental research questions. First, it asked whether a team of raters could develop a thematic typology of L/L programs, based upon programs' stated goals and objectives. Second, it explored whether students from different racial/ethnic backgrounds or socioeconomic statuses participate in various thematic types of L/L programs at a rate proportional to their presence in the larger L/L community. Finally, it sought to determine whether students underrepresented in a given thematic type (or types) of L/L program accrued differential benefits by participating in

L/L programs when compared to their peers living in traditional residence hall environments, net of other characteristics constant.

This chapter begins with a discussion of the results of each research question, and places those findings within the context of prior research related to L/L programs. Then, it considers limitations that should be kept in mind when considering those results. It concludes with the implications of this work for both researchers and residential educators, and then discusses directions for further research.

Instrumental Research Question One: A Thematic Typology of LLPs

This study's first research question asked whether, to allow for more accurate and nuanced comparisons between types of programs, a team of raters could develop a thematic typology of L/L programs based upon programs' stated goals and objectives. To do so, Neuendorf's (2002) technique of descriptive content analysis was employed by six raters, split in to two panels of three members each. In total, content describing the 611 programs participating in the 2007 NSLLP, including their names, brief descriptions, and intended outcomes, was analyzed.

Raters used an electronic rating program, developed in Microsoft Access by the researcher, to conduct their content analyses. Inkelas et al.'s (2004) original thematic typology of living learning programs, developed as part of the 2004 NSLLP, was used as the initial codebook. That codebook was refined and expanded throughout three waves of coding, including pilot coding and two waves of full coding. Krippendorf's alpha for pilot coding was .67, and ranged from .65 to .70 in the first wave of full coding, above the .60 threshold of acceptability established by Lombard et al. (2002). The second wave of full coding was conducted with input from all panelists, with the goal of reaching consensus on the appropriate typological membership of programs not typed in prior rounds. Ultimately, 555 of 611 programs were coded (91%). The remaining nine percent of cases did not contain enough information for coders to reach a unanimous decision on the program's likely theme.

The final thematic typology identified 41 specific types of L/L programs, compared to the 26 developed by Inkelas et al. in 2004. Of the 15 new thematic types identified by coders, five were disciplinary in nature (i.e., agriculture/pre-veterinary medicine, communication and journalism, interdisciplinary, law and criminal justice, and mathematics). The remaining ten ran the gamut from leisure programs (e.g., programs focused on sports or activities) to transition programs focused on distinct student groups (e.g., students identifying as LGBT or allies) to programs for students participating in ROTC. The complete thematic typology appears in Appendix G.

Among L/L participants for who program membership could be definitively established, half were in manifestly academic programs: 26% were in disciplinary programs and another 24% were in programs identified as being connected to academic honors programs. Transition programs for first-year students enrolled another 10% of participants. While it is not possible to generalize these findings to the population of L/L participants nationally, the prominence of programs focused on specific academic interests, honors students, and transition programs is consistent with the historical development of these programs (Gabelnick et al., 1990; Schoem, 2004) and with previous typologies of learning communities (n.b., Lenning & Ebbers, 1999; Love & Tokuno, 1999).

While the thematic typology of L/L programs derived from the 2007 NSLLP can be seen to suggest that many students participated in programs that addressed enduring themes, it also demonstrates diversity of this intervention. To be sure, one cannot infer that the new thematic types of L/L programs identified by raters, such as environmental sustainability and spirituality and faith-based programs, means that no programs of this type existed at the time of Inkelas et al.'s (2004) typology: their "discovery" here could be simply an artifact of sampling and the expansion of the NSLLP between 2004 and 2007, in which the number of programs evaluated more than doubled. Still, the sheer breadth of types of programs offered to students is indicative of innovation at the campus level and practitioners' efforts to develop residential programs that will engage students with a wide range of interests.

In contrast to the thematic diversity evidenced by programs participating in the 2007 NSLLP, the number of student learning and development outcomes on which these programs focused was limited. In particular, all but five program types (i.e., political interest, research, ROTC, umbrella, and general wellness) rated one (or more, in the case of ties) of the following four outcomes as their most important: (a) cultivating a sense of belonging, (b) facilitating students' academic transition, (c) promoting liberal learning, and (d) encouraging diversity appreciation. With the possible exception of encouraging diversity appreciation, the key outcomes listed above are consistent with the purposes with learning communities broadly, and their residential incarnations in particular (Gabelnick et al., 1990; Shapiro & Levine, 1999; Inkelas & Weisman, 2003).

It is interesting to note, however, that these findings suggest programs may not be placing equal emphasis on both the living *and* the learning components of L/L programs.

While this analysis did not explicitly investigate the structure of L/L programs (e.g., the presence and number of academic courses or the number of faculty participants), by looking across the ratings of importance of NSLLP outcomes, it is possible to infer something about the general motivations of program types. For example, while only one program type (language programs) did not have an overall rating of at least "important" for sense of belonging, increasing students' ability to explore ideas did not rate as important for 17 types, and growth in cognitive skills did not rate as important for 12 types. Similarly, the development of program types that have at least the potential to be manifestly non-academic, such as general leisure programs, culinary arts programs, and spirituality programs, suggests that the nature of the learning occurring in some L/L programs might differ from that expected by proponents of traditional learning communities.

Instrumental Research Question Two: Participatory Equity in LLPs

This study's second research question investigated the status of participatory equity among students from different racial/ethnic and SES groups for each thematic type of L/L program. Specifically, it asked: Did students from different racial/ethnic backgrounds or socioeconomic statuses participate in various thematic types of L/L programs at a rate proportional to their presence in the larger L/L community?

To answer that question, Hao's (2002) equity indices were calculated to identify instances in which participatory inequity on the basis of race/ethnicity and socioeconomic status might be present. Three steps constituted this analysis: (a) identifying students by thematic type of L/L program, (b) verifying or creating variables needed for equity calculations (i.e., race/ethnicity and SES group), and (c) generating the equity indices themselves.

The initial step—essentially "recoding" the variable that indicated the specific program in which a L/L participant was enrolled to instead indicate the specific thematic type of that program—was simple on face. Unfortunately, because the variable indicating program membership was taken from the student questionnaire as opposed to institutional records, an unexpectedly large proportion of students (18.3%) could not be connected to a particular program and had to be excluded from the analysis. Cases that could not be matched were flagged for further analysis to determine whether a relationship existed between program missingness and students' race/ethnicity or SES group.

In the second step, variables representing students' race/ethnicity and SES group were reviewed and created, respectively. Because the race/ethnicity variable already existed on the NSLLP dataset, only simple descriptive statistics need be computed. A chi-square test of independence indicated that L/L participation was not independent of race/ethnicity, such that students who were identified as African American were underrepresented in the L/L sample overall (compared to the TRH sample), while students who were identified as Asian/Pacific Islander or Multiracial/multiethnic were underrepresented in the TRH sample overall (compared to the L/L sample). As noted above, the relationship between program missingness and race/ethnicity was tested, revealing that African American students were disproportionately represented among those students for whom program enrollment information was not available.

A composite variable representing SES was created, using students' self-reports of their parents' highest level of educational attainment and family income. Once the composite was created, a robust independent sample t-test was computed to determine whether a statistically significant mean difference existed between L/L and TRH participants' SES. No difference was found. Because the distribution of SES was generally equivalent across both L/L and TRH students, three SES groups of roughly equivalent size were created, using cut-points defined by the 33rd and 66th percentiles of SES. Once again, the relationship between program missingness and student characteristics was tested, revealing that students in the lowest SES group were disproportionately represented among L/L participants for whom program enrollment information was not available. This may have biased subsequent EI calculations and, to the extent that excluded low-SES students reported experiences and outcomes that varied from included low-SES students, results of latent mean modeling.

In the final step of this analysis, Hao's (2002) EIs were computed (see Tables 5.4 and 5.5, for data concerning participatory equity related to race/ethnicity and SES, respectively). Some level of participatory inequity was noted for all racial/ethnic groups. Specifically, I found EIs below Bensimon et al.'s (2003) cut-off of .800 for: (a) African American students in five program types, including general science, honors, residential college, general wellness, and education programs; (b) Asian/Pacific Islander students in six program types, including leadership, general academic, general sciences, honors, business, and humanities; (c) Latino students in three program types, including general science, honors, and health science; (d) White students in three program types, including domestic multicultural, mathematics, and international global; and (e) Multiracial/multiethnic students in five program types, including education, health science, general sciences, business, and leadership.

It should be noted that not all of the participatory inequities noted above appeared to be of the same magnitude. Of the 22 total race/ethnicity × program type combinations that evidenced participatory inequity, eight had EIs of .60 or below. While the lowest EI was found for White students in domestic multicultural programs (.32), African American students had four of the lowest EIs: general science programs (.30), honors programs (.41), residential colleges (.42), and general wellness programs (.58). Students identified as Asian/Pacific Islander had particularly low EIs in leadership programs (.40), students identified as Multiracial/multiethnic had low EIs in education programs (.43) and health sciences programs (.50), and Latino students had low EIs in general science programs (.52). While in some regards "inequity is inequity," the statistics above may reflect particular, persistent inequalities that deserve special attention. Indeed, to ignore them in a time where our nation's economic security and competitiveness is increasingly linked to attainment at the postsecondary level seems ill-advised at best.

Thirteen SES-related inequities were noted, of which five had EIs of .60 or below. Although participatory inequities were found for low SES group students in five program types, none were below .60. Each of the middle SES group's three inequities was below .60, including political interest (.55), domestic multicultural programs (.56) and law and criminal justice programs (.58). Finally, two of the high SES group's four inequities were below .60, including education (.52) and upper-division programs (.56).

Although not a focus of this research, the prevalence of EI's well *above* 1.0—indicating that a group is overrepresented in a particular analysis—is notable. Drawing upon Bensimon et al.'s (2003) logic that 0.8 indicates underrepresentation, the informal adoption of an EI of 1.2 to indicate overrepresentation suggests several group × program type combinations where this phenomenon may exist. While some of these apparent inequities are likely an artifact of low absolute numbers of participants in a given group (e.g., students identifying as American Indian), others may be valid indicators of instances where a particular type of student is clustered in a type of program at a rate far greater than would ordinarily be expected, given that type of student's representation in the larger pool of L/L participants.

Placing these findings within the context of past scholarship is complicated by two factors: the dearth of equity-interested research in L/L programs and changes between the 2004 and 2007 NSLLP, the only source of data on which such analyses could be conducted. My finding that African American, American Indian, Asian/Pacific Islander, Latino, and low-SES students were all underrepresented in Honors programs is consistent with the work of Soldner et al. (2007), who reached substantively similar conclusions using a prior round of NSLLP data. New research in this vein, particularly research that examined other types of L/L programs using other data sources, could lead additional credence to the present findings.

Whether the present study conflicts with Johnson's (2007) assertion that women of color are underrepresented in STEM-focused L/L programs is less clear. Reanalysis of her data using EIs suggest that the true underrepresentation Johnson identified was localized in women-only STEM L/L programs, not STEM programs more generally. My analysis suggests that students of color are *overrepresented* in women-only STEM L/L programs, but that students (although not necessarily women) of color are inconsistently represented in types of L/L programs that might be considered STEM. For example, no racial group is under equity in engineering and Computer Science programs, but three of six racial groups were under equity in general science programs. It is important to remember, however, that the 2007 NSLLP had an explicit focus on women in STEM fields. Whether these findings are idiosyncratic to the 2004 collection (because we did not have enough women in STEM programs to accurately measure them) or to the 2007 collection (because we "oversampled" STEM programs and somehow confounded my results) cannot be readily determined.

Instrumental Research Question Three: Differences in Student Outcomes

This study's final research question sought to understand whether students underrepresented in a given thematic type (or types) of L/L program accrued differential benefits by participating in L/L programs when compared to their peers living in traditional residence hall environments, holding other characteristics constant. In total, three sets of analyses were conducted—one for Asian/Pacific Islander students, one for White students, and one for low-SES students—comparing L/L students in particular thematic types of L/L programs to non-participants. To address the final research question, this study employed a form of structural equation modeling known as latent mean modeling (LMM). The particular implementation of LMM used here is conceptually related to a familiar statistical procedure, analysis of covariance (ANCOVA), which seeks to identify statistically significant differences between two or more groups net of other important predictor variables.

Latent mean modeling took place in several steps. First, in a traditional confirmatory factor analysis, the factor structure of key latent constructs used in the analysis (that is, the relationship between survey items on the NSLLP and the college environments and learning outcomes they were believed to represent) was verified to ensure they were appropriate for the subpopulations under study, both separately and when analyzed simultaneously. Doing so allowed the development of latent measures that were free of measurement error and provided psychometric evidence about the extent to which the relationships between NSLLP survey items and the latent constructs they were designed to reflect were invariant across student groups.

Then, using standard structural equation modeling techniques, the hypothesized relationships among latent structures (e.g., college environments and student outcomes) and measured covariates (e.g., student characteristics and outcome pretests) depicted in Figure 3.1 were statistically imposed and tested. This stage of modeling provided an opportunity to evaluate the relationships among key living and learning environments—such as peer and faculty interaction— and student outcomes so often noted in the literature (n.b., Inkelas, Johnson, et al., 2006; Inkelas & Weisman, 2003; Lenning & Ebbers, 1999; Pascarella & Terenzini, 1980; Pike, 1999; Shapiro & Levine, 1999) at the student group × program type level.

Finally, the prior two stages of analysis were combined to yield information about the differences, if any, in the learning outcomes reported by students in a particular thematic type of L/L program and their non-participating peers. Broadly, stage one generated the best possible measurements of the outcomes to be compared, while stage two incorporated important covariates that, had they not been controlled, might otherwise confound my efforts to find differences between L/L participants and TRH students. Because these analyses were conducted within the framework of covariance and mean structure analysis, as opposed to the generalized linear model, strict statistical assumptions that often complicate AN(C)OVA-like analyses could be avoided (Hancock

& Mueller, 2009), generating more—and hopefully better—information about the role L/L programs play in promoting student learning.

Overall. Three separate analyses were conducted for: (a) first-year Asian/Pacific Islander students in disciplinary, general academic, and honors L/L programs and their TRH peers, exploring academic transition; (b) White students in international/global L/L programs and their TRH peers, exploring appreciation of diversity; and (c) first-year low-SES students in honors programs and their TRH peers, exploring academic transition to college.

All models evidenced good fit, based on my *a priori* selection of Hu and Bentler's (1999) joint fit statistics criteria. For Asian/Pacific Islander students, 17.4% and 15.1% of the variance in TRH and L/L students' outcome scores, respectively, were explained. Among White students, the percentages of variance explained in TRH and L/L students' outcome scores were 20.1% and 21.8%, respectively. For low-SES students, 24.1% and 26.3% of the variance in TRH and L/L students' outcome scores, respectively, were explained.

Academic transition. No statistically significant mean difference in academic transition between L/L and TRH participants was found for either Asian/Pacific Islander or low-SES students. For both the analysis of Asian/Pacific Islander students in disciplinary, general academic, and honors L/L programs and of low-SES students in honors programs, this is consistent with Inkelas and Weisman's (2003) finding that there was no statistically significant difference in self-reported academic transition scores between participants in "curriculum-based" programs or "academic honors" programs and students in the TRH environment. Indeed, in their analysis, only participation in a "transition program" resulted in higher mean scores on the academic transition outcome measure.

Although, on face, the present findings are at odds with those documented by Inkelas et al. (2007) and in the 2004 and 2007 NSLLP technical reports (Inkelas, Brower, & Associates, 2004; 2008), two considerations must be kept in mind when interpreting prior results. Although in all three reports the authors suggest that there is a main effect for L/L participation, the magnitude of the effect was either clearly characterized as small (Inkelas et al., 2007) or was likely to have been small, had effect sizes been computed: Inkelas, Brower and Associates reported differences of only between .06 and .10 on a six point scale. At least in the case of the NSLLP technical findings, the observed difference may only have been detectable due to the substantially larger statistical power afforded by using the whole NSLLP dataset. Additionally, while the present study chose to focus on first-year students because of the relative recency of their experience, both Inkelas et al. and Inkelas, Brower, and Associates included all students in their analysis. Taken together, the present findings and past literature suggest that it may well be that although general academic, honors, and disciplinary programs intend to smooth their participants' academic transition to college, they do not. (Or, at the very least, they did not for the populations under investigation here.)

Diversity appreciation. White students who participated in international/global programs reported means on the NSLLP diversity appreciation outcome measure that were statistically indistinguishable from their TRH peers. This is wholly consistent with findings documented by Inkelas, Brower and Associates in the 2004 (2004) and 2007 (2008) NSLLP technical reports. Given that the present study statistically controlled for several covariates prior research has suggested should be related to residential outcomes generally, and diversity appreciation in particular, it seems to represent a more thorough investigation that offers further evidence that L/L programs—at least as currently implemented and measured—do not contribute to this important student learning outcome.

My observation that there were no differences in the outcomes reported by L/L and TRH students in the present study is a cause for speculation—speculation that can be guided by investigating the larger conceptual framework in which those outcomes were embedded. First, I consider the relationship of each of the six key living and learning environments that served as the foundation for that framework to Asian/Pacific Islander and low-SES students' transition to college and White students' appreciation for diversity. Then, I consider the relationship between the student background characteristics and other covariates included in the conceptual framework and the outcomes explored here.

Academic/vocational peer conversations. For the Asian/Pacific Islander students in the sample, conversations with their peers about academic and vocational issues were positively related to perceptions of a smooth academic transition, irrespective of whether students participated in an L/L program or the TRH environment. This is consistent with Inkelas and Weisman's (2003) finding that academic conversations with peers were positively related to smoother academic transition for students in honors L/L programs and in the TRH environment, and is consistent with reviews of the general college impact literature (e.g., Pascarella & Terenzini, 1991, 2005). Its variance from Inkelas et al.'s (2007) findings, which would have suggested no relationship, may be a function of either different populations (i.e., first-generation students versus those whose parents were exposed to postsecondary education) or an interaction with program type.

No difference in the strength of the relationship was noted between L/L and TRH students, although L/L participants reported more frequent academic/vocational conversations with their peers than did students in the TRH environment (medium effect), consistent with Inkelas and Weisman's finding (2003) that students in honors programs were more likely to discuss academic issues outside of class with their peers, Pike's (1999) finding that L/L students reported more academically integrative conversations than their non-participating peers, and national norms presented in the 2004 and 2007 NSLLP technical reports (Inkelas, Brower, & Associates, 2004, 2008).

While academic/vocational conversations with peers was significantly positively related to low-SES students' smooth academic transition for the TRH sample, no such relationship was noted for L/L participants. Although low-SES L/L participants continued to report this type of interaction more frequently than their TRH peers (small effect), it appeared to have no effect on their transition. An extrapolation of Inkelas et al.'s (2007) findings vis-à-vis first-generation students to the low-SES population might have been taken as evidence that such an effect should not have been expected. However, this finding remains puzzling because the thematic type of program explored for low-SES students was honors programs, where one might expect this type of peer interaction to be particularly salient.

One possible reason for this finding could be the distinction between *frequency* and *quality*. It may be that low-SES students in honors L/L programs are experiencing frequent—but not necessarily positive—academically-oriented interactions with their peers. More detailed investigation into the specific experience of low-SES students in residential honors programs seems indicated.

For White students, the conceptual framework posited a relationship between academic/vocational peer conversations and students' appreciation for diversity. No such relationship was noted for either L/L or TRH sample members. This is consistent with Longerbeam's (2005) finding that discussions of academic and career issues with peers were unrelated to L/L participants' openness to diversity.

Socio-cultural peer conversations. Prior research by Inkelas and Weisman (2003) and Inkelas et al. (2007) found that socio-cultural peer conversations have a neutral to negative relationship to students' perceived academic transition to college, with the authors suggesting that the negative relationship observed for the honors students in their sample might result from conversations "gone awry" that inadvertently alienate students. The present study noted no statistically significant relationship between students' socio-cultural peer conversations and perceived academic transition for either Asian/Pacific Islander or low-SES students. In both cases, however, L/L participants reported more frequent socio-cultural conversations than their TRH peers (medium effects), a finding that is consistent with both the 2004 and the 2007 NSLLP national norms (Inkelas, Brower, & Associates, 2004, 2008).

While the general notion that socio-cultural peer interaction would not be related to academic transition may, on face, be difficult to grapple with, it is helpful to remember how these two constructs are operationalized. As Inkelas and Weisman note, socio-cultural peer conversations, as measured by the NSLLP, *are* challenging, and include navigating issues of political, religious, and racial/ethnic difference. In contrast, academic transition is measured by a set of fairly pragmatic behaviors, including communicating with faculty outside of class, seeking assistance when needed, and forming study groups. The disconnection between these constructs may, in part, explain these findings.

In contrast, socio-cultural peer conversations might be expected to be directly related to students' openness to diversity. Longerbeam (2005) concluded as such, and, indeed, it was the *strongest* predictor in her model net of dozens of other variables representing student and programmatic characteristics. Longerbeam's findings are confirmed in whole by the present study, which also documented the importance of socio-cultural peer conversations for White TRH students' diversity appreciation. Consistent with the 2004 and 2007 NSLLP national norms (Inkelas, Brower, & Associates, 2004, 2008), White students participating in international/global diversity programs reported higher levels of socio-cultural peer conversations than did their TRH peers (medium effect).

Course-related faculty interactions. Inkelas and Weisman's (2003) and Inkelas et al.'s (2007) prior research on academic transition suggested that, for both L/L and TRH students, discussions with faculty members about academic issues was positively related to students' smooth academic transition to college. Indeed, the authors concluded that—for L/L participants in the case of Inkelas and Weisman and for all students in the case of Inkelas et al.—it was the strongest predictor of smooth transition net of other key living learning environments and student background characteristics.

The present study only partially confirms these findings. For low-SES students in academic honors programs—and for their peers in the TRH environment—course-related faculty interaction *was* the living and learning environment most strongly related to a smooth academic transition. However, for neither Asian/Pacific Islander students in academically-focused L/L programs nor their TRH peers was course-related faculty interaction related to transition. Consistent with Inkelas and Weisman's (2003) findings, there was no statistically significant difference in the amount of course-related faculty interaction reported by L/L participants in these academically-focused programs and their non-participating peers.

Whether or not the discrepant finding for Asian/Pacific Islander students is an artifact of either analysis or an actual race-related phenomenon is not clear. Neither Inkelas and Weisman's (2003) nor Inkelas et al.'s (2007) analyses disaggregated students by race, so although race was included as a covariate in their model, it may be that race-related differences were obscured.

However, it is equally as tenable that the relatively small sample size encountered here made it impossible to detect a "true" relationship between Asian/Pacific Islander students' course-related faculty interaction and academic transition. Work by Kim, Chang, and Park (2009) have suggested that Asian/Pacific Islander students not only report lower levels of interaction with faculty, but also that those interactions are of lower quality than their White or Latino peers. Because this study did not compare across racial/ethnic groups, it is impossible to determine whether this was the case with our L/L respondents or their TRH peers. However, it may help explain why course-related faculty interaction failed to benefit Asian/Pacific Islanders' academic transition to college.

Prior work by Longerbeam (2005) suggested that course-related faculty interaction should be unrelated to the development of diversity appreciation for students in residential environments. Among the White students in the present study, that was indeed the case. Prior to this research, whether one would expect to find differences in the amount of course-related faculty interaction reported by White L/L and TRH students was an open question. Without past work that explored outcomes associated with this particular type of program (unlike the academically-focused programs above, which had been studied by Inkelas and Weisman [2003]), the 2004 and 2007 NSLLP national norms were the only source of relevant data, and they disagreed: no difference had been noted between L/L and TRH students in 2004, but an statistically significant difference was found in 2007 (Inkelas, Brower & Associates, 2004, 2008). Although it appears to have been to no particular end based on a lack of relevance to the outcome of interest, White students participating in international/global L/L programs reported significantly more course-related faculty interaction than did their TRH peers (medium effect).

Non-course-related faculty mentorship. On face, non-course-related faculty mentorship, which includes student-faculty discussions of personal and academic concerns and informal visits with faculty members during social occasions, might tenably be thought to be related to more positive academic transition, as that concept includes communicating with instructors outside of class and ease of seeking assistance and personal help. Prior to this effort, however, no study had borne out that intuitive relationship in the residential environment (Inkelas & Weisman, 2003; Inkelas et al. 2007), and the results of this study were consistent in that it found non-significant relationships among Asian/Pacific Islander students (both L/L and TRH) and among low-SES L/L participants.

Interestingly, although a statistically significant relationship was observed between low-SES TRH students and non-course-related faculty mentorship, the actual direction of the effect was negative. The most tenable explanation for this finding is that the “conventional wisdom” above mis-specifies the temporal ordering of the two phenomena: non-course-related interaction does not lead to a smooth academic transition, but instead a rocky academic transition results in TRH students having discussions with faculty about their problems. Longitudinal data could confirm or deny this supposition, which cannot be formally tested here.

No difference in the amount of non-course-related mentorship was detected between L/L and TRH students for Asian/Pacific Islander students, which is inconsistent with both the 2004 and the 2007 NSLLP national norms (Inkelas, Brower, & Associates, 2004, 2008). In comparison, low-SES students in L/L programs reported significantly more non-course-related mentorship than their TRH peers (medium effect). Taken together, my findings vis-à-vis low-SES students suggest that the extra non-course-related mentorship L/L participants receive offers no benefit, per se, but the comparatively small amount experienced by TRH students is actually associated with negative outcomes. This gives further credence to the hypothesis that, in the TRH environment, whatever small amount of non-course-related faculty mentorship goes on is in response to transition difficulties.

Longerbeam’s (2005) exploration of non-course-related faculty mentorship and diversity appreciation suggested that any relationship between the two was unlikely, but possible (she reported the p-value associated with the relationship to be significant at the $p \leq .10$ level). Interestingly, while the relationship was not statistically significant in the present analysis at $p \leq .05$ level, the associated t-statistic was 1.771, which would have been significant at the $p \leq .10$ level. Whether this relationship has been detected twice due to chance, or whether there is something about informal faculty contact that promotes diversity appreciation, merits further exploration.

Academically supportive residence hall climates. For both Asian/Pacific Islander and low-SES students in the sample, academically supportive residence hall climates were positively related to a smooth academic transition. No difference was noted in the strength of those relationships based upon participation in an L/L or residence in the TRH environment. Inkelas and Weisman's (2003) prior finding that students participating in honors L/L programs or living in the TRH environment evidenced a relationship between this climate measure and their transition is generally consistent with the findings of this study.

While Inkelas et al. (2007) found no relationship between academically supportive hall climates and students' transition, it is important to remember that the authors restricted their sample to first-generation college students. The extent to which their population is similar to the low-SES sample studied here is not known, but given that parental education was one half of this study's SES variable (income was the other), overlap should be expected. Given that income had no relationship with transition in Inkelas et al.'s models for either L/L or TRH students, it may be that there is some minimum threshold of parental education, below which the benefits of supportive academic climates do not accrue. Alternatively, it may be that the approximately 500 additional cases used in the present analysis were sufficient to detect an effect when Inkelas et al.'s could not.

Although academically supportive residence hall climates were similarly and positively related to students' academic transitions, low-SES L/L students reported significantly higher levels of perceived support than did their TRH peers (medium effect), while no residential-environment-related difference was noted for Asian/Pacific Islander students. Inkelas and Weisman (2003) had previously noted L/L participants in honors programs reported higher mean academic support scores than their TRH peers, consistent with my findings for low-SES students. Given that NSLLP national norms in both 2004 and 2007 (Inkelas, Brower & Associates, 2004, 2008) have consistently shown a statistically significant difference between L/L and TRH populations on this dimension, further specific investigation into the experience of Asian/Pacific Islander students could help discern whether this finding is a function of thematic type or may be race-related.

The level of perceived academic support felt by White students was also positively related to their level of diversity appreciation, irrespective of residential environment. No difference in the strength of the relationship by residential environment was noted. This relationship is consistent with Longerbeam's (2005) findings for L/L participants, and suggests that the effect may be able to be generalized to the TRH environment as well. In the light of the NSLLP national norms cited above, it is not surprising that the White L/L participants in this sample reported substantially greater levels of academic support than did their TRH peers.

Socially supportive residence hall climates. No relationship between Asian/Pacific Islander or low-SES students' perception of their residence hall as socially supportive and their perceived academic transition was found in this study. This runs counter to Inkelas and Weisman's (2003) prior findings about honors L/L program participants, but is consistent with their findings about students participating in disciplinary programs and TRH students more generally. It is also consistent with Inkelas et al.'s (2007) subsequent findings about first-generation students. Asian/Pacific Islander and low-SES L/L participants reported residential environments that were more socially supportive than did their TRH peers (small and medium effects, respectively).

In contrast, low-SES students in both L/L and TRH environments reported that socially supportive residence hall climates were positively related to their level of diversity appreciation. Although Longerbeam (2005) did not include TRH students in her work, this is consistent with her findings vis-à-vis L/L participants. As was the case with Asian/Pacific Islander and low-SES students, L/L participants reported more socially supportive halls than did their TRH peers (small effect). Taken together, these findings are consistent with the 2004 and 2007 NSLLP national norms (Inkelas, Brower, & Associates, 2004, 2007).

Given how this latent factor was operationalized across each of these analyses, the lack of a relationship to academic transition but the presence of a relationship to diversity appreciation is not particularly surprising. Each of the items that loaded on the social support factor dealt with interpersonal concordance, particularly across social differences. Indeed, it may be that "socially supportive hall environment" might be more aptly named "hall diversity climate," which could clearly bear on students' subsequent appreciation for diversity. In contrast, with the possible exception of facilitating study group formation (the lowest loading item on the academic transition factor), hall diversity climate is unlikely to affect academic transition, operationalized to include faculty interaction and ease of accessing help when needed.

Covariates. The present study was primarily concerned with the relationship between six key living and learning environments and students' learning and development outcomes. However, consistent with Astin's (1993) IEO framework, covariates were included to control for student-level characteristics that might tenably be thought to be related to either living and learning environments or the outcome measure. As depicted in Figure 3.1, direct relationships were posited between each covariate and each key living and learning environment and the outcome measure. This allows us to calculate the total effect of a covariate on the outcome measure, which consists of both its direct (unmediated) and indirect (mediated through environments) effect. Below, I discuss the relationship between covariates and the outcome measure only. More detailed information about the direct relationship between covariates and environment measures can be found in Chapter 6.

Socioeconomic status. SES, which was operationalized as the equally-weighted linear composite of parental education and family income, had no direct effect on the academic transition of either Asian/Pacific Islander students or low-SES students, irrespective of residential environment. This is consistent with Inkelas & Weisman's (2003) finding that family income and aid package had no relationship to students' subsequent rating of academic transition, though somewhat counterintuitive in the light of general theories of social capital. SES did, however, have a positive, indirect effect for Asian/Pacific Islander students participating in L/L programs, operating through the relationship between SES and frequency of academic/vocational peer conversations. While this finding merits further research, at least two rationales seem tenable: (a) higher-SES Asian/Pacific Islander students are simply more prone to engage with their peers on academic and vocational issues, or (b) SES operated in this model as a proxy for acculturation, distinguishing between Asian/Pacific Islander students who were more (or less) prone to discuss potentially personal issues with others in their residential environments.

SES exhibited a direct, negative effect on diversity appreciation for White students in both L/L and TRH environments. This was partially offset by a positive indirect effect in both environments, but a net negative effect remained. As this was not an element of Longerbeam's (2005) model, the relationship between SES and diversity appreciation within the residential environment is a fruitful subject for further inquiry.

The most likely explanation for this finding is the comingling of White privilege, generally low levels of White racial identity development among majority students, exposure to difference, and SES. Higher SES, particularly when it is reflective of higher levels of income, affords White individuals the ability to "shape" their lived experience. They may choose to live in more affluent areas, for example, or attend more well-resourced schools. Residential segregation and income disparity, then, may prevent White students from encountering students of color on a routine basis. White students' infrequent (or unproductive) contacts with students of color may also allow them to remain at lower levels of racial identity development. High-SES Whites may be particularly vulnerable to information processing strategies typified by reintegration, in which a "I've got mine" mentality is reinforced by White *and* economic privilege, further depressing openness to diversity.

High school grades. High school grades did not exhibit statistically significant direct or indirect relationships with the outcome variable in any of the models evaluated in this study. This is consistent with Inkelas and Weisman's (2003) prior findings vis-à-vis academic transition. While Longerbeam (2005) noted a negative relationship existed between SAT/ACT score and students' openness to diversity, its relative strength within the context of the other predictive variables in her model was only moderate. The disconnect between her work and the present study may be due to a variety of factors, including her use of a multiracial sample (as opposed to a monoracial sample here), a lack of variability for the students in this study's White sample, or a true predictive difference between high school performance measured by grades and scholastic aptitude measured by a standardized test.

Gender. Being female had a significant, positive direct relationships to White students' diversity appreciation, consistent with Longerbeam's (2005) prior findings. However, the indirect effect of gender in the TRH environment almost completely offset its direct effect. Unfortunately, the absence of a TRH comparison group in Longerbeam's work makes it impossible to place this finding in context. Taken together, I conclude gender is salient only in the White L/L population. Consistent with the prior literature (Inkelas & Weisman, 2003; Inkelas et al., 2007), no effect was noted on the academic transition of Asian/Pacific Islander or low-SES study participants.

Academic class level. Because only first-year students were included in analyses for Asian/Pacific Islander and low-SES students, academic class level was only included as a covariate in the White models exploring diversity appreciation. For both White L/L and TRH participants, class level exerted a significant, positive, and direct effect on diversity appreciation. This is consistent both with Longerbeam's (2005) prior findings vis-à-vis L/L populations, and consonant with a developmental interpretation of diversity appreciation. First, the length of time a student has been exposed to the college environment is directly related to the extent of cross-racial contact he or she can experience, which Allport (1954) would suggest should promote appreciation for diversity. Alternatively, students' increasing cognitive development over time, from more dualistic to more relativistic, may facilitate their capacity to appreciate those from other racial/ethnic backgrounds (King & Baxter Magolda, 2005).

Pretest. Based on Astin's (1993) IEO model, there is an *a priori* assumption that pre-test measures will be strongly related to the outcome variable of interest. Such was the case in each of my analyses, irrespective of racial/ethnic or SES group membership or residential arrangement.

For Asian/Pacific Islander students, the total effect for the pre-test was approximately 93% (.182/.196) as strong as the strongest predictor in the TRH environment (academic/vocational peer conversations), and approximately 98% (.194/.197) as strong as the strongest predictor in the L/L environment (academically supportive hall climate). For White students, the total effect for the pre-test was approximately 171% (.293/.171) as strong as the strongest predictor in the TRH environment (socio-cultural peer conversations), and approximately 173% (.291/.168) as strong as the strongest predictor in the L/L environment (socio-cultural peer conversations). Finally, for low-SES students, the total effect for the pre-test was approximately 87% (.272/.314) as strong as the strongest predictor in the TRH environment (course-related faculty mentorship), and approximately 92% (.284/.310) as strong as the strongest predictor in the L/L environment (course-related faculty mentorship).

While these findings are consistent with the importance Astin (1993) has placed on students' pre-college characteristics, they are particularly striking when one considers how the NSLLP RES-B instrument framed its proxy pre-test items. In each case, students were not asked to rate their current *level* of a given outcome, but instead their pre-entry expectation that development of a given outcome was *important*. This strongly suggests a motivational or dispositional component to the outcomes observed here, which begs a critical question: do students "select" development, implicitly or explicitly choosing the direction and magnitude of their growth? Almost 30 years ago, Pace (1984) argued that learning and development required investments of student time and energy. This study seems to corroborate his basic argument, and to pose challenges to residential educators seeking to promote "difficult" learning. If a student enters a L/L program unwilling to fully engage with an outcome such as diversity appreciation, for example, what can residential educators do to challenge that belief? This problem may be intensified when institutions place L/L programs in highly desirable residential environments, access to which students might not otherwise have.

Limitations Concerning These Findings

The findings reported in this study should be considered in the light of its unique limitations, as well as those limitations that are inherent in the larger program of research in which it is embedded and educational research more generally. Below, I consider eight factors that pose threats to the study's internal and external validity (that is, the extent to which I can consider the outcomes in this study related to students' participation in LLPs and can generalize those findings to the larger population of LLPs and LLP participants, respectively).

Low absolute number of racial/ethnic minority students. More than three-quarters of NSLLP participants were White. (The non-White L/L population included about 2700 students.) The low absolute number of racial/ethnic minority students combined with the high number of thematic types of L/L programs across which those students were dispersed made the development of racial/ethnic minority groups for analysis challenging. The resulting solution—aggregating thematic types—was done with caution, focusing only on programs with shared outcomes. However, this aggregation may have muddled the effects of any given type of program, making the detection of difference impossible.

Self-selection of institutions to the study. Institutions that participated in the NSLLP did so voluntarily, “opting-in” after having heard of the study from one of any number of potential sources, including conference presentations, journal articles, professional colleagues, or informational calls from the study’s data collection contractor. Although it is possible to speculate about institutions’ reasons for participating, the process of self-selection introduces non-random error in to this study’s findings that cannot be quantified.

Self-selection of students to LLPs. Seemingly more problematic than an institution’s self-selection in to the NSLLP is a given student’s choice to live in an LLP, rather than a TRH environment. Indeed, one can imagine any number of possible motivations to participate in an LLP and can hypothesize endlessly about the relationships between those motivations and subsequent outcomes. What can be said with surety, however, is that this unobserved variable confounds the findings of this study. Because the classic experimental solution to this problem—the randomized controlled trial—would likely be impractical from the perspective of most housing officers, additional analytic approaches are introduced later in this chapter as potential alternatives as implications for future research are considered.

Cross-sectional studies and inferences related to change. The NSLLP was a cross-sectional study, employing a proxy pre-test design. In this type of research, change in the dependent variable is “controlled for” by a respondent’s contemporaneous assessment of his or her development prior to exposure to the intervention of interest. While this may be preferable to post-test-only designs, the concern that the intervention may affect both the proxy pre-test *and* the post-test, thereby threatening internal validity, remains.

Developing and using a typology of LLPs. Producing a new typology of LLPs was motivated by the desire to ensure that programs with similar foci were being aggregated in latent means analyses, decreasing the likelihood that grouping dissimilar programs would mute students’ learning and development gains. Although the typology of LLPs developed as part of this study is larger (or, at least, more detailed) than the one originally offered in Inkelas et al. (2004), it cannot be said to be definitive. Because the NSLLP did not (and could not) attempt to take a census of LLPs nation-wide, it is possible that certain types were not included among the more than 600 evaluated by this study’s raters. However, even if a census of programs had been available, it is not evident that a truly definitive typology could have been developed: the rating task documented here demonstrated that two raters can evaluate similar descriptive information and draw dissimilar conclusions, and that “insider knowledge” can bring to light instances where institutions inaccurately (or poorly) described programs that would have otherwise resulted in misclassification. As a result, while the typology developed in this study is thought to combine like programs for the sake of enhancing comparability, it may be that it still is not sufficiently granular to ensure a "living-learning effect" is not being masked.

Calculating equity indices. The calculation of EIs was the critical step in the identification of student group × program type combinations where students from various racial/ethnic or socioeconomic backgrounds may have been underrepresented. These combinations were then considered for further analysis. This research demonstrates that while the use of EIs is simple in theory, the integrity of the method can be compromised when population-level information is not known with certainty. Ideally, EI analyses would have been based on program-level counts of students by race/ethnicity and SES. Because this information was not available, I had to rely upon program-level counts of *responding* students and to adopt the assumption that any non-response bias related to my grouping variables was independent of, and identically distributed across (IID), program type.

The IID assumption would have been testable had program-level non-response weights been developed for the NSLLP. However, during the design phase of the umbrella study, only institution-level weights were planned for and, ultimately, created. While this problem can easily be avoided in future collections by requesting program-level demographic data from the institution, it may be that non-response bias at the program level in the present study introduced non-random error in to the EIs calculated in Chapter 5. Also complicating the EI analysis was a surprisingly large (18%) percentage of LL respondents that did not indicate the specific program in which they participated. Because data from these students could not be analyzed, any race/ethnicity or SES-related non-response on *this* variable could bias EI results. Again, this problem is easily soluble in future iterations of the NSLLP through better institution-level data collections.

Ordinal data. The latent factors used in this model were identified using indicator variables that had between four and six ordinally-scaled response options. The resulting data are by definition not normally distributed, violating a standard assumption of many analytic techniques including structural equation modeling (Byrne, 2006). Two problems arise from the use of non-normally distributed data: (a) attenuated correlations, and (b) biased fit statistics and standard errors (Hancock & Mueller, 2009).

When confronted with ordinal data, two solutions are commonly employed. First, analyses based on ordinal data should be conducted using robust estimators, like the Satorra-Bentler chi-square, so that fit statistics and standard errors are appropriately adjusted (Byrne, 2006). Second, to more accurately represent the covariance structure of ordinal data, polychoric correlations—not Pearson correlations—should be analyzed.

Both robust estimators and polychoric correlations are available in EQS, the software used in this dissertation. Although robust estimators were used in all analyses, polychoric correlations could only be used in the measurement model (confirmatory factor analysis) phase of this work. EQS does not allow polychoric correlations when measured independent variables—like the covariates included in this model—are used. As a result, some error may exist within the parameter estimates of the structural components of the model.

Multi-level data. The analyses documented in this study consider covariance and mean structures at one level: the student. In fact, respondents in this study were participating in living learning programs that were nested within institutions. (Indeed, in some large programs that spanned more than one residence hall, students were participating in programs nested in *communities* further nested in institutions.) As a result, this study cannot disentangle program-level effects from effects that operated at the level of the institution. My choice to analyze all data at a single level was done largely out of necessity. Particularly in the race/ethnicity-focused analyses, the low absolute number of cases available for analysis would have made multi-level modeling impossible.

Construct validity. This is the first study to have used confirmatory factor analytic techniques on NSLLP data. Although the NSLLP's original measurement scales were constructed by experts with an eye toward face validity, it is perhaps not surprising that the modeling process described here did not—at least with the samples used—find strong connections between all RES-B items and the latent factors they were meant to represent. As items were removed from latent constructs as originally hypothesized in the 2004 and 2007 NSLLPs, it may be that statistical consistency was improved at the expense of construct validity.

For example, the original academic/vocational peer conversation (ACADPEER) factor piloted in the 2004 NSLLP contained four items, one of which involved "talking about future plans and career ambitions." However, due to low loadings by 2007, that item was removed from ACADPEER factor. Despite its removal, the name of the measurement scale never changed—even though it no longer consisted of any career-related items. In this dissertation, I employed the latent construct names originally developed by the 2004 and 2007 NSLLPs, despite having dropped several items during the process of CFA. My failure to consider renaming those factors to better characterize the constructs they represent seems, in retrospect, a notable omission. As such, care should be taken to remember what each factor does—and does not—represent in this study.

Implications for Researchers

The implications of the present study for education researchers fall in to four broad categories, including improvements to: (a) study designs, (b) institutional sampling, (c) the use of administrative data, and (d) measurement.

Design. The "gold standard" in any quantitative research is the classic experiment. Within the context of L/L programs, this would imply measurement of student characteristics well-prior to entry to the institution (perhaps even before a student's *choice* of institution, in case that choice is conditioned in some manner on his or her desire to enter a L/L program), random assignment of students to the L/L or TRH environment, rigorous control over the specific conditions of both, and one or more measures of the outcome of interest post-exposure. Of course, with the exception of the latter, these criteria are unlikely to be feasible in practice. I consider improvements to each briefly below.

As noted earlier, the NSLLP used proxy pre-tests to control for students' eventual outcome scores. Despite not being *genuine* pre-tests for the outcome of interest, these items elicited interesting findings about the relationship between motivation and achievement. However, the tenor of that relationship in the presence of an actual outcome pre-test remains unclear. This implies that researchers should continue to find ways to implement true tests of students' pre-entry levels of learning and development (perhaps at orientation or during a pre-scheduled institutional "assessment day"), but continue to include items that characterize the extent to which students intend to direct their energies toward increasing complexity.

The absence of random assignment in the present study means that variables not explicitly included in the model, in particular those that are simultaneously related to choice of L/L environment and learning and development, are biasing its results. While there are several quasi-experimental methods available to social science researchers to address this concern, only two seem potentially likely for those whose interests lie in better understanding interventions at the college level.

The first is through the use of a wait-list control. In this approach, researchers intentionally "over-recruit" for the program that they wish to study, randomly selecting only a subset of applicants for participation and relegating the other to the control condition. Some may find this ethically dubious, and it should be noted that this approach only really provides information about students who had an intention to participate in the initial treatment. A third group, "no intent to receive treatment," would be needed to truly understand the intervention's effect. Of course, the use of wait-list control depends upon demand: unless a housing officer is willing to have empty beds in a living-learning program, an unlikely situation, demand must outstrip supply two to one.

Alternatively, researchers can opt to leave assignment to treatments uncontrolled, but attempt to address selection statistically. The most common method for doing so is through the use of propensity scores. This promising approach is discussed in more detail below, as I discuss directions for future L/L research.

Statistical sampling of institutions. Because no definitive national registry of L/L programs exists, it is impossible to attempt to draw the random sample of programs that would be needed to eliminate institutional self-selection into a study like the NSLLP. The general problem—that I am left studying programs that want to be studied (and that we want to include), and that their motivation to participate threatens both internal validity and my ability to generalize to the larger population—is likely pervasive, and certainly difficult to surmount. As an example for L/L-related research, analysts could consider the following sampling strategy: (a) screening all postsecondary institutions (or a specific subset to within them to which they want to generalize) to ascertain if they have residential life programs, and, if so, whether L/L programs are among them, (b) stratifying the resulting eligible sample on the basis of meaningful institution-level characteristics (e.g., institutional sector, Carnegie classification, or region), and (c) randomly selecting potential institutional participants from within those strata.

Administrative data. As noted earlier, the computation of EIs was hindered in this study by an unexpectedly high (18%) rate of missingness on the variable representing the specific program in which a student participated. Given the centrality of that data element for the NSLLP generally and this study in particular, it seems evident that strategies should have been developed to ensure that it was collected accurately. The implication for researchers is clear: if administrative data exists on data elements key to your work, better to solicit it from institutions directly than to rely solely on student reports.

Measurement. As noted earlier, my *a priori* measurement assumption was that the factor structure identified in the 2007 NSLLP Technical Report (Inkelas, Brower, & Associates, 2008) was accurate, based upon several improvements made after the 2004 administration. That structure, developed through exploratory factor analysis, asserted the two-dimensional nature of peer conversations (i.e., academic/vocational and socio-cultural), faculty interactions (i.e., course-related and non-course related), and hall climates (i.e., academically and socially supportive). While the authors' dimensionality assertion was not tested here, CFA was undertaken to verify survey items were satisfactory indicators of the latent constructs they purported to measure.

As presented at some length in Chapter 6, CFA results indicated that prior exploratory factor analyses may have yielded inappropriate conclusions for a subset of items. Significant cross-loadings were noted for several questions, resulting in those items being dropped from the analysis. A notable example is the original assignment of an "intellectually stimulating" residence hall environment to the factor representing a *socially* supportive hall climate. Additionally, some items exhibited differential functioning *within* a specific analytic group. For example, "ease of forming study groups" loaded much more strongly on the academic transition outcome measure for Asian/Pacific Islander L/L participants than it did for their TRH peers.

While the reason for the prior finding (and others like it) is not immediately evident, there are two key implications for researchers. First, producers of data should seek to confirm the factor structure of their instruments and, as necessary, engage in cognitive testing with respondents to help ensure construct validity. Second, consumers of data should be willing to interrogate the scales and/or constructs they are provided as part of a secondary data analysis. Failure to do so can impugn the face validity of a research effort, and might yield inappropriate findings.

Implications for Practitioners and L/L Theoreticians

Taken together, these findings have implications for scholar-practitioners interested in how (and whether) L/L programs can be used to improve student outcomes. Below, I consider three such implications, including those that bear on: (a) the importance of equity, irrespective of whether participation in L/L programs is genuinely related to outcome attainment, (b) the design and implementation of programs, and (c) our collective understanding of how and why L/L programs "work." I conclude with a call for increased local assessments to determine the value-added of L/L programs at the campus level.

Inequity appears to exist, but its effects are unclear. To greater or lesser degrees, the equity indices reported in this study suggest that students from various racial/ethnic backgrounds and SES groups are disproportionately dispersed across types of L/L programs. Despite the occasional over-representation, the low overall participation rate of racial/ethnic minority students (~24%) seems to dictate that most L/L programs will be overwhelmingly White.

Simply on the basis of the LMM findings that suggest that L/L and TRH students within those groups experience key living and learning environments differently, the paramount implication for practitioners is to engage in a continuous process of assessing equity in program participation, identifying barriers to participation, and remediating those barriers when found. For example, each L/L group reported mean levels of socio-cultural peer conversations and socially supportive residence hall climates that were greater than their TRH peers. Another two out of three L/L groups reported mean levels of academic/vocational peer conversations and academically supportive residence hall climates greater than their TRH peers.

However, given that in no case could I find evidence of differences in learning and development outcomes between L/L and TRH participants, a stark question emerges: is the unequal distribution of a resource truly bad if that resource is not of objective value? And, in the light of my finding that some group × thematic type combinations had equity indices greater than one, are all forms of inequity—both over- and under-representation—equally problematic under this scenario?

I would contend that, at least in the case of under-representation, the answer remains "yes." First, although differences in outcomes were not observed, it well may be that measures were poor or that students accrued other outcomes not considered in this analysis. Furthermore, it is not particularly difficult to construe the key living and learning environments measured here as educational goods in and of themselves, akin to Astin's (1991) notion of an intermediate outcome, In that regard, L/L participants at least appear to have accrued gains not seen by their TRH peers (this point is discussed in more details in *Counterfactual analysis*, next section). Even if one were to adopt the apparently untenable position that there is *no* benefit to L/L participation on any dimension, there is the issue of perceived inequity: to the extent that students ascribe disproportional representation within an ostensibly enriched educational environment, pre-existing beliefs about domination and subordination may be reinforced (n.b., discussions about tracking in elementary and secondary education, such as Hallinan, 1994; Loveless, 1999; Oakes, 2005).

The implications of over-representation are somewhat less clear. Initially, problems stemming from the appearance of under-representation might also be triggered by apparent over-representation. (Indeed, NSLLP manuscripts in preparation for publication suggest that students are keenly aware when some groups, particularly dominant ones, are clustered within certain types of high-impact programs.) Perhaps more problematic is the situation in which a vulnerable student population is intentionally over-represented, but the ostensible benefits of the program do not actually accrue, such as might be the case in a transition program for an underserved population.

Many thematic types, comparatively fewer outcomes. The driving forces behind the first research question were the promotion of comparability and the identification of salient outcome measures for more targeted study. However, a key finding from that effort was the discovery that although the number of thematic types of L/L programs appears to be quite large, the number of truly salient outcome measures across program types appears to be quite small.

From a strictly practical standpoint, this suggests that while a L/L program's theme may be its "hook" to engaging students—and as such can be as distinctive as needed to recruit participants—a program's academic, social, and residential components can be comparatively generic, tailored to a small number of potential learning outcomes and based upon a set of best practices known to promote them. For the researcher, this suggests that thematically variant programs *can* be meaningfully compared, so long as attention is paid to identifying an outcome that is shared by all. This finding may also serve to sharpen the focus of living-learning assessment efforts: while the NSLLP sought to measure fifteen distinct student learning outcomes, it seems clear that a much more parsimonious set of outcomes measures can be developed, each of which may be able to be explored at greater depth.

Which environments are associated with outcomes remains murky. The underpinnings of this study's conceptual framework include Astin's (1991) IEO model, which states the general form of college impact, and Pascarella and Terenzini's (1980) structural mediation model, which privileges peer and faculty interactions and hall climate as the primary causal forces within residential environments. As noted in Figure 3.1, my conceptual framework dictated that six key living and learning environments—academic/vocational and socio-cultural peer conversations, course-related and non-course-related faculty interaction, and academically and socially supportive residence hall climates—would be related to any given residential outcome of interest.

In the present analysis, only one living and learning environment was consistently related to the outcomes I studied, irrespective of residential arrangement: an academically supportive residence hall climate. While it is true that other environments were idiosyncratically related to particular outcomes (see the outcome-by-outcome discussions above), this study cannot be read to support the notion that all six "key" living and learning environments are deserving of such a label. Indeed, non-course-related faculty mentorship was found to have a negative relationship to low-SES TRH students' diversity appreciation, and no NSLLP-based study has found *any* relationship between this construct and a student outcome of interest. This latter finding may be due to generally low levels of faculty interaction—course-related or otherwise—reported in the 2007 NSLLP overall, where the means of faculty interaction measures were generally only 60% of measures for interactions with peers (Inkelas, Brower, & Associates, 2008).

While these findings generally support the structural mediation model (Pascarella & Terenzini, 1980), they suggest some combination of the following: (a) there may be some minimum threshold below which an environment cannot influence student development, (b) some environments bear no relationships to some outcomes, irrespective of their level, (c) students are differentially disposed to environments and outcomes, or (d) true environmental effects are obscured by limitations inherent in the study. While the latter is certainly the case to some degree, the former implications are worthy of consideration.

It seems appropriate to question the extent to which L/L environments are truly "powerful," at least as currently implemented and in the aggregate. Among Asian/Pacific Islander students, no mean differences in environment scores were noted on three of the six living and learning environments (course-related faculty interaction, non-course-related faculty interaction, and academically-supportive hall environment). Such was the case for two of six environments for White students (academic-vocational peer conversations and non-course-related faculty interaction), and one of six for low-SES students (course-related faculty interaction). So long as one believes the residential environments identified here have the capacity to influence student development at some level, the implication for educators is persevere in their efforts to strengthen programs with an eye toward increasing the power of these environments, and to continue assessment efforts to identify the "dose" at which the outcome needle is moved.

To the extent that some environments were not shown to be positively related to *any* student outcome in the present study, residential educators and theoreticians may choose to not focus on efforts to promote them. A likely candidate is non-course-related faculty interaction, which as operationalized in this study includes the extent to which students seek out faculty assistance in addressing personal or academic concerns or have informal social interaction. Prior NSLLP-based research has yet to find a link between this construct and an outcome of interest, and the present study questioned an observed negative relationship between non-course-related faculty mentorship and academic transition for one subset of students. To be sure, conventional wisdom and prior research dictates that non-course-related faculty interaction *should* promote student development. That I have yet to detect this "elusive" relationship may be due to misspecification of the phenomenon (i.e., perhaps the relationship does not hold *as* true in residential environments as in others), the use of a largely first-year sample, or poor measurement tools.

The motivational character of the proxy pre-tests used in the present study and their consistently positive relationship to the student outcomes studied here suggest that, despite residential educators intentions, not all students will be equally disposed to attain learning and development outcomes believed to be important. This may also extend to students' capacity to benefit from environments believed to be educationally powerful, as one plausible interpretation of this study's findings vis-à-vis non-course-related faculty interaction is that, given two of the three groups considered here consisted entirely of first-year students, they simply lacked some dimension of interpersonal competence or confidence to benefit from informal faculty contact. The implication for residential educators is somewhat unsatisfying: it may be that the students who could benefit the most from L/L programs are those who, before participation, already want to benefit. To the extent that L/Ls are meant to be enrichment programs, this may be palatable. To the extent that they are meant to be remedial (in its most positive sense), this may be personally and professionally challenging, and suggests the need to simultaneously attend to outcomes and students' motivation to attain them.

The importance of local assessment. Although a strength of the NSLLP is its bringing-together of data from a national sample of L/L programs, neither its findings nor those of this dissertation can serve as substitutes for local assessments as bases for local judgments. Institutional contexts are simply too specific, and even *within* institutions year-by-year variability due to changing student, staff, and faculty populations can make generalization difficult.

What this research can do, however, is serve as a guide to practitioners as to the specific elements of their program to explore in more detail. The present study, for example, suggests that practitioners should begin to assess participatory equity in their residential programs and engage in purposeful discussions as to whether the students who are participating in a program are those for whom such participation could be most valuable. It also suggests that L/L programs *are* capable of positively influencing a subset of residential environments, each of which may be of value irrespective of their relationship to terminal NSLLP outcomes. As such, practitioners could use qualitative or quantitative methods to identify "environmental exemplars" from among their own portfolio of L/L programs, using assessment results to spread promising practices across L/L and TRH environments.

Suggestions for Future Research

The limitations and implications identified above suggest any number of directions for future research. Three lines of future analysis seem particularly important, as they represent significant impediments to better understanding the true benefits of L/L participation.

Counterfactual analysis. To the extent that random assignment to L/L programs is not likely to take place at scale, techniques that allow researchers to estimate what *would have happened* to TRH students had they been exposed to the L/L environment is an important next step in understanding their effectiveness. Absent that work, accurately estimating treatment effects net of confounding selection bias is virtually impossible. Two of the most common methods of counterfactual analysis include propensity score matching and instrumental variable approaches (Morgan & Winship, 2007).

Within the context of L/L research, propensity score approaches would attempt to estimate counterfactual effects by first predicting the likelihood that any given student would participate in an L/L program, and, on the basis of those predicted probabilities, compare the outcomes of highly similar treated and untreated students. Propensity scores are created via some binary linear model (e.g., logit or probit), where treatment participation is predicted by student-level characteristics. Those scores are then used either to weight cases in subsequent analyses (good for small datasets where each case must be retained for analysis) or to create more homogeneous samples by matching treatment cases to one or more control cases with similar propensity scores (good for large datasets where some cases can be discarded).

The use of propensity scores in general is complicated by difficulty in ensuring that the propensity model itself accounts for important sources of bias beyond already observed covariates (Morgan & Winship, 2007). Unfortunately, the 2007 NSLLP RES-B represents a missed opportunity in advancing the use of propensity scores in this domain of inquiry: While it included a strong set of items that attempted to understand why students opted to participate in L/L programs, no analogous set of items was administered to TRH students. Absent those, the primary candidates for inclusion in a propensity model are the same covariates already included in the conceptual framework (e.g., race/ethnicity, gender, pre-test scores, or ability), and bias may not be appreciably reduced. Future research, then, should include not only the exploration of this method, but also the identification of a solid model for predicting treatment.

Instrumental variables (IV) approaches are another technique designed to help address selection bias. Generally, they seek to isolate the causal effect of a treatment on an outcome by identifying variables (instruments) that are strongly associated with the treatment but unrelated to other factors that influence the outcome, such as covariates or error. Properly specified, any residual relationship between the instrument and the outcome variable must, then, be evidence of causality, and a more accurate estimate of the treatment effect can be calculated (Morgan & Winship, 2007).

Pike, Hansen, and Lin (2010) have recently demonstrated an IV approach to measure the effect of participation in non-residential, disciplinary-themed learning communities on first-year students' fall grades. The authors identified two instruments—participation in a summer bridge program and having chosen a major—for participation, on the basis that summer bridge program participants had been encouraged to participate in learning communities and that those who had chosen a major would be more likely to participate in a discipline-based program, but neither bridge participation nor major declaration would drive Fall grades. Prior to adjustment, Pike et al. reported that there was a statistically positive relationship between learning community participation and fall grades. After adding IVs to their analysis—presumably attenuating for selection bias—the authors could not detect a treatment effect.

Although future research should explore IV approaches as a method to attenuate selection bias, it seems particularly challenging in the context of the NSLLP. It is difficult to imagine what factors could be identified that would drive L/L participation but not affect outcomes of interest. Taking a cue from Pike et al. (2010), one could consider using major declaration as an IV for participation, although it seems tenable that having chosen a major would be directly related to the developmental outcomes the NSLLP generally hopes to assess, violating the tenets of IV methods. Given that so-called "weak instruments" can introduce more bias into an analysis than they remove (Kennedy, 2003), researchers who hope to use this particular quasi-experimental method will have to exercise both creativity and caution in their analyses.

Motivation. Better understanding the relationship between motivation, L/L participation, and student outcomes is critical to unpacking the efficacy of the intervention itself. First, one can imagine that there is some facet of motivation that disposes a student to want to participate in an L/L program compared to the TRH environment. Since these environments are generally thought of as being enriched in some capacity (e.g., courses, seminars, relationships, activities), they are likely to appeal to a particular subset of students who are qualitatively different from their peers who prefer the TRH environment. To the extent that those differences are directly or indirectly related to outcomes, they confound my ability to measure the impact of L/L participation.

That first motivational effect—motivation to participate—may be distinct from a second motivational driver—motivation to engage with the treatment or to benefit from it. As noted earlier, Pace (1984) identified the connection between a student's direction of energy and the gains he or she ultimately accrues, and it follows generally from developmental theory that individuals must exhibit at least some openness or readiness for growth if it is to occur. In two of the three student groups studied, no statistically significant difference was noted between L/L and TRH students' pre-test means. The extent to which that lack of difference at entry corresponded to a lack of differences at the time of measurement should continue to be explored. More explicit measures of willingness to engage with key living and learning environments could also be of value, particularly among those where little difference and/or effect was found (e.g., non-course-related faculty interaction).

Tying learning and development outcomes to other critical outcomes. The learning and development outcomes in this study included smooth academic transition and diversity appreciation. Presumably, the latter is important for students' functioning in increasingly diverse university communities and the broader American society. The former is valued for its capacity to influence students' persistence during the first year of college.

While these outcomes are seemingly good in and of themselves, their relationship to the paramount outcome of college attendance—degree attainment—was not measured in this cross-sectional study. Future efforts should attempt to develop longitudinal research programs that can associate L/L participation with more distal gains, such as persistence, satisfactory academic progress, and, ultimately, credentialing. The advent of longitudinal data systems—and more advanced analytic techniques—stand to demonstrate these programs' contribution to the outcome upon which all stakeholders can agree: attaining a degree, with all the public and private goods that generates.

Conclusion

Living-learning programs have been developed in response to calls to better serve students, and outcomes ranging from greater student persistence to developmental gains have been ascribed to them (Blimling, 1993; National Leadership Council, 2007; Pike, 1999; Rowan-Kenyon et al. 2008; Stassen, 2003; Terenzini, Pascarella, & Blimling, 1996). Attaining these benefits may be particularly important for student subtypes who are less successful within postsecondary education, including underrepresented racial/ethnic minority students and students from lower socioeconomic strata (Kinzie, Palmer, Hayek, Hossler, Jacob, & Cummings, 2004; Knapp, Kelly-Reid, and Ginder, 2009; Leslie & Brinkman, 1988). If participation in L/L or other high impact programs *is* associated with meaningful benefits, then diffusing the best elements of those programs to students most in need is in the best interest of students, institutions, and society generally (n.b., Kuh et al., 2007).

The efficacy of L/L programs is far from proven, however (Pascarella & Terenzini, 2005), and their seeming widespread adoption may be a function of external pressures to "do something" to improve educational outcomes (Inkelas, Vogt, et al., 1996; Shapiro & Levine, 1999) or to maintain institutional viability by bolstering social legitimacy (Powell & DiMaggio, 1991; Scott, 1987, 2003). The present study sought to shed light on whether students from diverse backgrounds were equitably distributed across types of L/L programs, and, for those students who were underrepresented, whether their outcomes differed from their TRH peers. Were a difference to have been noted, it would be taken as evidence that increasing participation in that form of program could be of particular benefit to students.

Underrepresentation across L/L types were found in this study (as was over-representation, in some instances), although a low absolute number of underrepresented students generally complicated analysis. Differences in student outcomes were not detected by this study, although, to greater or lesser extents, L/L participants reported higher mean scores on several measures of key living and learning environments. While I conclude that L/L and TRH participants may not have exhibited differential scores on NSLLP learning and development scales, it may well be that the environmental differences they reported are independently beneficial.

While it is not evident that I have found “what works” (p. 70) in high impact practices sought by Kuh et al. (2007), it may be—even in light of the limitations identified in this chapter—that I have confirmed something is *happening* with L/L programs. Whether that happening can be made to translate to outcome differences, and whether those outcome differences translate to tangible benefits for students, remains to be seen. Disciplined inquiry, including the future research suggested previously, may one day help us find out. Hopefully, disciplined, equitable practice will help us apply what we learn to help the students who need that assistance the most.

Appendix A

The Residence Environment Survey – Baseline (RES-B)

2007 Residence Environment Study 0

NSLLP 2007
NEW BASELINE
QUESTIONNAIRE

ABOUT YOU

1. **What is your gender?** (Circle one.)
   1. Male
   2. Female
   3. Transgendered

2. **Please indicate your sexual orientation.** (Circle one.)
   1. Bisexual
   2. Gay or Lesbian
   3. Heterosexual

3. **Are you:** (Circle all that apply.)
   1. African American/Black (not of Hispanic origin)
   2. Asian or Pacific Islander (includes the Indian sub-continent)
   3. American Indian or Alaskan Native
   4. Hispanic/Latino (Spanish culture or origin)
   5. White/Caucasian (Persons not of Hispanic origin, having origins in any of the original peoples of Europe, North Africa, or the Middle East)
   6. Race/ethnicity not included above

4. **Were the following individuals in your family born in the United States?** (Circle one response for each)

1. You	Yes	No
2. Mother	Yes	No
3. Father	Yes	No

[If only "You" in Q4 is marked "no"]
Please indicate the country where you were born: __________

[If only "Mother" in Q4 is marked "no"]
Please indicate the country where your mother was born: ____

[If only "Father" in Q4 is marked "no"]
Please indicate the country where your father was born:______

[If only "You" and "Mother" in Q4 are marked "no"]
Please indicate the country where you were born: __________
Please indicate the country where your mother was born: ____

[If only "You" and "Father" in Q4 are marked "no"]
Please indicate the country where you were born: __________
Please indicate the country where your father was born:______

[If only "Mother" and "Father" in Q4 are marked "no"]
Please indicate the country where your mother was born: ____
Please indicate the country where your father was born:______

[If only "You," "Mother, and "Father" in Q4 are marked "no"]
Please indicate the country where you were born: __________
Please indicate the country where your mother was born: ____

Q4 cont.

Please indicate the country where your father was born:______

[If "You" in Q4 is marked "no" (regardless of mother and fat

5. **Which of the following statements applies to you? (Circle one)**

I came to the United States:
- Before age 6
- Between ages 6-12
- Between ages 13-17
- After age 17

6. **What is your current religious affiliation?**

____

7. **How important is your religion in your life?** (Circle one.)

1. Not at all important	3. Important
2. Somewhat important	4. Very important

8. **How would you describe your political views?** (Circle one.)

0. No political viewpoint	3. Middle of the road
1. Very liberal	4. Slightly conservative
2. Slightly liberal	5. Very conservative

9. **What is the highest level of education completed by one or both of your parent(s) or guardian(s)?** (Circle one in each column, if applicable.)

	Father or Male Guardian	Mother or Female Guardian
Don't know	0	0
High school or less	1	1
Some college	2	2
Associates degree	3	3
Bachelors degree	4	4
Masters degree	5	5
Doctorate or professional degree (JD, MD, PhD)	6	6

10. **What is your best estimate of your parents' total income last year? Consider income from all sources before taxes.** (Circle one.)

1. Less than $25,000	6. $125,000 to $149,999
2. $25,000 to $49,999	7. $150,000 to $174,999
3. $50,000 to $74,999	8. $175,000 to $199,999
4. $75,000 to $99,999	9. $200,000 or more
5. $100,000 to 124,999	

11. What were your average grades in high school? (Circle one.)

1. A+ or A
2. A- or B+
3. B
4. B- or C+
5. C or C-
6. D+ or lower
7. No high school GPA

12. Did you take the SAT and/or ACT? (Circle one for each.)

SAT	Yes	No
ACT	Yes	No

[Skip to SAT and/or ACT question, depending upon answer.]

13a. If you took the SAT, which version did you take?

1. SAT Critical Reading, Math, and Writing
2. SAT Verbal and Math

If above = 1,

Please indicate your scores:

Cumulative: ____________ Don't know

Critical Reading: ____________ Don't know

Math: ____________ Don't know

Writing: ____________ Don't know

If above = 2,

Please indicate your scores:

Cumulative: ____________ Don't know

Verbal: ____________ Don't know

Math: ____________ Don't know

13b. If you took the ACT, please indicate your score:

Cumulative: __________ Don't know

14. What is your current class level? (Circle one.)

1. First year
2. Sophomore
3. Junior
4. Senior
5. Graduate student
6. Other

15. Did you receive financial aid in 2003-2004 in the form of: (Circle all that apply.)

0. Not receiving financial aid
1. Loans

Q15 cont.

2. Need-based scholarships or grants
3. Non-need-based scholarships or grants
4. Work-study
5. Athletic scholarship
6. Other: ____________________

16. How many majors do you currently have?

0. Undecided/undeclared
1. 1
2. 2
3. 3 or more

[IF Q16 = 0, skip to Q26)
[IF Q16 = 1, go to Q17, then go immediately to Q18)
[IF Q16 = 2 or 3, go to Q17.5] [These students do Q17.5/Q17.7 thru 22 twice only (even if they have more than 2 majors (Q16 = 3)]

17. Please indicate your current major. [Go to Q18]

Provide major categories.

17..5. Please indicate current major #1. [Go to Q18]

Provide major categories.

17.7 Please indicate current major #2.

Provide major categories.

[IF Q17 or Q17.5 or Q17.7 = ANYTHING EXCEPT UNDECLARED]

18. Was [NAME OF MAJOR] your major when you started college? (Circle one.)

Yes No

19. How many classes in [NAME OF MAJOR] have you completed? (Circle one.)

None 1 2 3 4 5 6+

20. How many semesters, quarters, or terms have you been a [NAME OF MAJOR] major? (Circle one.)

None 1 2 3 4 5+

21. How likely is it that you will complete a degree in [NAME OF MAJOR]? (Circle one.)

1. Not at all likely
2. Somewhat likely

Q21 cont.

3. Likely
4. Very likely
5. I have already completed all requirements for this major.
9. Not sure

[IF Q18 = YES for student with one major (Q16=1), go to Q27] [IF Q18 = YES for student with multiple majors (Q16 = 2 or 3) and this is for 1st major, go to Q17.7 to indicate major #2, then go thru Q18 thru Q22 again with 2nd major, then go to Q27.] [IF Q18 = YES for student with multiple current majors (Q16 = 2 or 3) and this refers to 2nd major, go to Q27]

[IF Q18 = NO]
22. When you first started college, did you intend to major in [NAME OF MAJOR]?

1. Yes.
2. No, I changed from another major.
3. No, I was undecided when I began college.
4. No, I added this major in addition to another. *[only show if Q16 = 2 or 3]*

[IF Q22 = 1 for student with one current major (Q16=1), go to Q27] [IF Q22 = 1 for student with multiple current majors (Q16 = 2 or 3) and this refers to 1st major, go to Q17.7 to indicate major #2, then go thru Q18 thru Q22 again with 2nd major, then go to Q27.] [IF Q22 = 1 for student with multiple current majors (Q16 = 2 or 3) and this refers to 2nd major, go to Q27]

[IF Q22 = 2 for student with one current major (Q16 = 1), do Q22.5, Q23, Q24, Q25] [IF Q22 =2 for student with multiple current majors (Q16= 2 or 3) and this refers to 1st major, do Q22.5, Q23, Q24, Q25, then go to Q17.7 to indicate 2nd major] [IF Q22 = 2 for student with multiple majors (Q16 = 2 or 3), and this refers to 2nd major, do Q22.5, Q23, Q24, Q25]

[IF Q22 = 3 for student with one current major (Q16=1), go to Q27] [IF Q22 = 3 for student with multiple majors (Q16 = 2 or 3) and this refers to 1st major, go to Q17.7 to indicate 2nd major, then go through Q18 thru Q21 again with 2nd major, then go to Q27]

[IF Q22 = 4, go to Q27]

22.5. When you first started college, what was your intended major?

Use same list as for Q17, add "undeclared."

23. How many semesters, quarters, or terms were you a [NAME OF MAJOR from Q22.5] major before you changed your major?

None 1 2 3 4 5+

24. Why did you discontinue your pursuit of [NAME OF MAJOR from Q22.5]? (Circle all that apply)

1. My institution required choosing a major when I entered college, and I subsequently chose a different major.
2. I received information about a more interesting major.
3. It takes too long to finish degree in this major.
4. Faculty in this major were not supportive.
5. I could not meet GPA requirements.
6. The coursework in this major was too competitive.
7. There are too few women in this field.
8. There are too few minorities in this field.
9. I received inadequate or inappropriate advising.
10. I didn't think there would be jobs for me when I graduated.
11. I didn't want this major in the first place.
12. I lost interest in the subject area of my major.
13. Other: ___________________________

25. How many times, since your first major, did you change your major?

None 1 2 3 4 5+

[IF Q16 = 2 or 3 and this is for 1st major, go to Q17.7] [IF Q16 = 2 or 3 and this is for 2nd major, go to Q27]

[IF Q22 = UNDECLARED]
26. What major are you considering most seriously? [Same list as in Q17, add "I don't know" as an option]

27. What is your cumulative GPA on a 4-point scale? (Please enter the number (e.g., 3.6) in the space below.

____

BEFORE COLLEGE

28. Thinking back to *before you started college*, please rate how important you imagined these aspects of college would be. (Circle one response for each.)

1 = Not at all important **3 = Important**
2 = Somewhat important **4 = Very important**

Appendix B

Measurement Scales in the NSLLP, As Described by Inkelas, Brower, and Associates (2008)

Factor Item	α
Academic/Vocational Peer Conversations	.806
Shared your concerns about classes or assignments [with peers]	
Discussed something learned in class [with peers]	
Talked about current news events [with peers]	
Socio-cultural Peer Conversations	.885
Held discussions with students whose political opinions were very different than your own	
Held discussions with students whose religious beliefs were very different than your own	
Discussed major social issues, such as peace, human rights, and social justice [with peers]	
Discussed your views about multiculturalism and diversity [with peers]	
Held discussions with students whose personal values were very different than your own	
Course-Related Faculty Interaction	
Visited informally [with a faculty member] before or after class	.832
Made an appointment to meet in [a faculty member's] office	
Asked [a faculty member] for information related to a course you were taking	
Worked on a research project [with a faculty member]	
Non-course-related Faculty Mentorship	.738
Discussed personal problems or concerns [with a faculty member]	
Discussed academic problems or concerns [with a faculty member]	
Visited informally [with a faculty member] during a social occasion	
Supportive Hall Academic Climate	.800
My residence hall clearly supports my academic achievement	
Most students in my residence environment study a lot	
I think it's easy for students to form study groups in my residence environment	
I think the staff in my residence environment spend a great deal of time helping students succeed academically	
Supportive Hall Social Climate	.878
Students in my residence environment are concerned with helping and supporting one another	
I find that students in my residence environment have an appreciation for people from different religious backgrounds	
Life in my residence environment is intellectually stimulating	
I find that students in my residence environment have an appreciation for people from different races or ethnic groups	
I would recommend this residence environment to a friend	
I see students with different backgrounds having a lot of interaction with one another in my residence hall	
I have enough peer support in my residence environment to succeed academically	
Smooth Academic Transition to College	.762
Communicating with instructors outside of class	

Factor Item	α
Ease with seeking assistance or personal help when needed	
Ease with forming study groups	
Diversity Appreciation	.820
Since coming to college, I have learned a great deal about other racial/ethnic groups	
Since coming to college, I have become aware of the complexities of inter-group understanding	
I have gained a greater commitment to my racial/ethnic identity since coming to college	

Appendix C

The Living Learning Programs Survey

National Study of Living-Learning Programs

2007 LIVING LEARNING PROGRAM SURVEY

Please note:

The 2007 NSLLP living learning program survey will be fielded on the World Wide Web. Thus, the layout and format of the final version of the questionnaire will not be the same as this paper-and-pencil version. (For example, questions will not spill over onto subsequent pages, as they do in the paper-and-pencil version.) In addition, the paper-and-pencil version includes instructions for web-coding that will not appear on the final internet version. However, the content and order of the questions on the web version of the questionnaire will be the same as this version.

Please do not reproduce, distribute, or use any portion of this questionnaire in part or in full without the prior permission of the National Study of Living-Learning Programs. See www.livelearnstudy.net or contact info@livelearnstudy.net.

PART I: PROGRAM CHARACTERISTICS

1. Respondents (who will be the primary school contact at each NSLLP participating school) will be presented with a list of Living-Learning programs, as they were included in their sample. This list will allow respondents to work through the programs one by one, in any order they feel is easiest for them. As they complete each program section, the program name will be dynamically removed from the list. This "base" screen will looks something like this:

UNIVERSITY OF MOLDAVIA

- Honors Program
- Healthy Living Program
- Civil Society Program

2. For each of the programs, respondents will be asked to fill out the survey program information form, the questions for which are provided below. (The questions are the same for every program.) All of the questions below will be provided on one screen; respondents will need to use the scroll bar to move up and down the screen. However, this will make it very easy for respondents to print out the entire form for each program and ask another party to fill it out for them. Respondents can go in and out of each program's link as many times as they need to until they indicate that they are done with the program.

LIVING-LEARNING PROGRAMS SURVEY
INFORMATION FORM

[Name of university inserted here]

[Name of l/l program inserted here]

PART II: GENERAL INFORMATION

Opening Screen
Your responses to this survey assist us in understanding best practices in living-learning programs. Given the importance of this information for practitioners nationwide, we value your efforts to provide as complete and accurate responses as possible. You may be contacted by NSLLP staff if more information is needed about your program or programs.

1. Describe the primary goals and/or objectives of this living-learning program (LLP) in 50 words or less.

   [Fill-in space here; cut off response at 50 words.]

Q2. How important are the following outcomes of participation in this LLP? As a result of participation, students will:

1. Not at all important
2. Somewhat important
3. Important
4. Very important

1. Explore the meaning of facts when introduced to new ideas
2. Apply something learned in one class to another
3. Demonstrate growth in ability to critically analyze ideas and information
4. Demonstrate growth in developing values and ethical standards
5. Demonstrate openness to views that are different than their own
6. Learn about people from backgrounds other than their own
7. Volunteer and/or perform community service
8. Report increased confidence in their academic abilities
9. Report increased confidence in interpersonal relationships
10. Feel a sense of belonging to the institution
11. Experience a smoother academic transition to the institution
12. Experience a smoother social transition to the institution
13. Develop a greater enjoyment of challenging intellectual pursuits
14. Develop healthy behaviors around alcohol and other drugs
15. Develop greater self-awareness

3. For how many years has this LLP been in existence?

   1. 1 year

2. 2-4 years
3. 5-9 years
4. 10-14 years
5. 15 or more years

4. Is the focus of this LLP related to science, technology, engineering, or mathematics (STEM)?

1. Yes
0. No

5. Are students from any of the following majors represented in this LLP?
*(Select **all** that apply.)*

1. Agriculture
2. Biological Sciences (Biology, Botany, Zoology, etc.)
3. Computer of Information Sciences
4. Engineering
5. Health, Pre-Health, or Wellness
6. Mathematics & Statistics
7. Physical Sciences (Chemistry, Physics, etc.)

6. How many students participate in this LLP?
(Please provide a number in the space below.)

[Fill-in space here]

7. What is the academic class standing of students who are eligible to participate in this LLP?
*(Select **all** that apply)*

1. First-years/freshmen
2. Sophomores
3. Juniors
4. Seniors
5. Graduate students

8. How long are students REQUIRED to participate in this LLP?
(Select one response.)

0. No required length of time
1. One semester/term only
2. One year
3. Two years
4. Entire undergraduate duration
6. Other

(please specify how long students are REQUIRED to participate in this LLP: ______________________)

9. Where do most of the students in this LLP live?
(Select one response.)

1. Participants encompass the entire capacity of one residence hall.
2. Participants live together in a specific reserved portion of one residence hall, and there are other non-LLP students living in this building.
3. Participants live together in a specific reserved portion of one residence hall, alongside students in other LLPs housed in the same building.
4. Participants live in a self-contained community, but across more than one residence hall.
5. Participants live on campus, but not necessarily in the same residence hall or community.
6. None of the above apply
(Please specify the living arrangements for this LLP: ____________________________)

10. Do some of the students in this LLP live off-campus?

1. Yes
0. No

11. Is this LLP selective (i.e., there are selection criteria from which you base admission to the program) or open admissions (i.e., everyone who is interested can enroll)?
(Select one response.)

1 Open admissions
2. Selective

[If R selects "1", skip to Q13]

12. You indicated in the previous response that this LLP is selective. What criteria do you use to select among applicants?
*(Select **all** that apply.)*

1. Standardized test scores (e.g., SAT, ACT, AP scores)
2. High school GPA
3. High school rank
4. College GPA
5. High school recommendation (e.g., from teachers, counselors, etc.)
6. College recommendation (e.g., from professors, advisors, professional staff, etc.)
7. Prior extra-curricular involvement
8. Completed application
9. Essay written specifically for LLP
10. Other written material produced by applicant (not necessarily for LLP specifically)
11. Interview

12. Major or academic interest area
13. Admission to specific school/college
14. Class standing (e.g., must be a junior to apply)
15. Other
(please specify the criteria you use to select applicants for this LLP: ____________________)

REPORTING STRUCTURE

13. What is the budget source (i.e., source of funding) for this LLP? In your response, consider all costs for the LLP (e.g., salaries, stipends, programming)
(Select one response.)

1. 100% Student Affairs (including Housing)
2. 100% Academic Affairs
3. 50% Student Affairs (including Housing) / 50% Academic Affairs
4. More Student Affairs (including Housing) than Academic Affairs
5. More Academic Affairs than Student Affairs (including Housing)
6. Other
(please specify the budget source for this LLP: ____________________)

14. Are students charged a fee to participate in this LLP?

1. Yes (please specify the amount in the blank provided): $ *[Fill-in space here]*
0. No

15. Is this LLP funded, in whole or in part, by an external grant?

1. Yes (please specify the amount in the blank provided): $ *[Fill-in space here]*
0. No

16. Approximately what is the annual budget (in dollars) for this LLP, above the regular costs if the residence hall did not contain this LLP? Consider all sources, for example: staff and/or faculty salaries, programming, marketing, etc. Write in the amount in the space below.

$ *[Fill-in space here]*

17. What unit(s) does this LLP directly report to?
*(Select **all** that apply.)*

1. Office of Residence Life or Student Housing
2. Other Student Affairs unit (e.g., office of VPSA, Student Activities, etc.)
3. Academic Department or College (e.g., History Dept., College of Engineering)
4. Academic Administrative unit (e.g., Office of Provost, Dean of Undergraduate Studies)

5. Other
   (please specify the units to which this LLP directly reports: ____________________)

18. What is the primary professional affiliation of this LLP's Director (or chief administrator)? *(Select one response.)*

    1. Office of Residence Life or Student Housing
    2. Other Student Affairs unit (e.g., office of VPSA, Student Activities, etc.)
    3. Academic Department or College (e.g., History Dept., College of Engineering)
    4. Academic Administrative unit (e.g., Office of Provost, Dean of Undergraduate Studies)
    5. LLP is overseen by a multi-person "board" or "committee"
    6. Other
       (please specify the primary professional affiliation of this LLP's Director: ____________________)

    If R chose option "5", skip to Q19.

18a. What percent (out of a total of 100%) of the above person's professional appointment is spent directing this LLP?

    [Fill in box, amount may not exceed 100] %

18b. If this LLP has more than one director, what is the primary professional affiliation of the other director? *(Select one response.)*

    1. There is only one director in this LLP.
    2. Office of Residence Life or Student Housing
    3. Other Student Affairs unit (e.g., office of VPSA, Student Activities, etc.)
    4. Academic Department or College (e.g., History Dept., College of Engineering)
    5. Academic Administrative unit (e.g., Office of Provost, Dean of Undergraduate Studies)
    6. Other
       (please specify the primary professional affiliation of this LLP's other director: ____________________)

    If R chose option "1", skip to Q19.

18c. What percent (out of a total of 100%) of the above person's professional appointment is spent directing this LLP?

    [Fill in box, amount may not exceed 100] %

ACADEMIC COURSEWORK

Screen displayed to all respondents before proceeding to question 19

All of the remaining questions on this survey will help us to understand best practices among living-learning programs. Your responses are very important to us.

Please be advised that, if you skip any of the remaining questions, we will assume that your program does not include the components we are querying and will enter a value of "no" or "never" for these question items. If you do not wish for us to make this assumption, please fill in the appropriate response or use the "other" option to record an alternative.

Thank you very much for taking the time to provide as complete of information as you can. We sincerely appreciate your efforts.

19. Does this LLP offer any of the following types of courses?

		Yes	No
19a.	Courses for official academic credit developed and taught by LLP staff	1	0
19b.	Special sections of introductory or large classes (e.g., English Composition, Calculus, Introductory Psychology) taught by academic departments	1	0
19c.	Courses offered by departments or the university-at-large	1	0
19d.	Courses that do not carry academic credit but contain academic content	1	0

[DO IF Q19a = 1]

19a_1. How many courses for official academic credit developed and taught by the LLP staff are offered by this LLP?

1. 1
2. 2
3. 3
4. 4
5. 5 or more

19a_2. Who generally teaches these courses?
(*Select **all** that apply.)*

1. Faculty members
2. Student affairs staff
3. Academic affairs staff
4. Graduate students

5. Undergraduate students

[DO IF Q19b = 1]
19b_1. How many special sections of introductory or large classes taught by academic departments are offered by this LLP?

1. 1
2. 2
3. 3
4. 4
5. 5 or more

19b_2. Who generally teaches these courses?
(*Select **all** that apply.)*

1. Faculty members
2. Student affairs staff
3. Academic affairs staff
4. Graduate students
5. Undergraduate students

[DO IF Q19c = 1]
19c_1 How many courses offered by departments or the university-at-large are offered by this LLP?

1. 1
2. 2
3. 3
4. 4
5. 5 or more

19c_2. Who generally teaches these courses?
(*Select **all** that apply.)*

1. Faculty members
2. Student affairs staff
3. Academic affairs staff
4. Graduate students
5. Undergraduate students

[DO IF Q19d = 1]
19d_1 How many courses that do not carry academic credit but contain academic content are offered by this LLP?

1. 1
2. 2
3. 3

4. 4
5. 5 or more

19d_2. Who generally teaches these courses?
(*Select* ***all*** *that apply.)*

1. Faculty members
2. Student affairs staff
3. Academic affairs staff
4. Graduate students
5. Undergraduate students

FACULTY AND STAFF ROLES

20. Approximately how many faculty play a ***direct*** role in the functioning of this LLP?
(Select one response.)

0. None
1. 1
2. 2-3
3. 4-5
4. 6-10
5. 11 or more

If R chose option "1", skip to Q21.

20a. Do FACULTY MEMBERS at your institution fulfill the following roles in this LLP?
0. No
1. Yes

1. Serve as members of the LLP's steering/advisory board or committee
2. Have live-in roles (live in residence hall with LLP participants)
3. Supervise RAs or other undergraduate staff

20b. How often do FACULTY MEMBERS at your institution fulfill the following roles in this LLP?
0. Never
1. Once or more a year
2. Once or more a term
3. Once or more a month
4. Once or more a week

1. Serve as academic advisors to participants
2. Serve as mentors to participants
3. Conduct social/cultural outings (e.g., going to live performances, museums, etc.)
4. Eat with LLP participants in a campus eatery
5. Conduct special lectures/workshops
6. Facilitate service learning opportunities

7. Run tutoring sessions
8. Perform administrative responsibilities (e.g., program management, student staff training, selection of participants, budget issues)
9. Other
   (please specify what other roles faculty members at your institution fulfill in this LLP: ____________________)

20c. How are faculty who work with this LLP compensated?
*(Select **all** that apply.)*
0. Faculty in this LLP are not compensated
1. Salary
2. Non-salary stipend (e.g., professional development funds, research funds)
3. Course release
4. Housing
5. Meal plan
6. Other
   (please specify in what other ways faculty in this LLP are compensated: ____________________)

21. Are student affairs staff involved in this LLP?
1. Yes
0. No

If R chose option "1", skip to Q22.

21a. Do STUDENT AFFAIRS STAFF at your institution fulfill the following roles in this LLP?
1. Yes
0. No

1. Serve as members of the LLP's steering/advisory board or committee
2. Have live-in roles (live in residence hall with LLP participants)
3. Supervise RAs or other undergraduate staff

21b. How often do STUDENT AFFAIRS STAFF at your institution fulfill the following roles in this LLP?
0. Never
1. Once or more a year
2. Once or more a term
3. Once or more a month
4. Once or more a week

1. Serve as academic advisors to participants
2. Serve as mentors to participants
3. Conduct social/cultural outings (e.g., going to live performances, museums, etc.)
4. Eat with LLP participants in a campus eatery
5. Conduct special lectures/workshops
6. Facilitate service learning opportunities

7. Run tutoring sessions
8. Perform administrative responsibilities (e.g., program management, student staff training, selection of participants, budget issues)
9. Other
(please specify what other roles student affairs staff at your institution fulfill in this LLP: ____________________)

22. Are academic affairs staff involved in this LLP?
1. Yes
0. No

If R chose option "1", skip to Q22.

22a. Do ACADEMIC AFFAIRS STAFF at your institution fulfill the following roles in this LLP?
1. Yes
0. No

1. Serve as members of the LLP's steering/advisory board or committee
2. Have live-in roles (live in residence hall with LLP participants)
3. Supervise RAs or other undergraduate staff

22b. How often do ACADEMIC AFFAIRS STAFF at your institution fulfill the following roles in this LLP?
0. Never
1. Once or more a year
2. Once or more a term
3. Once or more a month
4. Once or more a week

1. Serve as academic advisors to participants
2. Serve as mentors to participants
3. Conduct social/cultural outings (e.g., going to live performances, museums, etc.)
4. Eat with LLP participants in a campus eatery
5. Conduct special lectures/workshops
6. Facilitate service learning opportunities
7. Run tutoring sessions
8. Perform administrative responsibilities (e.g., program management, student staff training, selection of participants, budget issues)
9. Other
(please specify what other roles academic affairs staff at your institution fulfill in this LLP: ____________________)

23. Are graduate student employees involved in this LLP?
1. Yes
0. No

If R chose option "1", skip to Q24.

23a. Do GRADUATE STUDENT EMPLOYEES at your institution fulfill the following roles in this LLP?
1. Yes
0. No

1. Serve as members of the LLP's steering/advisory board or committee
2. Have live-in roles (live in residence hall with LLP participants)
3. Supervise RAs or other undergraduate staff

23b. How often do GRADUATE STUDENT EMPLOYEES at your institution fulfill the following roles in this LLP?
0. Never
1. Once or more a week
2. Once or more a month
3. Once or more a term
4. Once or more a year

1. Serve as academic advisors to participants
2. Serve as mentors to participants
3. Conduct social/cultural outings (e.g., going to live performances, museums, etc.)
4. Eat with LLP participants in a campus eatery
5. Conduct special lectures/workshops
6. Facilitate service learning opportunities
7. Run tutoring sessions
8. Perform administrative responsibilities (e.g., program management, student staff training, selection of participants, budget issues)
9. Other
(please specify what other roles graduate student employees at your institution fulfill in this LLP: ____________________)

24. Are undergraduate student employees involved in this LLP?
1. Yes
0. No

If R chose option "1", skip to Q25.

24a. Do UNDERGRADUATE STUDENT EMPLOYEES at your institution fulfill the following roles in this LLP?
1. Yes
0. No

1. Serve as members of the LLP's steering/advisory board or committee
2. Have live-in roles (live in residence hall with LLP participants)

3. Supervise RAs or other undergraduate staff

24b. How often do UNDERGRADUATE STUDENT EMPLOYEES at your institution fulfill the following roles in this LLP?

0. Never
1. Once or more a week
2. Once or more a month
3. Once or more a term
4. Once or more a year

1. Serve as academic advisors to participants
2. Serve as mentors to participants
3. Conduct social/cultural outings (e.g., going to live performances, museums, etc.)
4. Eat with LLP participants in a campus eatery
5. Conduct special lectures/workshops
6. Facilitate service learning opportunities
7. Run tutoring sessions
8. Perform administrative responsibilities (e.g., program management, student staff training, selection of participants, budget issues)
9. Other
   (please specify what other roles undergraduate student employees at your institution fulfill in this LLP: ____________________)

ACTIVITIES AND RESOURCES

25. What special resources are available only to participants in this LLP (i.e., not available to students who do not participate in this LLP)?
*(Select **all** that apply.)*

1. Academic advisors in residence hall
2. Additional programming funds
3. Classes held in or near residence hall/community
4. Computer labs
5. Counselors in residence hall
6. Faculty offices in residence hall
7. Internship opportunities
8. Residence hall in popular or prime location on campus
9. Scholarships
10. Separate dining facilities
11. Single occupancy rooms
12. Special amenities (e.g., individual bathrooms, complimentary cable TV, etc.)
13. Special facilities dedicated to program theme (e.g., practice rooms, art studios, library)
14. Special residence hall configurations (e.g., suite-style, apartment-style, etc.)
15. Study space
16. Other (please specify what other special resources are offered by this LLP: ________________________)

17. None of the above

26. Which of the following, if any, are curricular or co-curricular activities offered in this LLP? *(Select **all** that apply.)*

0. Not offered
1. Required for participants
2. Optional for participants

a. Academic advising
b. Arts/music performances
c. Capstone experience (e.g., capstone seminar, senior thesis)
d. Career workshops
e. Cultural outings
f. Group projects
g. International programming
h. Internship, field experience, co-op experience, clinical assignment, etc.
i. Intramural or club sports
j. Multicultural programming
k. Orientation programs (designed specifically for LLP students)
l. Outdoor recreation
m. Research project
n. Service learning (i.e., community service done in conjunction with a course)
o. Community service (i.e., not done in conjunction with course)
p. Speaking foreign languages
q. Study abroad
r. Study groups
s. Team/community building activities (e.g., retreats)
t. Tutoring
u. Other (please specify what other curricular or co-curricular activities are offered by this LLP: ______________________________)

27. Which of the following, if any, are awarded upon successful completion of this LLP? *(Select **all** that apply.)*

1. Baccalaureate degree conferred by program
2. Transfer directly to specific major (e.g., matriculation into School of Engineering)
3. Completion of major
4. Completion of minor, citation, or notation on transcript
5. Special awards or certificates of participation (e.g., medallions)
6. None of the above
7. Other (please specify the award granted upon successful completion of this LLP: ______________________________)

If R answered "0" to Q4, the survey is complete. [Non-STEM]

If R answered "1" to Q4, R completes next three questions. [STEM]

ADDITIONAL QUESTIONS FOR STEM-FOCUSED PROGRAMS

Screen displayed to all respondents before proceeding to q28.

The final set of questions refers to any STEM (science, technology, engineering, & mathematics) components of your living-learning programs.

Just a reminder, if you leave any of the remaining questions blanks, we will record your response as "no" or "not offered by program."

28. Which of the following STEM resources, if any, do you offer in this LLP? *(Select **all** that apply.)*

a. STEM academic advising with faculty members in residence hall
b. STEM academic advising with professional staff
c. STEM academic advising with peers
d. STEM career counseling by faculty members in residence hall
e. STEM career counseling by professional staff
f. STEM career counseling by peers
g. STEM career information/materials
h. Mentoring from STEM alumni
i. Mentoring from STEM professional staff
j. Mentoring from peers in STEM disciplines
k. Mentoring from professional in STEM field
l. Mentoring from STEM professor/faculty member
m. Special sections of introductory STEM courses (e.g., Calculus, Biology, Chemistry, etc.)
n. Special laboratory sections for introductory STEM courses
o. Support for laboratory sections
p. Study groups
q. Support groups
r. Tutors
s. None of the above
t. Other (please specify what other STEM resources are offered in this LLP): ____________________

29. Which of the following workshops related to STEM issues, if any, do you offer in this LLP? *(Select **all** that apply.)*

a. Career preparation in STEM fields
b. Graduate school options in STEM
c. Guest speakers from STEM faculty
d. Guest speakers from STEM field
e. Stress management

f. Study skills
g. Test-taking and/or problem solving skills
h. Time management
i. None of the above
j. Other (please specify what other workshops related to STEM issues are offered in this LLP): ________________________

30. Which of the following activities, if any, do students in this LLP participate in ?
(Please answer each question below.)

0. Not offered by program
1. Required for participants
2. Optional for participants

a. Work on a research project with a STEM focus
b. Complete an experiment or project using scientific methods
c. Participate in an internship in STEM field
d. Be a mentor or "buddy" to another student in STEM
e. Be a tutor in STEM courses
f. Attend lecture/presentation by a STEM professional
g. Participate in field trips to a STEM industry site
h. Work with outreach to high school students interested in STEM
i. Participate in STEM professional associations or honor societies (e.g., IEEE, ASME, SWE, NSBE, SHPE)

Appendix D

Codebook and Rater Instructions

Instructions for National Study of Living-Learning Programs Coders

First, let me thank you for agreeing to participate in this task. In addition to helping me complete my own research, the coding scheme you help develop will be used in future living-learning (L/L) scholarship that is designed to aid practitioners and, ultimately, improve the student experience. Your time and efforts are appreciated!

Context.
As you may know, the National Study of Living Learning Programs (NSLLP) is a multi-year, multi-institutional study into the efficacy of L/L programs. The initial NSLLP administration in 2004 surveyed more than 22,000 students at 34 institutions, which among them offered almost 300 L/L programs.

As a result of the 2004 administration, principal investigator Karen Kurotsuchi Inkelas and her associates developed a **thematic typology** of L/L programs, based upon those programs' stated goals and objectives. This typology helped the researchers group similar programs—at least in terms of theme—for subsequent analyses. A total of 14 **primary** types of programs were identified, many with several subcategories.

The following types of programs were identified:

- **Disciplinary Programs**: Disciplinary programs represented the greatest number of L/L programs. This group of programs, while distinct, all centered on one curricular or disciplinary focus. There were seven sub-categories of Disciplinary Programs, including: (a) *Business*, (b) *Education*, (c) *Engineering & Computer Science*, (d) *Health Science*, (e) *Humanities*, (f) *General Science*, and (g) *Social Science*.

- **Cultural Programs**: The second largest category of programs focused on cultural issues. These programs were divided into three sub-categories. *International/Global Programs* provided students with an internationally oriented environment in which many countries and nationalities are studied and celebrated. *Language Programs* afforded their participants the opportunity to learn more about a specific language and culture (e.g., German, Japanese, etc.). *Multicultural/Diversity Programs* focused on domestic diversity issues, which included topics related to race/ethnicity, one specific race or ethnicity, sexual orientation, or the hearing impaired.

- **Transition Programs**: The first-year experience is emphasized in this type of L/L program. Transition programs generally focused on the shift from high school to college, and/or offer students an overall introduction to the university and to college life. Two sub-categories were included in this theme: (a) *New Student Transition Programs* closely adhere to this theme's definition, and (b) *Career/Major Exploration Programs* introduced students to different academic disciplines or vocations in order to assist them in major or career choices.

- **Honors Programs**: These programs provided a rigorous and enriched curricular environment for the university's most academically talented students. Generally, these students were identified initially by their academic records (e.g., high school grades and/or SAT score) and were invited to join the program.

- **Fine & Creative Arts Programs**: These programs celebrated different forms of the fine or creative arts, including visual arts, music, architecture, film, prose, cuisine, and photography.

- **Civic/Social Leadership Programs**: Programs in this category were divided into three varieties. The first sub-category represents *Civic Engagement Programs*, or programs that focused on active participation in the political process or in public service. The second sub-category represents *Leadership Programs*, which focused on public leadership through community service or service-learning. The third sub-category represented *Service-Learning/Social Justice Programs*, which emphasized community service or service-learning initiatives while striving to facilitate greater social responsibility among participants.

- **Women's Programs**: These programs catered specifically to women students. *Women in Leadership Programs* focused on facilitating leadership development through several opportunities, such as Greek life, service learning, and cooperative living. *Women in Math, Science, and Engineering Programs* provided resources for women interested in pursuing majors and careers in science, technology, engineering, and mathematics – fields that have traditionally been male-dominated.

- **Wellness/Healthy Living Programs**: These programs fostered healthy lifestyles through emphases such as substance free residence environments, fitness programs, and/or health education programs.

- **General Academic Programs**: These L/L programs emphasized academic excellence and provide general support for academics in the residence hall. However, they have no particular curricular or co-curricular focus or theme.

- **Residential Colleges**: These programs generally spanned several years of the participants' college experiences, offering a broad range of course offerings in artistic, social, and cultural pursuits. They often emphasized exploration and creativity, and most closely mirrored the classic liberal arts tradition of colleges in the 18^{th} and early 19^{th} centuries.

- **Multi-Disciplinary Programs**: These programs were often broader umbrella organizations which house several smaller communities clustered around a specific theme. An example would be a "living-learning center" that hosts 10 or more smaller programs of varying disciplinary themes. **IMPORTANT: This category does not describe programs that are "interdisciplinary."**

- **Upper-Division Programs**: While most of the other types of L/L programs focus on the first and second years of college, these programs catered directly to junior and senior undergraduates with the intention of providing out-of-class experiences that complement students' academic interests. These co-curricular involvements may have included service learning projects, independent research studies, entrepreneurial business pursuits, and internships.

- **Research Programs**: These programs offered students the opportunity to conduct academic research, either in conjunction with a faculty member's research project or in teams with fellow student participants under faculty supervision.

- **Outdoor Recreation Programs**: These programs emphasized sporting and outdoor skills. This is the only strictly non-academic theme in the L/L typology.

In 2007, the NSLLP was repeated. However, in this administration, the number of programs surveyed nearly doubled. The significant expansion of programs surveyed makes it at least tenable that the original typology may have "missed" one or more types of programs, simply because they were not present in the 2004 sample.

Purpose of your work.
The purpose of this coding effort is to code 2007 programs, using the 2004 typology as a foundation. Because the 2004 typology was never intended to be "exhaustive," you should be guided, but not restricted by, its outlines. Through this process, existing 2004 codes may be refined or deleted, or new codes developed. Our goal is to reach reasonable consensus in the coding effort, keeping in mind that perfect inter-rater reliability is unlikely.

An overview of the coding process.
The coding process will take place in two phases: (a) a single-step trial coding of approximately 10% of the sample, and (b) one or more rounds of full coding. The purpose of trial coding is for raters to develop a familiarity with the coding software. Because real data is used in trial coding, it is possible (and perhaps likely) that one or more new codes will emerge. After trial coding has been completed, some preliminary inter-rater reliability statistics will be calculated, and potential new codes brought forth to the coding group. Depending on the number of new codes identified and preliminary reliability estimates, a group meeting may be needed before full coding can begin.

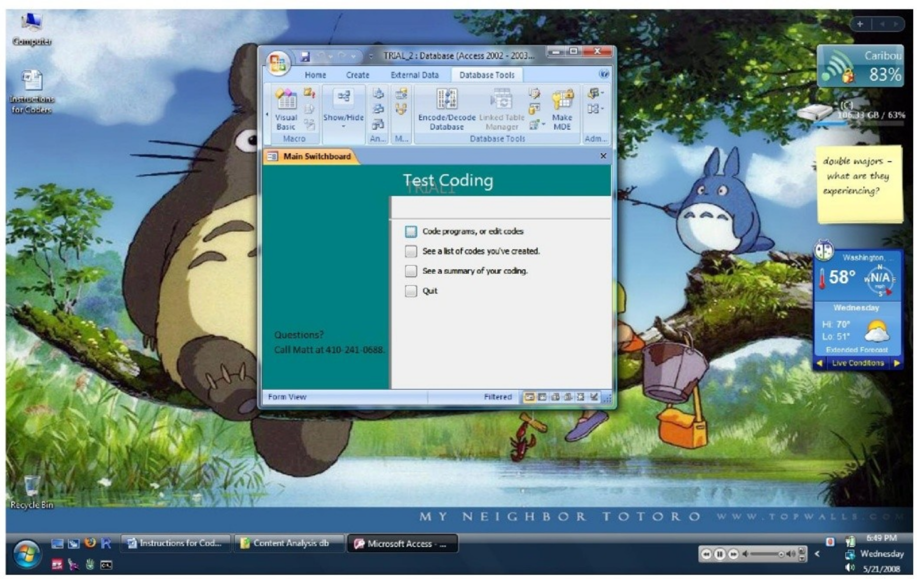
Once trial coding has concluded, full coding will begin. At that time, each rater will receive their portion of the remaining programs to code, along with any revisions to the coding scheme suggested by trial coding. Once the first round of full coding has concluded, inter-rater reliability statistics will be computed and "problematic" programs or codes will be reviewed. Given the number of raters, it is unlike ly that we will achieve consensus after the first round of full coding. Depending upon the number of new codes identified and reliability estimates, a group meeting may be needed, followed by one or more smaller rounds of recoding.

Using the coding software.
You will use a Microsoft Access 2003 database to do your coding. This database will run on computers with either Office 2003 or 2007 installed. Here are the specific steps you should follow while coding:

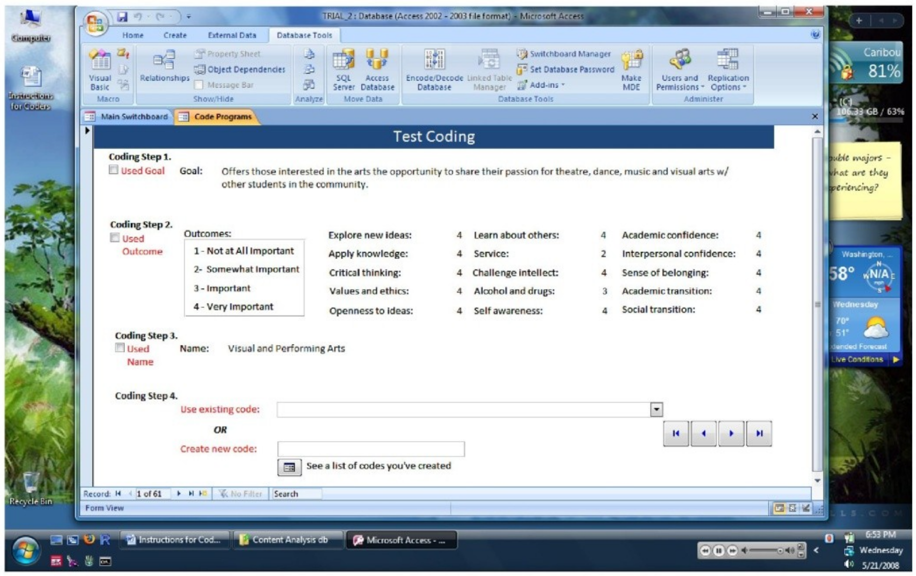
(1) You will receive the coding database via e-mail, in a ZIP file. Save the ZIP file to your desktop, and then double-click on it to open the archive. Drag the database onto your desktop and delete the original ZIP file.

(2) Open the database. Depending upon the version of Office you are using and the settings on your individual computer, you will receive one or more security warnings. You will need to read each warning and dismiss it with the option that will *open* the database. Once the database opens, you may note *another* security warning indicating that "certain content has been disabled." If this is the case, hit "options" and enable it.

(3) You will know you have successfully opened the database when the following screen labeled "Test Coding" appears:



(4) To begin coding, click the first button marked "Code programs, or edit codes." The following screen will appear:



(5) You will note that the screen indicates there are four "Steps" to coding. Step 1 lists the program's goal, as stated by the program's director. Note that only the first 50 words are displayed. Step 2 lists program director's rating of the importance of fifteen learning outcomes assessed by the NSLLP. Step 3 lists the programs name. The final "step" is reserved for the type code *you* assign the program.

(6) To code:

a. Begin by reading the description. If on the basis of the description you believe you can code the program, check the box labeled "used goal" and skip to "d"
b. Review the goal ratings to see if you can identify the program's theme. If on the basis of the goal ratings you can code the program, check "used goal" and skip to "d"
c. Read the program name. If on the basis of the program name you can code it, check "used name" and skip to "d".
d. Now, you must indicate whether the program fits an EXISTING code, or you wish to suggest your own code.
    i. To assign an existing code, click on the downward pointing triangle at the end of "existing code" and select the appropriate code.

ii. To suggest a new code, choose "Not Listed" **and**, in the box next to "Create New Code" type in your suggestion.

(7) Your results are automatically saved. To move through the database, use the arrows on the right side of the screen.

(8) When done, or if you need a break, click on the small "x" in the upper right hand side of the window labeled "Code Programs." **If you are stopping mid-way and need to resume, note the record number you are leaving off with in the lower left hand corner. You can return to that record when you re-open the database by typing that number directly in the record number box.**

Other tasks from the main menu.
You may view a list of all the codes you have generated by choosing the option marked "See a list of codes you've created" (this option is also available on the coding screen). This may be helpful to avoid duplication.

You may view a list of all programs, sorted by code, by choosing the option marked "See a summary of your coding." If your screen width is too narrow, you may get an error message. Programs with NO code are listed first and will have no main category on the far left. Programs you have coded with your own code will be under "Not Listed," split out by "Suggested Category, if applicable."

To edit.
You may always go back in the database and edit your codes. You can return to programs by record number, or by words in their name. To search by name, position your cursor in the program name field and hit CTRL-F.

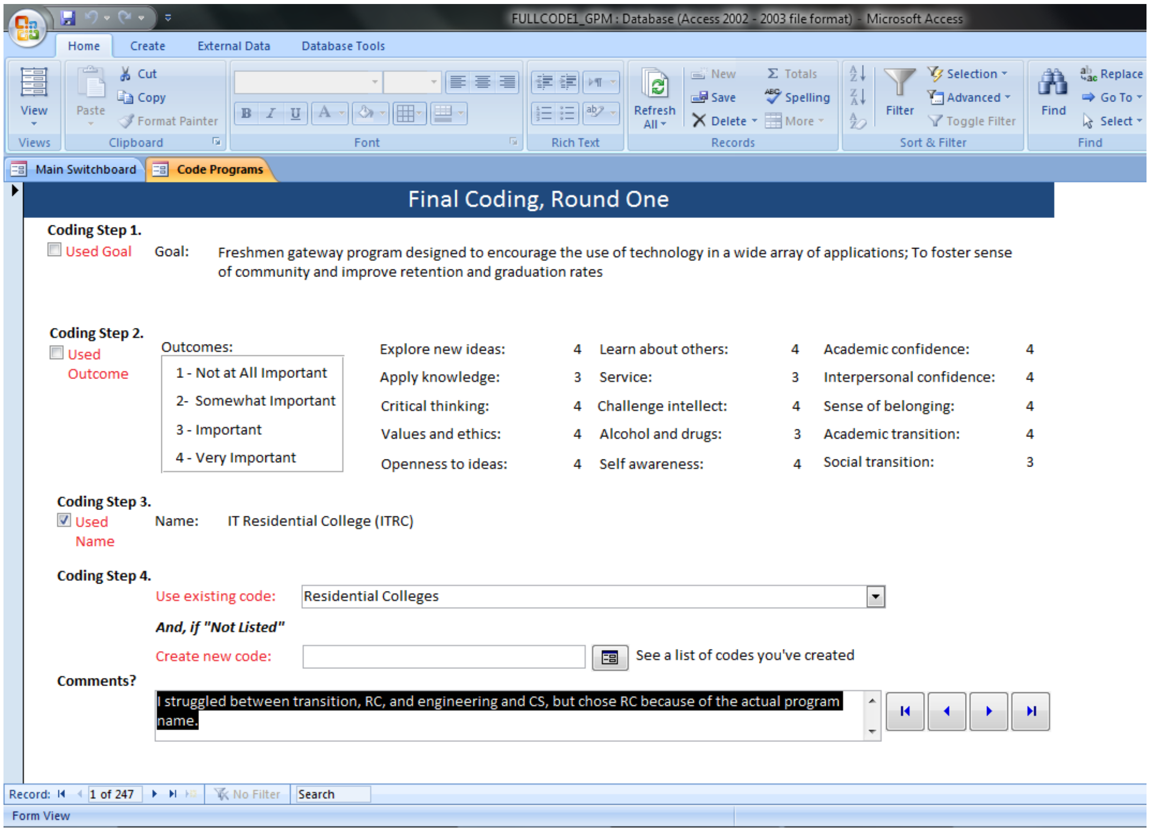
If you have a crisis.
Call me! I can always be reached at 410-241-0688.

Your deadline.
Preliminary testing suggests that coding takes less than one minute per record. Preliminary coding should take less than one hour, and full coding should take less than four hours.

Preliminary coding deadline	May 27, 2008
Final coding deadline	TBD

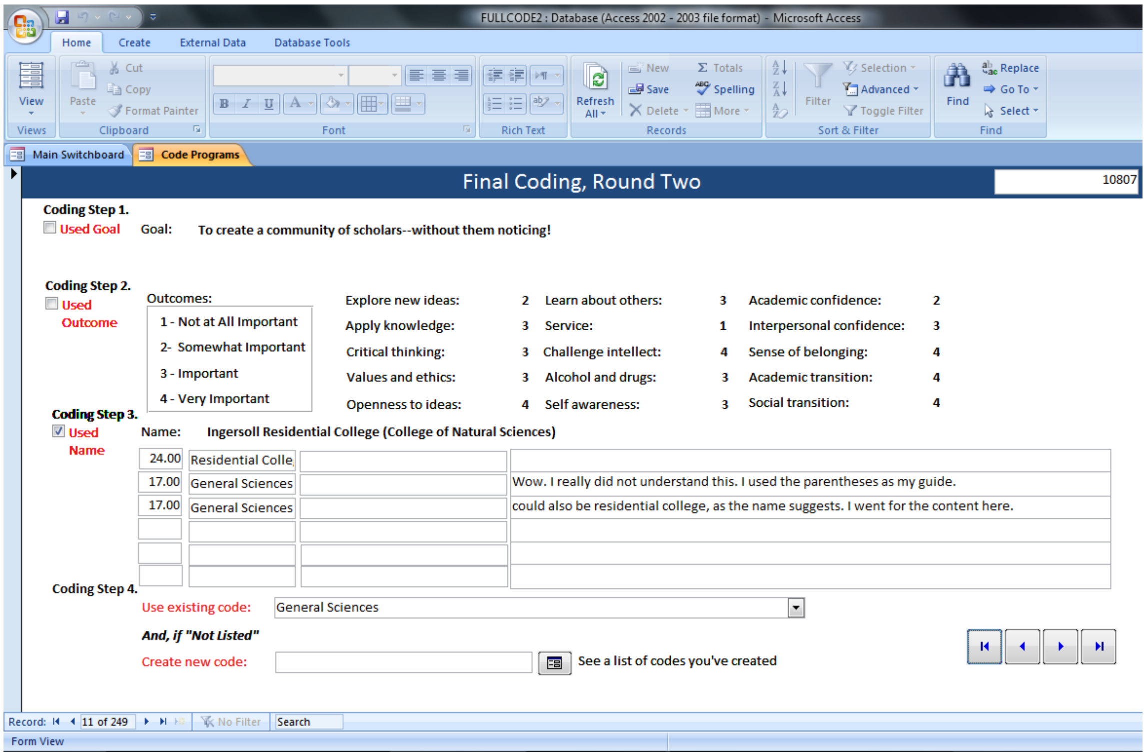
Appendix E

Exemplar Screen from Microsoft Access Coding Database (Raters' Version)



Appendix F

Exemplar Screen from Microsoft Access Coding Database (Final Version)



Appendix G

Newly Created Thematic Typology of L/L Programs

1. Civic and Social Leadership
    1.1. Civic engagement
    1.2. Environmental sustainability
    1.3. Leadership
    1.4. Service learning and social justice
2. Disciplinary
    2.1. Agriculture or veterinary medicine
    2.2. Business
    2.3. Communications or journalism
    2.4. Education
    2.5. Engineering and computer science
    2.6. General science
    2.7. Health sciences
    2.8. Humanities
    2.9. Interdisciplinary
    2.10. Law or criminal justice
    2.11. Mathematics
    2.12. Social sciences
3. Fine and Creative Arts
    3.1. Culinary arts
    3.2. Fine and creative arts
4. General Academic
5. Graduate Student
6. Honors Programs
7. Cultural
    7.1. International/global
    7.2. Multicultural/domestic diversity
8. Leisure
    8.1. General leisure
    8.2. Local community exploration
    8.3. Outdoor recreation
9. Political Interest
10. Residential College
11. Research
12. ROTC
13. Transition
    13.1. Career or major exploration
    13.2. First-year students
    13.3. Transfer students
    13.4. Transition programs for diverse populations
14. Umbrella
15. Upper Division
16. Wellness
    16.1. Spirituality or faith based
    16.2. General wellness or healthy living
17. Women's Programs
    17.1. Women's leadership
    17.2. Women-only science, technology, engineering, and math
18. Unknown

Appendix H

Rubin's Rules, as Described by Schafer & Olsen (1998)

As described by Schafer and Olsen (1998), let $\hat{Q}$ represent the estimate of a single quantity based upon one multiply imputed dataset, $\hat{U}$ represent that estimate's variance, and *m* represent the number of multiple imputations conducted. The aggregated point estimate, $\bar{Q}$, is a simple average described as:

$$\bar{Q} = \frac{1}{m}\sum_{i=1}^{m}\hat{Q}_i$$

That estimate has within-imputation variance, $\bar{U}$, described as a simple average of variances:

$$\bar{U} = \frac{1}{m}\sum_{i=1}^{m}\hat{U}_i$$

As well as between-imputation variance, *B*, described as the variance of the estimates themselves:

$$B = \frac{1}{m-1}\sum_{i=1}^{m}(\hat{Q}_i - \bar{Q})^2$$

And is adjusted to account for "simulation error," (p. 557), forming the quantity *T*:

$$T = \bar{U} + \left(1 + \frac{1}{m}\right)B$$

The square root of *T* is the estimate's standard error. That standard error can then be used to create a 95% confidence interval around the point estimate, $\bar{Q}$. That confidence interval is defined as:

$$\bar{Q} \pm t_{df}\sqrt{T}$$

Where the *t*-value is based upon with the Student's *t*-distribution with degrees of freedom:

$$df = (m-1)\left(1 + \frac{m\bar{U}}{(m+1)B}\right)^2$$

All point estimates, standard errors, and confidence intervals were automatically computed using Harding & Tussell's (2009) CAT package for R, using information gathered from EQS outputs generated from analyzing each multiply imputed dataset.

Appendix I

Nested Model Comparisons Using the Satorra-Bentler χ^2,

as Described by Satorra & Bentler (2001)

As described by Satorra and Bentler (2001), let $\mathcal{M}_0$ represent the initial model and represent the model nested within it. Each model has its own Satorra-Bentler scaling correction factor which defined as the ratio of its unscaled chi-square value and its Satorra-Bentler chi-square:

$$c_o = \frac{\mathcal{M}_{0_{\chi^2}}}{\mathcal{M}_{0_{S-B\chi^2}}} \quad and \quad c_1 = \frac{\mathcal{M}_{1_{\chi^2}}}{\mathcal{M}_{1_{S-B\chi^2}}}$$

Those correction factors are then combined to create a difference test scaling correction (Satorra & Bentler, 1999) based upon the number of degrees of freedom associated with each model, such that:

$$c_d = \frac{\left(\mathcal{M}_{0_{df}} \times c_o\right) \times \left(\mathcal{M}_{1_{df}} \times c_1\right)}{\mathcal{M}_{0_{df}} - \mathcal{M}_{1_{df}}}$$

The Satorra-Bentler chi-square difference test statistic is then computed:

$$T^*{}_d = \frac{\left(\mathcal{M}_{0_{\chi^2}} - \mathcal{M}_{1_{\chi^2}}\right)}{c_d}$$

In this dissertation, this difference test was frequently employed when testing cross-model constraints. Because only one constraint was released at a time in this testing, the degrees of freedom associated with each test was one, yielding a χ^2 critical value of 3.84 at $\alpha = 0.05$.

Appendix J

Calculating Standardized Effect Sizes

Standardized effect sizes provide one way for the reader to make a judgment about the practical difference between two group means (Thompson & Greene, 2006). Translating latent mean differences, κ, into standardized effect sizes, *ES*, is simple and relies upon three pieces of information: (a) the mean difference between two groups (κ), (b) the sample sizes of both groups, and (c) the relevant disturbance variance in both groups. Specifically:

$$ES = \frac{\kappa}{\sqrt{\frac{n_1\psi_1 + n_2\psi_2}{n_1 + n_2}}}$$

Translating differences between covariates into effect sizes is accomplished in a similar manner, substituting the mean's standard errors for disturbance variances for the measure of variability. The formula for Cohen's (1988) commonly used *d* statistic is as follows:

$$d = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{(n_1 - 1)(SE_1\sqrt{n_1})^2 + (n_2 - 1)(SE_2\sqrt{n_2})^2}{n_1 + n_2}}}$$

As can be noted by comparing the two equations, Thompson and Greene's (2006) formula is conceptually analogous to Cohen's (1988) *d* statistic, in which a mean difference—say from a t-test—is divided by a pooled standard deviation. As such, we might interpret both effect sizes similarly, considering those below .2 to be small, those approaching .5 to be medium, and those above .8 to be large.

References

Aarons, G. A. (2004). Multiple indicator multiple cause modeling: Application to children's mental health research. In C. Newman, C. Liberton, K. Kutash, & R. Friedman (Eds.), *Proceedings of the 16th Annual Research Conference: A System of Care for Children's Mental Health: Expanding the Research Base, March 2-5, 2003,* (pp. 455-458). Tampa, FL: The University of South Florida.

Aiken, L. S., Stein, J. A., & Bentler, P. M. (1994). Structural equation analyses of clinical subpopulation differences and comparative treatment outcomes: Characterizing the daily lives of drug addicts. *Journal of Consulting and Clinical Psychology, 62,* 488-499. doi:10.1037/0022-006X.62.3.488

Allport, G.W. (1954). *The nature of prejudice.* Reading, MA: Addison-Wesley.

American College Personnel Association & National Association of Student Personnel Administrators. (2004). *Learning reconsidered: A campus-wide focus on the student experience*. Washington, DC: Authors. Retrieved from http://www.myacpa.org/pub/documents/learningreconsidered.pdf

Ansalone, G. (2001). Schooling, tracking, and inequality. *Journal of Children & Poverty, 7*(1), 33-47. doi: 10.1080/10796120120038028

Arms, J. H., Cabrera, A. F., & Brower, A. (2008). Moving in to students' spaces: The impact of location of academic advising on student engagement among undecided students. *NACADA Journal, 28,* 8-18.

Arum, R., & Roksa, J. (2011). *Academically adrift: Limited learning on college campuses*. Chicago, IL: University of Chicago Press.

Association of American Colleges and Universities. (2002). *Greater expectations: A new vision for learning as a nation goes to college*. Washington, DC: Author.

Astin, A. W. (1991). *Assessment for excellence: The philosophy and practice of assessment and evaluation in higher education*. New York, NY: American Council on Education/Macmillan.

Astin, A. W. (1993). *What matters in college? Four critical years revisited*. San Francisco, CA: Jossey-Bass.

Astin, A. W., & Oseguera, L. (2004). The declining "equity" of American higher education. *The Review of Higher Education, 27*, 321-341. doi: 10.1353/rhe.2004.0001

Bagozzi, R. P., & Yi, Y. (1989). On the use of structural equation models in experimental designs. *Journal of Marketing Research, 26*(3), 271-284. doi: 10.2307/3172900

Belia, S., Fidler, F., Williams, J., & Cumming, G. (2005). Researchers misunderstand confidence intervals and standard error bars. *Psychological Methods, 10*, 389-396. doi: 10.1037/1082-989X.10.4.389

Bensimon, E. M., Hao, L., & Bustillos, L. T. (2003, October). *Measuring the state of equity in public higher education*. Paper presented at the Harvard Civil Rights and UC Conference on Expanding Opportunity in Higher Education: California and the Nation. Retrieved from http://www.wisconsin.edu/edi/equity/pdf/measuringthestateofequity.pdf

Bettinger, E. P., Evans, B. J., & Pope, D. G. (2011). *Improving college performance and retention the easy way: Unpacking the ACT exam*. (NBER Working Paper No. 17119). Retrieved from http://papers.nber.org/papers/w17119

Bok, D. (2006). *Our underachieving colleges: A candid look at how much students learn and why they should be learning more*. Princeton, NJ: Princeton University Press.

Bollen, K. A. (1989). *Structural equation models with latent variables.* New York, NY: Wiley.

Bonisteel, S. (2006). From treehouses to RV parks, students embrace dorm alternatives [Electronic Version]. *Fox News.com*. Retrieved January 9, 2007, from http://www.foxnews.com/story/0,2933,208180,00.html

Borg, W. R., & Gall, M. D. (1989). *Educational research: An introduction* (5th ed.). New York, NY: Longman.

Brower, A. M. (1997). *End-of-year evaluation on the Bradley Learning Community*. Madison: University of Wisconsin-Madison.

Byrne, B. M. (2006). *Structural equation modeling with EQS: Basic concepts, applications, and programming*. Mahwah, NJ: Lawrence Erlbaum Associates.

Chickering, A. W. (1969). *Education and identity*. San Francisco, CA: Jossey-Bass.

Chickering, A. W., & Gamson, Z. F. (1991). *Applying the seven principles for good practice in undergraduate education*. New Directions for Teaching and Learning, No. 47. San Francisco, CA: Jossey-Bass.

Chickering, A. W., & Reisser, L. (1993). *Education and identity* (2nd ed.). San Francisco, CA: Jossey-Bass.

Clarke, J. H., Miser, K. M., & Roberts, A. O. (1988). Freshman residential programs: Effects of living-learning structure, faculty involvement, and thematic focus. *Journal of College & University Student Housing, 18*(2), 7-13.

Cohen, A. M. (1998). *The shaping of American higher education: Emergence and growth of the contemporary system*. San Francisco: Jossey-Bass.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

Cook, B. J., & Cordova, D. I. (2006). *Minorities in higher education: Twenty-second annual status report.* Washington, DC: American Council on Education.

Conger, D., & Long, M. C. (2010). Why are men falling behind? Gender gaps in college performance and persistence. *The ANNALS of the American Academy of Political and Social Science, 627,* 184-214. doi: 10.1177/0002716209348751

Crawford, S. D., Couper, M. P., & Lamia, M. J. (2001). Web surveys: Perceptions of burden. *Social Science Computer Review, 19*, 146-162. doi: 10.1177/089443930101900202

Darrity, Jr., W., Castellino, D., Tyson, K., Cobb, C., & McMillen, B. (2001). *Increasing opportunity to learn via access to rigorous courses and programs: One strategy for closing the achievement gap for at-risk and ethnic minority students*. Raleigh, NC: North Carolina State Board of Education, Department of Public Instruction. Retrieved from http://www.ncpublicschools.org/docs/academicservices/gifted/resources/reports/increasingopportunities.pdf

DeVellis, R. F. (2003). *Scale development: Theory and applications* (2nd ed.). Thousand Oaks, CA: Sage.

Edwards, K. E., & McKelfresh, D. A. (2002). The impact of a living learning center on students' academic success and persistence. *Journal of College Student Development, 43*, 395-402.

Ehrenberg, R. G. (2003). Reaching for the brass ring: The *U.S. News & World Report* rankings and competition. *Review of Higher Education, 26*, 145-162. doi: 10.1353/rhe.2002.0032

Enders, C. K. (2001). The performance of the full information maximum likelihood estimator in multiple regression models with missing data. *Education and Psychological Measurement*, *61*, 713-740. doi: 10.1177/0013164401615001

Enders, C. K. (2006). Analyzing structural equation models with missing data. In G. R. Hancock & R. O. Mueller (Eds.), *Structural equation modeling: A second course* (pp. 315-344). Charlotte, NC: Information Age Publishing.

Fan, W., & Hancock, G. R. (2006). Impact of post hoc measurement model overspecification on structural parameter integrity. *Journal of Educational and Psychological Measurement, 66*, 748-764. doi: 10.1177/0013164405285908

Finney, S. J., & DiStefano, C. (2006). Non-normal and categorical data in structural equation modeling. . In G. R. Hancock & R. O. Mueller (Eds.), *Structural equation modeling: A second course* (pp. 269-314). Charlotte, NC: Information Age Publishing..

Foderaro, L. W. (2005, April 24). Under one roof: Learning communities. *The New York Times, Education Life Supplement,* p. 26.

Gabelnick, F., MacGregor, J., Matthews, R. S., & Smith, R. S. (1990). *Learning communities: Creating connections among students, faculty, and disciplines*. New Directions for Teaching and Learning No. 41. San Francisco: Jossey-Bass.

Graham, J. W. & Schafer, J. L. (1999). On the performance of multiple imputation for multivariate data with small sample size. In R. H. Hoyle (Ed.), *Statistical Strategies for Small Sample Research* (pp. 1-29). Thousand Oaks, CA: Sage.

Hallinan, M. T. (1994). Tracking: From theory to practice. *Sociology of Education, 67*, 79-84. doi: 10.2307/2112697

Hancock, G. R. (1997). Structural equation modeling methods of hypothesis testing of latent variable means. *Measurement & Evaluation in Counseling & Development, 30,* 91-106.

Hancock, G. R. (2001). Effect size, power, and sample size determination for structured means modeling and MIMIC approaches to between-groups hypothesis testing of means on a single latent construct. *Psychometrika, 66*, 373-388. doi: 10.1007/BF02294440

Hancock, G. R. (2003). Fortune cookies, measurement error, and experimental design. *Journal of Modern and Applied Statistical Methods, 2,* 293-305.

Hancock, G. R., & Mueller, R. O. (2009). *Advanced topics in structural equation modeling.* The Department of Measurement, Statistics, and Evaluation, University of Maryland, College Park: Authors.

Hao, L. (2002). *The Equity Index: A method to measure equity in educational outcomes for minority students*. Poster session presented at the annual meeting of the Association for the Study of Higher Education in Sacramento, CA.

Harding, T., & Tussell, F. (2009). *Cat: Analysis of categorical-variable datasets with missing values* (R Package Version 0.0-6) [Computer software]. Vienna: Department of Statistics and Mathematics, Vienna University of Economics and Business Administration. Retrieved from http://cran.r-project.org/web/packages/cat

Hayes, A. F., & Krippendorff, K. (2007). Answering the call for a standard reliability measure for coding data. *Communication Methods and Measures, 1*, 77-89. doi: 10.1080/19312450709336664

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling,* 6, 1-5. doi: 10.1080/10705519909540118

Inkelas, K. K., Brower, A. M., & Associates. (2004). *The National Study of Living Learning Programs: 2004 report of findings*. Retrieved from http://www.livelearnstudy.net/images/ NSLLP_2004_Final_Report.pdf

Inkelas, K. K., Brower, A. M., & Associates. (2008*). The National Study of Living Learning Programs: 2007 report of findings*. Retrieved from http://www.livelearnstudy.net/images/ 2007_NSLLP_Final_Report.pdf

Inkelas, K. K., Johnson, D., Lee, Z., Daver, Z., Longerbeam, S. D., Vogt, K., & Brown Leonard, J. (2006). The role of living-learning programs in students' perceptions of intellectual growth at three large universities. *NASPA Journal, 43*, 115-143.

Inkelas, K. K., Longerbeam, S., Brown Leonard, J., & Soldner, M. (2005, November). *Understanding differences in student outcomes by types of living-learning programs: The development of two typologies.* Paper presented at the Annual Meeting of the Association for the Study of Higher Education, Philadelphia, PA.

Inkelas, K. K., Soldner, M., Longerbeam, S. D., & Brown Leonard, J. (2008). Differences in student outcomes by types of living-learning programs: The development of an empirical typology. *Research in Higher Education, 49*, 495-512. doi: 10.1007/s11162-008-9087-6

Inkelas, K. K., Vogt, K. E., Daver, Z. E., & Brown Leonard, J. (2007). Living-learning programs and first-generation college students' academic and social transition to college. *Research in Higher Education, 48*, 403-423. doi: 10.1007/s11162-006-9031-6

Inkelas, K. K., Vogt, K. E., Longerbeam, S. D., Owen, J., & Johnson, D. (2006). Measuring outcomes of living-learning programs: Examining college environments and student learning and development. *Journal of General Education, 55*(1), 40-76. doi: 10.1353/jge.2006.0017

Inkelas, K. K., & Weisman, J. L. (2003). Different by design: An examination of student outcomes among participants in three types of living-learning programs. *Journal of College Student Development, 44*, 335-368. doi: 10.1353/csd.2003.0027

Jamshidian, M., & Bentler, P. (1999). ML estimation of mean and covariance structures with missing data using complete data routines. *Journal of Educational and Behavioral Statistics, 24*(1), 21-41. doi: 10.2307/1165260

Johnson, D. R. (2007). *Sense of belonging among women of color in science, technology, engineering, and math majors: Investigating the contributions of campus racial climate perceptions and other college environments.* (Doctoral dissertation). Retrieved from http://hdl.handle.net/1903/7723

Jones, P. R., Levine Laufgraben, J., & Morris, N. (2006). Developing an empirically based typology of attitudes of entering students toward participation in learning communities. *Assessment & Evaluation in Higher Education, 31*, 249-265. doi: 10.1080/02602930500352766

Kamin, M. A. (2009). *Academic self-efficacy for sophomore students in living-learning programs* (Masters thesis). Retrieved from http://hdl.handle.net/1903/9480.

Kano, Y. (2000). Structural equation modeling for experimental data. In S. DuToit, R. Cudeck, & D. Sörbom (Eds.,), *Structural equation modeling: Present and future* (pp. 381-402). Lincolnwood, IL: Scientific Software International.

Keeling, R. P. (Ed.). (2006). *Learning reconsidered 2: Implementing a campus-wide focus on the student experience.* Washington, DC: American College Personnel Association, Association of College and University Housing Officers – International, Association of College Unions – International, National Academic Advising Association, National Association for Campus Activities, National Association of Student Personnel Administrators, National Intramural-

Recreational Sports Association. Retrieved from http://www.myacpa.org/pub/documents/LearningReconsidered2.pdf

Kennedy, P. (2003). *A guide to econometrics*. Cambridge, MA: The MIT Press.

Kim, Y. K., Chang, M. J., & Park, J. J. (2009) Engaging with faculty: Examining rates, predictors, and educational effects for Asian American undergraduates. *Journal of Diversity in Higher Education, 2,* 208-218. doi: 10.1037/a0017890

King. P. M., & Baxter Magolda, B. M. (2005). *Toward a developmental model of intercultural maturity.* In M. E. Wilson & L. E. Wolf-Wendel (Eds.), *ASHE Reader on College Student Development Theory* (pp.101-113). Boston, MA: Pearson Custom Publishing.

Kinzie, J., Palmer, M., Hayek, J., Hossler, D., Jacob, S. A., & Cummings, H. (2004). *Fifty years of college choice: Social, political, and institutional influences on the decision-making process.* Indianapolis, IN: Lumina Foundation for Education.

Kline, R. B. (2006). *Reverse arrow dynamics.* In G. R. Hancock & R. O. Mueller (Eds.), *Structural equation modeling: A second course* (pp. 43-68). Charlotte, NC: Information Age Publishing.

Knapp, L. G., Kelly-Reid, J. K., & Ginder, S. A. (2009). *Enrollment in Postsecondary Institutions, Fall 2007; Graduation Rates, 2001 and 2004 Cohorts; and Financial Statistics, Fiscal Year 2007* (NCES 2009-155). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education: Washington, DC.

Kohl, J. L. (2009). *The association of critical thinking and participation in living and learning programs: Residential honors compared to civic/social leadership programs and non-participation in living and learning programs* (Doctoral dissertation). Retrieved from http://hdl.handle.net/1903/9871

Kuh, G. D. (2008). *High impact educational practices: What they are, who has access to them, and why they matter*. Washington, DC: Association of American Colleges and Universities.

Kuh, G. D., Kinzie, J., Cruce, T., Shoup, R., & Gonyea, R. M. (2007). *Connecting the dots: Multi-faceted analyses of the relationships between student engagement results from the NSSE, and the institutional practices and conditions that foster student success*. Bloomington, IN: Center for Postsecondary Research, Indiana University Bloomington.

Laufgraben, J. L., Shapiro, N. S., & Associates. (2004). *Sustaining & improving learning communities*. San Francisco, CA: Jossey Bass.

Leinwall, F. (2006). *A residential learning community and its affect on the psychosocial development of college students* (Doctoral dissertation). Retrieved from ProQuest Dissertations & Theses: Full Text. (AAT 3235107)

Little, R. J. A. (1988). A test of missing completely at random for multivariate data with missing values. *Journal of the American Statistical Association, 83*, 1198-1202. doi: 10.2307/2290157

Lenning, O. T., & Ebbers, L. H. (1999). *The powerful potential of learning communities: Improving education for the future.* (ASHE-ERIC Higher Education Report

Volume 26, No. 6). Washington, DC: The George Washington University, Graduate School of Education and Human Development.

Leslie, L. L., & Brinkman, P. T. (1988). *The economic value of higher education*. New York, NY: Macmillan.

Lombard, M., Snyder-Duch, J., & Bracken, C. C. (2002). Content analysis in mass communciation: Assessment and reporting of intercoder reliability. *Human Communication Research, 28*, 587-604. doi: 10.1111/j.1468-2958.2002.tb00826.x

Longerbeam, S. D. (2005). *Exploring the relationships among living-learning programs, peer interactions, critical thinking, and civic engagement on college student openness to diversity*. (Doctoral dissertation). Retrieved from http://hdl.handle.net/1903/2635

Loyola College of Maryland. (2008). *Strategic planning for 2008-2012*. Retrieved from http://www.loyola.edu/about/stratplan/

Loveless, T. (1999). *The tracking wars: State reform meets school policy*. Washington, DC: The Brookings Institution.

McEwen, M. K. (2003). The nature and uses of theory. In S. R. Komives & D. B. Woodard Jr. (Eds.), *Student services: A handbook for the profession* (4th ed., pp. 153-178). San Francisco, CA: Jossey-Bass.

Meredith, M. (2004). Why do universities compete in the ratings game? An empirical analysis of the effects of the U.S. News and World Report college rankings. *Research in Higher Education, 45*, 443-461. doi: 10.1023/B:RIHE.0000032324.46716.f4

Midden, R. (2011). *The residential learning communities international clearinghouse*. Retrieved from http://pcc.bgsu.edu/rlcch/

Monks, J., & Ehrenberg, R. G. (1999, November/December). U.S. News and World Report rankings: Why they do matter. *Change,* 43-51. doi: 10.1080/00091389909604232

Morgan, S. L., & Winship C. (2007). *Counterfactuals and causal inference: Methods and principles for social research*. New York, NY: Cambridge University Press.

Museus, S., Lutovsky, B., & Colbeck, C. (2007). Access and equity in dual enrollment programs: Implications for policy formation. *Higher Education in Review, 4,* 1-19.

National Leadership Council. (2007). *College learning for the new global century*. Washington, DC: Association of American Colleges and Universities.

National Survey of Student Engagement. (2007). *Experiences that matter: Enhancing student learning and success (Annual Report 2007).* Bloomington: Center for Postsecondary Research, Indiana University Bloomington.

Neuendorf, K. A. (2002). *The content analysis guidebook*. Thousand Oaks, CA: SAGE.

Oakes, J. (2005). *Keeping track*. New Haven, CT: Yale University Press.

Pace, R. (1984). *Measuring the quality of college student experiences: An account of the development and use of the College Student Experiences Questionnaire*. Los Angeles: Higher Education Research Institute; Graduate School of Education; University of California, Los Angeles.

Pascarella, E. T. (2001). Using self-reported gains to estimate college impact: A cautionary tale. *Journal of College Student Development, 42*, 488-492.

Pascarella, E. T., Cruce, T., Umbach, P. D., Wolniak, G. C., Kuh, G. D., Carini, R. M., et al. (2006). Institutional selectivity and good practices in undergraduate education: How strong is the link? *Journal of Higher Education, 77*, 251-285. doi: 10.1353/jhe.2006.0016

Pascarella, E. T., & Terenzini, P. T. (1980). Student-faculty and student-peer relationships as mediators of the structural effects of undergraduate residence arrangements. *Journal of Educational Research, 73*, 34-353.

Pascarella, E. T., & Terenzini, P. T. (1991). *How college affects students: Findings and insights from twenty years of research*. San Francisco, CA: Jossey-Bass.

Pascarella, E. T., & Terenzini, P. T. (2005). *How college affects students: A third decade of research* (Vol. 2). San Francisco, CA: Jossey-Bass.

Pasque, P. A., & Murphy, R. (2005). The intersections of living-learning programs and social identity as factors of academic achievement and intellectual engagement. *Journal of College Student Development, 46*, 429-441. doi: 10.1353/csd.2005.0041

Pike, G. R. (1999). The effects of residential learning communities and traditional residence hall arrangements on educational gains during the first year of college. *Journal of College Student Development, 40*, 269-284.

Pike, G. R., Hansen, M. J., & Lin, C-H. (2010). Using instrumental variables to account for selection effects in research on first-year programs. *Research in Higher Education, 52*, 194-214. doi: 10.1007/s11162-010-9188-x

Pike, G. R., Schroeder, C. C., & Berry, T. R. (1997). Enhancing the educational impact of residence halls: The relationship between residential learning communities and first-year college experiences and persistence. *Journal of College Student Development, 38*, 609-621.

Powell, W. W., & DiMaggio, P. J. (1991). *The new institutionalism in organizational analysis*. Chicago, IL: University of Chicago Press.

Purdie II, J. R. (2007). *Examining the academic performance and retention of first-year students in living-learning communities, freshman interest groups and first year experience courses attainment* (Doctoral dissertation). Available from ProQuest Dissertations & Theses: Full Text database. (UMI No. 3322736).

Rowan-Kenyon, H., Soldner, M., & Inkelas, K. K. (2008). The contributions of living-learning programs on developing sense of civic engagement in undergraduate students. *NASPA Journal, 44*, 750-778.

Satorra, A. & Bentler, P. M. (2001). A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika*, *66*, 507-514. doi: 10.1007/BF02296192.

Schafer, J. (1987). *Analysis of incomplete multivariate data.* London, England: Chapman & Hall.

Schafer, J. L., & Olson, M. K. (1998). Multiple imputation for multivariate missing-data problems: A data analyst's perspective. *Multivariate Behavioral Research, 33*, 545-571. doi: 10.1207/s15327906mbr3304_5

Schneider, D., & Dorans, N. (1999). *Concordance between SAT I and ACT scores for individual students.* New York, NY: The College Board.

Schoem, D. (2004). Sustaining living-learning programs. In J. L. Laufgraben, N. S. Shapiro & Associates (Eds.), *Sustaining & improving learning communities* (pp. 130-156). San Francisco: Jossey-Bass.

Scholnick, E. K. (1996). *A two-year longitudinal study of science and math students in the College Park Scholars program* (Unpublished manuscript). University of Maryland, College Park.

Schroeder, C. C. (1994). Developing learning communities. In C. C. Schroeder, P. Mable & Associates (Eds.), *Realizing the educational potential of residence halls* (pp. 165-189). San Francisco, CA: Jossey-Bass.

Scott, W. R. (1987). The adolescence of institutional theory. *Administrative Science Quarterly, 32*, 493-511. doi: 10.2307/2392880

Scott, W. R. (2003). *Organizations: Rational, natural, and open systems* (5th ed.). Upper Saddle River, NJ: Prentice Hall.

Shapiro, N. S., & Levine, J. H. (1999). *Creating learning communities: A practical guide to winning support, organizing for change, and implementing programs*. San Francisco, CA: Jossey-Bass.

Shields, P. M., & Tajalli, H. (2006). Intermediate theory: The missing link in student scholarship. *Journal of Public Affairs Education, 12*, 313-334.

Sim, J., & Wright, C. C. (2005). The kappa statistic in reliability studies: Use, interpretation, and sample size requirements. *Physical Therapy, 85*, 257-268.

Singh, A., & Nocerino, J. (1994). *Robust estimation of mean and variance using environmental data sets with below detection limit observations.* Retrieved from http://www.epa.gov/esd/cmb/research/papers/jn103.pdf

Smith, T. B. (1994). Integrating living and learning in residential colleges. In C. C. Schroeder, P. Mable & Associates (Eds.), *Realizing the educational potential of residence halls* (pp. 241-265). San Francisco, CA: Jossey-Bass.

Soldner, M., McCarron, G. P., & Inkelas, K. K. (2007, November). *Honors living-learning programs for first-year students: Educational benefits for whom?* Paper presented at the Annual Meeting of the Association for the Study of Higher Education, Louisville, KY

Soldner, M. & Szelenyi, K. (2008). A national portrait of today's living-learning programs. *Journal of College and University Student Housing*, *35,* 14-31.

Sörbom, D. (1978) An alternative to the methodology for analysis of covariance. *Psychometrika, 43*, 381-396. doi: 10.1007/BF02293647.

SPSS Inc. (2007). *SPSS missing value analysis™ 16.0.* Chicago, IL: Author.

Stassen, M. L. A. (2003). Student outcomes: The impact of varying living-learning community models. *Research in Higher Education, 44*, 581-613. doi: 10.1023/A:1025495309569

Terenzini, P. T., Cabrera, A. F., & Bernal, E. M. (2001). *Swimming against the tide: The poor in American higher education* (College Board Research Report No. 2001-1). New York, NY: College Entrance Examination Board.

Terenzini, P. T., & Pascarella, E. T. (1994, January/February). Living with myths: Undergraduate education in America. *Change, 26*, 28-32. doi: 10.1080/00091383.1994.9938488

Terenzini, P. T., Pascarella, E. T., & Blimling, G. S. (1996). Students' out-of-class experiences and their influence on learning and cognitive development: A literature review. *Journal of College Student Development, 37,* 149-162.

The College Board. (2006). How the SAT can help your institution. Retrieved from www.collegeboard.com/prod_downloads/highered/ra/sat-admission-guide.pdf

The Secretary of Education's Commission on the Future of Higher Education. (2006). *A test of leadership: Charting the future of U.S. higher education*. Washington, DC: Department of Education.

Thompson, M. S., & Green, S. B. (2006). Evaluating between-group differences in latent variable means. In G. R. Hancock & R. O. Mueller (Eds.), *Structural equation modeling: A second course* (pp. 119-170). Charlotte, NC: Information Age Publishing.

Thomson Peterson's. (2006). *Peterson's four-year colleges 2007* (37th ed.). Lawrenceville, NJ: Author.

Tinto, V. (1993). *Leaving college: Rethinking the causes and cures of student attrition* (2nd ed.). Chicago, IL: University of Chicago Press.

Tinto, V. (2000, March/April). What have we learned about the impact of learning communities on students? *Assessment Update, 12*, 1-2, 12.

University of California-Irvine (2008). *A focus on excellence: A strategy for academic development at the University of California, Irvine 2005-2015*. Retrieved from http://www.strategicplan.uci.edu/unitplans/?p=28

Walsh, W. B. (1978). Person/environment interaction. In G. D. Kuh, J. P. Bean, D. Hossler & F. K. Stage (Eds.), *ASHE reader on college students* (1989 ed., pp. 105-113). Needham Heights, MA: Ginn Press.

Widaman, K. F. (2006). Missing data: What to do with or without them. *Monographs of the Society for Research in Child Development, 71*(3), 42-64. doi: 10.1111/j.1540-5834.2006.00404.x

Willse, J. T., & Goodman, J. T. (2008). Comparison of multiple-indicators, multiple-causes- and item response theory-based analyses of subgroup differences. *Educational and Psychological Measurement, 68,* 587-602. doi: 10.1177/0013164407312601

Wolfe, R., & Hanley, J. (2002). If we're so different, why do we keep overlapping? When 1 plus 1 doesn't make 2. *Canadian Medical Association Journal, 166*(1), 65-66.

Yoo, J. E. (2009). The effect of auxiliary variables and multiple imputation on parameter estimation in confirmatory factor analysis. *Educational and Psychological Measurement,69,* 929-947. doi: 10.1177/0013164409332225.