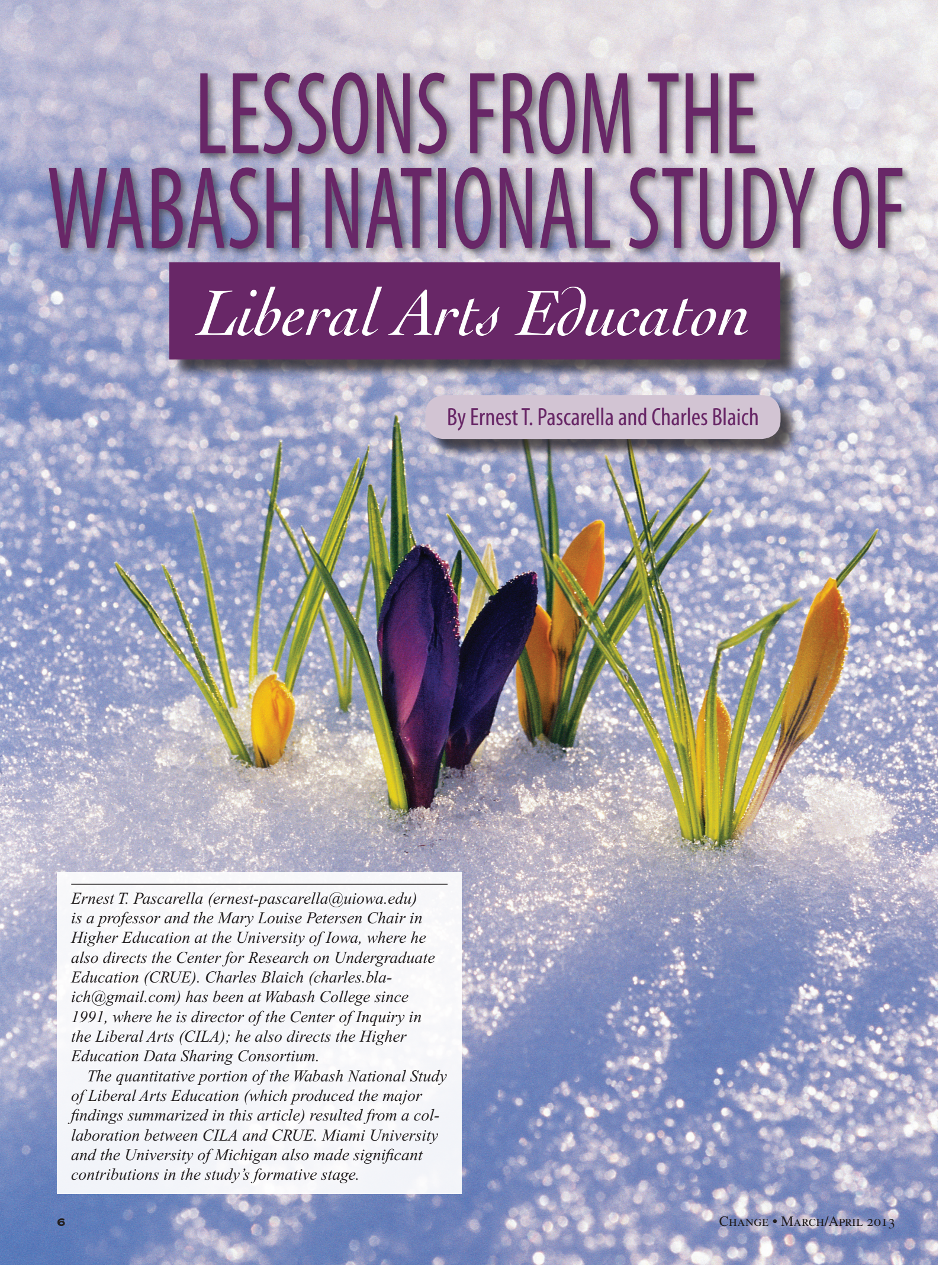
LESSONS FROM THE WABASH NATIONAL STUDY OF *Liberal Arts Educaton*

By Ernest T. Pascarella and Charles Blaich



Ernest T. Pascarella (ernest-pascarella@uiowa.edu) is a professor and the Mary Louise Petersen Chair in Higher Education at the University of Iowa, where he also directs the Center for Research on Undergraduate Education (CRUE). Charles Blaich (charles.blaich@gmail.com) has been at Wabash College since 1991, where he is director of the Center of Inquiry in the Liberal Arts (CILA); he also directs the Higher Education Data Sharing Consortium.

The quantitative portion of the Wabash National Study of Liberal Arts Education (which produced the major findings summarized in this article) resulted from a collaboration between CILA and CRUE. Miami University and the University of Michigan also made significant contributions in the study's formative stage.

Since its inception, American postsecondary education has had a core belief in the desirability of a kind of individual development that prepares one for a complete life, and it has placed great faith in a liberal arts education to provide a transformative undergraduate experience that would lead to such development. Liberal learning has typically been characterized by a broad range of compelling and desirable outcomes—among others, effective reasoning and problem solving, a capacity for lifelong learning, the ability to interact with people from different cultures, moral character, and leadership. Although it may be difficult to argue with the educational importance of such outcomes, it has not always been clear how they are influenced by college. A recent longitudinal study, however, has addressed this issue.

Funded by the Center of Inquiry in the Liberal Arts (CILA) at Wabash College, the Wabash National Study of Liberal Arts Education (WNS) is a multi-institution, multi-year, longitudinal study designed to identify the academic and non-academic collegiate experiences that foster liberal learning. In what follows, we describe how the study was done and highlight some of our major findings to date, with an emphasis on those related to cognitive outcomes.

The Study

Data Collection

The first WNS cohort, upon which most of the findings we report are based, consisted of 19 institutions (see Box 1). The individuals in the study were first-year, full-time undergraduates—either selected randomly from each institution's incoming class or, in the case of most small liberal arts colleges, composed of the entire incoming class.

The students were assessed three times: upon matriculation in fall 2006, at the end of the first year of college in 2007, and at the end of the fourth year in 2010. Each assessment lasted between 90 and 120 minutes and was administered by ACT. Students were paid a stipend of $50 for their participation in each assessment.

The 2006 pre-college assessment collected information on student demographic characteristics, family backgrounds, secondary school experiences, and the like. Students also completed a series of instruments that measured dimensions of cognitive and personal development theoretically associated with a liberal arts education (see Box 2).

The follow-up assessments in spring 2007 and spring 2010 gathered two types of data. First, we collected extensive information about students' academic and non-academic

Box 1: WNS Participating Institutions (2006 Cohort)

Institution	Sector of Country	State
Alma College	Midwest	Michigan
Bard College	Northeast	New York
Butler University	Midwest	Indiana
Coe College	Midwest	Iowa
Columbia College	Southeast	South Carolina
Connecticut College	Northeast	Connecticut
Gustavus-Adolphus College	Midwest	Minnesota
Hamilton College	Northeast	New York
Hampshire College	Northeast	Massachusetts
Hope College	Midwest	Michigan
Ivy Tech Community College	Midwest	Indiana
Kirkwood Community College	Midwest	Iowa
San Jose State University	Pacific Coast	California
University of Kentucky	Southeast	Kentucky
University of Michigan	Midwest	Michigan
University of North Carolina - Wilmington	Southeast	North Carolina
University of Notre Dame	Midwest	Indiana
Wabash College	Midwest	Indiana
Whittier College	Pacific Coast	California

Box 2: Liberal Arts Outcomes and Measurement Instruments

Liberal Arts Outcome	Measurement Instrument Used
Effective Reasoning and Problem Solving	Critical Thinking Test of the Collegiate Assessment of Academic Proficiency, (CAAP)
Inclination to Inquire and Life-Long Learning	Need for Cognition Scale and Positive Attitude toward Literacy Activities Scale
Intercultural Effectiveness	Miville-Guzman Universality/Diversity Scale and Openness to Diversity/Challenge Scale
Moral Character	Defining Issues Test – 2
Leadership	Socially Responsible Leadership Scales

The psychometric properties of each of these measures (i.e., reliability and validity) can be found at http://www.education.uiowa.edu/centers/crue/Publications.aspx#WNSLAE.

experiences during college. Second, students completed follow-up assessments of the elements of their cognitive and personal development first measured in the 2006 pre-college data collection (Box 2).

The end-of-first-year (2007) findings we report are based on a sample of about 3,100 four- and two-year college students, while the end-of-fourth-year findings are based on one of about 2,200 four-year college students. Using information provided by each institution, these samples were weighted up to the end-of-first- and end-of-fourth-year populations at each institution by sex, race, and ACT score (or its SAT or COMPASS equivalent).

Analytic Approach

Because our data were non-experimental, we employed various forms of regression and propensity-score analysis to estimate the causal relationships between college experiences and outcomes. We first hypothesized an association between some college experience and a liberal arts outcome (e.g., engagement in diversity experiences and growth in critical-thinking skills). Then we added potential confounding influences (see Box 3) to our model in an attempt to reduce the posited association to zero (i.e., statistical non-significance). If this occurred, we ruled out a causal relationship. However, if the association persisted (i.e., remained statistically significant) we concluded that there was a possible causal relationship.

> *Students' exposure to clear and organized instruction enhanced not only their general cognitive skills such as critical thinking but also their orientation toward inquiry and continuing intellectual development.*

Box 3: Confounding Influences

- ACT (or equivalent score)
- Pre-college measure of the liberal arts outcome
- Race
- Sex
- Parental education
- Secondary school involvement
- Pre-college academic motivation
- Institutional type—i.e., liberal arts college, research university, regional institution, and community college (in the first year of the study)
- Place of residence during college
- Work responsibilities during college
- Academic major or type of coursework taken

Nearly all analyses were based on samples weighted to estimate population characteristics, and we typically controlled for the clustering effect (i.e., the tendency for students at each institution to respond more like each other than they do like students at other institutions).

Findings

Broad-Based Impacts of Good Classroom Teaching

Based on extensive evidence at the classroom level indicating that effective instruction enhances course achievement, we constructed a 10-item scale (reliability: .89) based on two instructional characteristics: instructional organization (representative items include "Class time is used effectively" and "Presentation of material is well organized") and instructional clarity (representative items include "Teachers give clear explanations" and "Teachers make good use of examples and illustrations to explain difficult points"). Rather than asking students about instruction in a single course, we requested that they indicate the overall level of instructional clarity and organization that characterized the teaching they received at the institution they were attending.

We hypothesized that student exposure to clear and organized instruction over four years would foster gains in both effective reasoning/problem solving skills (measured by the CAAP critical-thinking test) and the inclination to inquire and to pursue life-long learning (measured by the Need for Cognition scale). And indeed, even after statistical controls were introduced for important confounding influences (Box 3), there was a statistically significant, positive association between students' perceptions of being exposed to clear and organized instruction in their coursework overall and four-year gains on both measures. Students' exposure to clear and organized instruction enhanced not only their general cognitive skills such as critical thinking but also their orientation toward inquiry and continuing intellectual development.

Based on research by John Braxton and his colleagues, we also hypothesized that exposure to such instruction would play a positive role in student persistence at the institution they were attending. So we expected to see a positive relationship between scores on the instructional clarity and organization scale during the first year of college and re-enrolling in the same institution for the second year.

This hypothesis was also supported. With controls in place for a variety of factors—pre-college ACT (or equivalent) score, gender, race, parental educational attainment, pre-college educational aspirations, type of institution attended, co-curricular involvement during college, and first-year college grades—students' perceptions of exposure to clear and organized instruction during the first year of college significantly improved the probability of their re-enrolling at the same institution for the second year, irrespective of the type of institution attended.

But the relationship was not direct: Instructional clarity and organization substantially enhanced student perceptions of *satisfaction* with college—which directly helped determine re-enrollment for the second year at the same institution. That such instruction exerted its influence indirectly does not diminish its importance; it still played a significant role in the causal chain that led to second year re-enrollment.

These findings clearly suggest that effective teaching practices count in ways that extend beyond achievement in individual courses to the fostering of general intellectual skills and orientations, as well as increased student retention. Furthermore, many if not all of the instructional skills constituting the clear and organized instruction scale are learnable by college faculty. Thus, institutional investments in programs designed to improve faculty's teaching skills may have institution-level returns beyond those realized in specific courses.

Box 4: Deep Learning Scales and Representative Items

Scale	Representative Items
Higher-Order Learning	"Applied theories or concepts to practical problems or in new situations"
	"Synthesized and organized ideas or experiences into new, more complex interpretations and relationships"
Integrative Learning	"Worked on a paper or project that required integrating ideas or information from various sources"
	"Put together ideas and concepts from different courses when completing assignments or during class discussions"
Reflective Learning	"Learned something that changed the way you understand an issue or concept"
	"Examined the strengths and weaknesses of your own views on a topic"

Reliabilities for the scales range from .72 to .82.

> Reflective learning had a significant positive link with four-year growth in critical-thinking skills, while all three deep-learning scales had significant unique and positive associations with four-year growth in the need for cognition.

The Developmental Impact of Deep-Learning Experiences

Guided by the emerging literature on deep approaches to learning, Thomas Nelson Laird of Indiana University developed three measures of "deep learning" from items embedded in the National Survey of Student Engagement (NSSE) filled out by all WNS participants in spring 2007 and spring 2010 (see Box 4): those in the categories of higher-order, integrative, and reflective learning.

We hypothesized that such deep-learning experiences, many of which can be purposefully integrated into coursework, would foster a growth in cognitive skills, an orientation toward cognitive activity, and the use of principled/post-conventional reasoning in judging moral dilemmas. This hypothesis was supported for the first year of college. Controlling for the list of confounding variables and for the other deep-learning scales, reflective learning was found to have a unique positive association with first-year growth in both critical-thinking skills and the need for cognition, while integrative learning had a unique positive link with growth in the need for cognition and principled/post-conventional moral reasoning.

Our findings for the deep-learning scales over four years of college, at least in the analyses we have been able to conduct so far, closely parallel the first-year year results. Controlling for the same set of confounding influences as in the first year but also for the clear-and-organized-instruction scale discussed above, reflective learning once again had a significant positive link with four-year growth in critical-thinking skills, while all three deep-learning scales had significant unique and positive associations with four-year growth in the need for cognition.

Thus, we found that deep-learning experiences may have positive cognitive effects that are independent of exposure to clear and organized instruction, while such exposure continued to have significant positive effects on four-year growth in critical-thinking skills and the need for cognition even when deep approaches to learning were taken into account. As far as we know, the WNS is the first longitudinal study to simultaneously consider the broad-based impacts of these academic experiences on general measures of cognitive growth over four years of college.

The Cognitive Impact of Liberal Arts Colleges

To date, the WNS has not uncovered compelling evidence that liberal arts college attendance has consistent *direct* effects on our measures of cognitive development during college. However, the evidence did suggest that the typical instructional/learning environment for liberal arts college students was significantly different from that of their counterparts at research universities or regional institutions.

Even with controls for student pre-college characteristics and academic major, students attending liberal arts colleges (as compared to their peers at research universities and regional institutions) reported significantly higher levels of clarity and organization in the instruction they received, as well as a significantly higher frequency of experiences on all three of the deep-learning scales. Since these kinds of experiences appeared to significantly enhance measures of four-year cognitive growth, we hypothesized that the most consistent cognitive impacts of liberal arts colleges might be indirect, transmitted through their instructional and learning environments.

Our testing of these indirect effects supported this hypothesis. Relative to their peers at both research universities and regional institutions, liberal arts college students realized significant advantages on both their critical-thinking skills and their need for cognition that were attributable to exposure to higher levels of instructional clarity and organization and more frequent deep-learning experiences.

> Students attending liberal arts colleges…reported significantly higher levels of clarity and organization in the instruction they received, as well as a significantly higher frequency of experiences on all three of the deep-learning scales.

In short, a significant part of the cognitive impact of liberal arts colleges may be exerted by their distinctive instructional and learning environments. But it is important to point out that students accrued the positive cognitive outcomes of both clear/organized instruction and deep-learning experiences irrespective of the type of institution they attended. The quality of instruction and deep approaches to learning may improve undergraduate learning at all types of institutions.

The Cognitive Impact of Diversity Experiences

The WNS put a strong emphasis on estimating the effects diversity experiences during college had on standardized measures of cognitive development—specifically on critical-thinking skills, the need for cognition, and a positive attitude toward literacy activities. Our end-of-fourth-year analyses are perhaps the most revealing.

From student responses on the 2010 experience-of-college instruments, we developed a nine-item scale (reliability: .80) that we termed "interactional diversity." This scale assessed the extent of students' engagement with diverse peers, ideas, and socio-political and religious perspectives during college. Some examples of items from the scale are how often respondents had had serious conversations with students with different races or ethnicities, religious beliefs, political opinions, or personal values, as well as how often they had participated in a racial- or cultural-awareness workshop.

In addition to controlling for our typical array of confounding variables, we took into account the effects of six other college experiences that have typically been considered in the discussion of high-impact practices in undergraduate education: worked with a faculty member on a research project, academic challenge and high expectations, co-curricular involvement and positive peer interactions, cooperative learning experiences, interactions with student affairs staff and faculty, and high-quality interactions with faculty.

Our findings were somewhat surprising. Net of the confounding variables and the influence of all other measures of good practices, only interactional diversity had a significant positive impact on all three cognitive outcomes by the end of the fourth year. Furthermore, interactional diversity was the only good-practice dimension to significantly foster four-year growth in critical-thinking skills.

Thus, the WNS findings contribute a reasonably solid piece of corroborating evidence to arguments that engagement with diversity in one's undergraduate experience is not merely the extension of a specific social/political agenda but may, in fact, have non-trivial implications for students' intellectual growth. As we shall see in the next section, however, it gets a bit more complicated.

The Conditional Impact of High-Impact / Good Practices

The findings from our analyses of the WNS data reinforce a major conclusion from the vast body of evidence on college impact: Irrespective of what happens during college, individual differences among entering students are typically the strongest predictors of college outcomes. If such differences exert such a compelling influence on college outcomes, it should come as no surprise that they might also play a significant role in moderating how college experiences shape the outcomes of college for different types of students.

"Engagement with diversity in one's undergraduate experience is not merely the extension of a specific social/political agenda but may, in fact, have non-trivial implications for students' intellectual growth.

When the estimated effects of a college experience differ in magnitude for different kinds of students, they are typically called *conditional* (or interaction) effects. These are different from the general effects we report above for clear and organized instruction, deep learning, and interactional diversity, which are based on the entire sample irrespective of individual variations in student characteristics.

As might be expected, conditional effects do not reveal themselves automatically; one has to take an extra step in the analysis to find them. When we took that additional step with the WNS data, we found conditional effects on four-year cognitive outcomes for a number of the good-practices dimensions measured in the WNS.

How much impact these practices had often depended on the characteristics of the students who experienced them. Race, ACT (or equivalent) score, level of pre-college development on the outcome being considered, and gender were all student traits that moderated the cognitive influence of one or more of these practices.

Race. The moderating effects of race were confined to four-year gains in critical-thinking skills and depended on the specific practice considered. For example, while interactional diversity had a significant effect on four-year growth in critical-thinking skills for the entire WNS sample, this finding masked conditional effects based on race.

Specifically, the significant positive impacts of interactional diversity on fourth-year critical-thinking skills accrued almost entirely to white students. Students of color derived no such statistically significant critical-thinking increases from their engagement with diverse peers.

> Diversity experiences tended to function in a compensatory manner in fostering the critical-thinking skills of students who seem to have been the least well prepared academically for college.

Conversely, we found no evidence to suggest that co-curricular involvement and positive peer interactions had a significant general effect on four-year critical-thinking skills for the WNS sample overall. However, this masked its significant positive effect on gains in critical thinking for students of color. Such findings suggest that there are situations in which white students and students of color benefit cognitively from different experiences.

ACT (or Equivalent) Score. Pre-college levels of academic preparation also moderated the impact of diversity experiences on students' four-year growth in critical-thinking skills. The positive effect of interactional diversity on critical-thinking skills for the entire WNS sample hid the fact that the greatest cognitive benefits of engagement with diverse others accrued to those students with ACT scores of 27 or less (or their SAT equivalent). For their counterparts with a score greater than 27, the effect of interactional diversity on critical-thinking growth was trivial. Thus, diversity experiences tended to function in a compensatory manner in fostering the critical-thinking skills of students who seem to have been the least well prepared academically for college.

Pre-College Levels of the Outcome. A similar compensatory pattern was found for low levels of pre-college critical-thinking skills and a positive attitude toward literacy activities. Interactional diversity and working with a faculty member on a research project counted the most for students who entered college in the lowest third on the pre-college measures. However, for students in the upper two-thirds of the pre-college distributions on the measures, diversity experiences and working with a faculty member on a research project had only non-significant effects by the fourth year.

The consistent pattern of these conditional effects was unequivocal. For example, the estimated positive effects of interactional diversity and working with a faculty member on four-year growth in critical-thinking skills were between

> Individual student characteristics will find a way to exert their influence and moderate the impact of college. While this makes our analysis less parsimonious, it also presents institutions with an opportunity to fashion programs and interventions targeted at specific groups.

three and ten times as strong for students who started college in the lowest one-third of the critical-thinking distribution as for students who began college in the upper two-thirds. Similarly, the positive effects of these practices on four-year increases in positive attitudes toward literacy activities were approximately five to six times as strong for students who entered college in the lowest one-third of the distribution on this measure as for students who began college in the upper two-thirds.

Gender. Despite the fact that men and women in the WNS made essentially the same net four-year gains in critical-thinking skills and a positive attitude toward literacy activities, men benefitted more than women from some high-impact practices. For example, the positive effect of high-quality interactions with faculty on men's critical-thinking skills by the fourth year was about .13 of a standard deviation, which was statistically significant. However, for women the corresponding effect was -.01 of a standard deviation and not significant. Similarly, interactional diversity increased positive attitudes toward literacy significantly for both men and women; however, the magnitude of the effect was almost three times as large for men as for women.

The WNS findings clearly suggest that high-impact practices accelerate student cognitive growth during college. However, they may do so in complex ways. The impact of some such practices (e.g., diversity experiences, high quality interactions with faculty, co-curricular involvement, and positive peer interactions) may depend on a student's race or gender, while some good practices (e.g., working with a faculty member on a research project and diversity experiences) may compensate for the academic disadvantages that some students bring to college. Such conditional influences should remind us that individual student characteristics will find a way to exert their influence and moderate the impact of college. While this makes our analysis less parsimonious, it also presents institutions with an opportunity to fashion programs and interventions targeted at specific groups that have the likelihood of benefiting those groups the most.

The Impact of Study Abroad on Intercultural Competence

Substantial human and financial resources have been invested in study-abroad programs for American undergraduates. Unfortunately, there is little systematic evidence concerning the effectiveness of these programs in fostering intercultural competence. What evidence does exist is characterized by weak cross-sectional research designs and a failure to take into account the selection effect (i.e., the likelihood that students who choose to study abroad may have higher levels of intercultural competence to begin with). The longitudinal, pretest-posttest nature of the WNS, as well as the fact that during the first year of college we asked students to indicate their "intent to study abroad," afforded us a unique opportunity to estimate the unique effect of study abroad on intercultural competence.

In a dissertation by Mark Salisbury, participation in study abroad was hypothesized to increase intercultural

> We found little evidence that Greek affiliation in the first year of college inhibits growth in liberal arts outcomes for either men or women, as has sometimes been claimed.

competence, as assessed by the Miville-Guzman Universality/Diversity Scale (MGUDS). The MGUDS measures the degree to which an individual is aware of and accepts both similarities and differences among people. Controlling for the selection effect, the good news was that students who studied abroad during college showed statistically significant greater increases in the MGUDS total score than similar students who did not.

The less-good news was that most of this effect was realized in only one of the MGUDS subscales: diversity of contact (a respondent's engagement in diverse social and cultural activities). Study abroad demonstrated only small and statistically non-significant net effects on relative appreciation of cultural differences or comfort with difference.

Of course study-abroad programs may vary dramatically in their rigor and efficacy, and the WNS was not able to take these differences among programs into account. The fact that the WNS uncovered these modest effects with the rather simple and global measure of participation in study abroad (the student did or did not participate) leaves open the possibility that some very effective and rigorously conducted programs may have substantially stronger impacts on intercultural competence than we report.

The Impact of Greek Affiliation

Controversies about Greek life in American colleges and universities have led to a growing body of research that attempts to understand its unique impacts. Most of this research, however, is cross-sectional and fails to take into account the selection effect. This means that it is easy to mistake the influence of the individual characteristics that lead students to join fraternities and sororities for the impact of fraternity or sorority membership itself.

As with study abroad, the pretest-posttest, longitudinal nature of the WNS permitted us to statistically control for a large number of these selection effects in estimating the impact of Greek affiliation. With such controls in place, we found little evidence that Greek affiliation in the first year of college inhibits growth in liberal arts outcomes for either men or women, as has sometimes been claimed. There was an essential parity between Greek-affiliated men and women and their independent same-sex counterparts in first-year gains in critical-thinking skills, need for cognition, positive attitude toward literacy activities, moral reasoning, and intercultural competence.

Moreover, we found additional evidence to indicate that fraternity members had significantly greater first-year growth than their non-fraternity counterparts in the citizenship and change scales of socially responsible leadership. Similarly, sorority members had significant net advantages over their non-sorority peers in first-year gains on the citizenship and common-purpose scales of the socially responsible leadership measure.

Concluding Thoughts

What we have learned has implications for institutional policies and practices. Most important is the finding that clear and organized classroom instruction counts, and not just in terms of course-level knowledge acquisition; it also fosters students' cognitive growth more generally and increases the probability that they will persist at the institution they are attending. And many of the constituent elements of clear and organized instruction are learnable by faculty. Thus, institutional programs designed to improve faculty instructional skills may pay off in broader ways than simply improving student achievement in individual courses.

Deep-learning experiences count. While they tend to be positively linked to clear and organized instruction, reflective and integrative learning activities had positive net impacts on growth in critical-thinking skills and the need for cognition that were independent of being exposed to such instruction. Additionally, integrative learning experiences appeared to independently enhance first-year growth in moral reasoning.

Even when student background characteristics and college academic major were taken into account, liberal arts colleges tended to expose students to higher levels of both clear and organized instruction and deep-learning experiences than did research universities or regional institutions. It was this distinctive teaching/learning environment that transmitted most of the cognitive benefits of attending a liberal arts college. However, these instructional approaches can have just as positive an influence at other types of institutions as well.

Our findings clearly support the argument that engagement with diversity during college has positive consequences for students' general cognitive development. However, as was the case with other high-impact practices that we considered (e.g., working with a faculty member on a research project, co-curricular involvement and positive interactions with peers, high-quality interactions with faculty), the impact of diversity experiences on cognitive growth was conditional: White students and those who entered college the least prepared academically appeared to benefit the most.

Similar conditional effects involving race, gender, and pre-college traits were uncovered in several other good practices. Such findings suggest that there may be no single silver bullet that works for all students, all the time. Rather, institutions should be prepared for the reality that individual differences among students have the potential to moderate the impact of any college experience.

The longitudinal, pretest-posttest design of the WNS permitted us to control for many of the selection-effect variables that have confounded the body of existing research on the impacts of study abroad and Greek affiliation. We found that study abroad positively influenced some dimensions of four-year growth in intercultural competence (diversity of contact) but not others (relative appreciation of cultural differences and comfort with difference). But the modest support offered by our findings could be attributable to our inability to consider qualitative differences in study-abroad programs.

Our analyses of the impact of Greek affiliation provided little evidence to indicate that fraternity or sorority membership typically inhibits gains in cognitive development, moral reasoning, or intercultural effectiveness—at least during the first year of college. Additional analyses we are now conducting will consider the impact of four years of Greek affiliation on these outcomes.

Finally, throughout our ongoing analyses of the WNS data we have been continually reminded of how important it is to take into account students' pre-college pretest scores on the outcome being studied. Time and time again, we found that, irrespective of the sophistication of the analytic procedures we used, controlling for the pretest was essential to achieving relatively unbiased estimates of the effects various college experiences had on any dependent variable. Indeed, when we experimented by deliberately failing to take the pretest into account, our choice of analytic procedure didn't matter; we still got substantially biased results. Research design appears to trump statistical approach. Thus, it would seem that, whenever possible, a pretest-posttest longitudinal design is best for studying the impact of college. C

Website

■ A more detailed description of the study's methodology and the sampling procedure, as well as the complete papers and research reports documenting the findings discussed, are available online at http://www.education.uiowa.edu/centers/crue/Publications.aspx#WNSLAE.

> **"The longitudinal, pretest-posttest design of the WNS permitted us to control for many of the selection-effect variables that have confounded the body of existing research.**