

Overview of Findings from the First Year of the Wabash National Study of Liberal Arts Education

Introduction

In this report, we provide an overview of the findings from the 19 institutions that participated in the Wabash National Study in 2006–2007. While reviewing the findings, we would like to call your attention to several key points you should keep in mind:

1. We have only followed the students for six months so far, a relatively short period of time. We both hope and expect to see larger changes after four years of college when we collect data again.
2. Although the study is longitudinal and we have statistically controlled for many important characteristics and qualities that students bring to college,¹ it is still important to exercise caution in drawing conclusions about the extent to which experiences cause students to change on learning outcomes. Since we cannot randomly assign students to colleges, or to experiences within those colleges, it is *possible* that the relationships we have identified between teaching conditions and institutional practices and student growth are mediated by some other as yet unidentified causes.
3. We have used well-known instruments and surveys to measure student learning; however, as we have indicated throughout this project, these measures are not as authentic as the exams, papers, and projects students complete in their work at their institutions. They do provide, however, an additional source of information that we can use in our efforts to enhance our institutions' impact on students.
4. The surveys that we use to examine what students are experiencing inside and outside of the classroom are all based on students' reports of how they are interpreting their experience. We do not tout such reports as "objective measures" of what they are encountering at their institutions. However, these self-reports are still informative because they give an idea of how students are experiencing their institutions, and because, self-report or not, they correlate with the extent to which students are changing on many of our outcome measures.

Review of National Findings

I – How much did students change on the outcome measures?

In the fall of 2006, 4,501 students from nineteen different institutions completed the first phase of the study, and in the following spring, 3,081 students returned to complete the second phase. In the discussion that follows below, we will look at the 3,081 students who completed *both* the fall and the spring phases of the study.

It is fair to say that we were surprised and somewhat disappointed by the lack of change on many of our outcome measures. The largest positive change nationally was students' growth in moral reasoning, which improved by about 10% over the first two semesters of college. We used the

¹ See footnote 4 for more detail on the background characteristics for which we have statistical controls.

Defining Issues Test-2² to gauge student’s moral reasoning, and students improved on both the traditional (P-score) and more recent (N2 score) measures of their ability to employ theoretical frameworks and moral ideals in their reasoning about complex moral dilemmas. On the other hand, although students’ improvement on the CAAP Critical Thinking test was statistically significant, the change was so small (less than 1% increase) that it was practically meaningless. (See Table 1 below)

Table 1: Spring/Fall Comparison – Critical Thinking and Moral Reasoning (average/standard deviation)

	Spring 2007	Fall 2006	Difference
CAAP Critical Thinking Test			
Critical Thinking score	64.36 / 5.47	63.79 / 5.15	0.57*
Defining Issues Test 2 (DIT2)			
P-score	41.44 / 15.93	37.49 / 15.22	3.96*
N2 score	40.56 / 15.40	35.89 / 15.05	4.67*

* statistically significant change at or below the 0.05 level

As shown in Table 2 below, students also changed very little on the different scales that we used to measure well-being and leadership. The Ryff Scales of Psychological Well-Being focus on six different components of well-being and positive psychological functioning, while the Socially Responsible Leadership Scale examines eight different aspects of leadership focused on creating positive social change. Although there are several statistically significant changes in the scales in Table 2, these changes are all very small.

Table 2: Spring/Fall Comparison – Leadership and Well-Being (average/standard deviation)

	Spring 2007	Fall 2006	Difference
Ryff Scales of Psychological Well-Being			
Self-Acceptance	4.55 / 0.81	4.59 / 0.78	-0.03
Positive Relations with Others	4.68 / 0.78	4.67 / 0.77	0.01
Autonomy	4.34 / 0.74	4.32 / 0.76	0.02
Environmental Mastery	4.35 / 0.75	4.38 / 0.69	-0.03
Purpose in Life	4.66 / 0.74	4.71 / 0.70	-0.05*
Personal Growth	4.69 / 0.67	4.71 / 0.64	-0.02
Socially Responsible Leadership Scale			
Consciousness of Self	3.98 / 0.55	3.94 / 0.55	0.03*
Congruence	4.17 / 0.57	4.17 / 0.56	0.00
Commitment	4.39 / 0.52	4.43 / 0.51	-0.04*
Collaboration	4.04 / 0.49	4.04 / 0.48	0.00
Common Purpose	4.02 / 0.48	4.01 / 0.46	0.00
Controversy with Civility	3.90 / 0.47	3.88 / 0.47	0.01
Citizenship	3.97 / 0.59	3.98 / 0.59	-0.01
Change	3.75 / 0.54	3.73 / 0.53	0.03*

* statistically significant change at or below the 0.05 level

² Please see the “Guide to Outcome Measures” at http://www.liberalarts.wabash.edu/storage/assessment-instruments/Guide_to_Outcome_Measures.pdf for detailed descriptions of all of the outcome measures.

Of greater concern were the small but consistent declines that we saw in students' attitudes about diversity. Table 3 below shows the changes in students' scores on the scales that we use to measure views about diversity. The questions in these scales focus on the interest, comfort, and appreciation that students have for people from diverse intellectual, cultural, and ethnic backgrounds. While these changes are often small, there is a consistent downward pattern across all of these measures.

Table 3: Spring/Fall Comparison – Attitudes toward Diversity (average/standard deviation)

	Spring 2007	Fall 2006	Difference
Miville-Guzman Universality-Diversity Scale			
Full Scale score	4.59 / 0.66	4.64 / 0.62	-0.06*
Diversity of Contact Subscale score	4.19 / 0.98	4.27 / 0.95	-0.08*
Relativistic Appreciation Subscale score	4.74 / 0.74	4.83 / 0.66	-0.09*
Comfort with Differences Subscale score	4.82 / 0.81	4.83 / 0.78	-0.01
Orientation toward Learning Scales			
Openness to Diversity and Challenge	3.78 / 0.70	3.92 / 0.62	-0.15*

* statistically significant change at or below the 0.05 level

Finally, we also observed a somewhat larger and consistent decline in students' academic motivation and their interest in academic subject matter, community involvement, and professional success. We found these declines to be especially troublesome because these attitudes and values may shape the extent to which students engage with subsequent college work.

Table 4: Spring/Fall Comparison – Life Goals and Academic Orientation (average/standard deviation)

	Spring 2007	Fall 2006	Difference
Life Goals and Orientation toward Learning Scales			
Contribution to the Arts	1.76 / 0.79	1.85 / 0.78	-0.09*
Contribution to the Sciences	1.75 / 0.86	1.90 / 0.83	-0.15*
Political and Social Involvement	2.58 / 0.58	2.68 / 0.51	-0.10*
Professional Success	2.35 / 0.70	2.46 / 0.68	-0.11*
Academic Motivation	3.35 / 0.60	3.60 / 0.55	-0.25*
Positive Attitude toward Literacy	3.30 / 0.79	3.35 / 0.74	-0.05*
Need for Cognition Scale			
Need for Cognition score	3.50 / 0.63	3.51 / 0.61	-0.02

* statistically significant change at or below the 0.05 level

How do we explain these findings? First, it is important to remember that we are only looking at students' first six months of college. Given both the very short amount of time and the challenges of adjusting to college life, it may not be surprising to see such small changes on our outcome measures. Colleagues who have seen these results have also suggested that students might be "underwhelmed" by their college experience. Students might have entered college with unrealistic expectations and found the reality of college life somewhat disappointing. Others have suggested that some experiences, such as thinking about things from a different perspective and interacting with people from different backgrounds, might have appealed to students in the

abstract when they entered college, but then turned out to be quite challenging—and possibly even a little uncomfortable—once they encountered the reality of college. We have not yet evaluated these possibilities, but our findings about the levels at which students are experiencing good teaching practices and supportive institutional conditions may help us understand part of the reason that student growth is, overall, so low.

II – What teaching practices and institutional conditions appear to be influencing student growth?

Despite the fact that students do not appear, overall, to be changing in a positive direction in the first year, we have identified a number of teaching practices and institutional conditions that predict gains on virtually all of the outcomes we are measuring in the study. This finding is especially important because it points to conditions that may, over the course of the next two years, boost the small changes we have seen to date.

We asked students many different questions about their experiences in college. In order to make sense of this large amount of information, we first had to find a way to reduce the data into meaningful and more manageable chunks. We used a statistical technique called factor analysis to examine student responses to questions about their college experiences and to locate “clusters” or subsets of questions that focused on core elements of student experience. All of the questions in an identified cluster measure some underlying component of student experience. Using common social science language, we refer to these clusters as “scales” because we can add the student responses on each question in the scale to get an overall score for the scale. We can then look to see if higher and lower scores on the scale predict the extent to which students changed over the course of the first year on our outcomes.

We ultimately identified six distinct scales of teaching practices and institutional conditions:

- *Good Teaching and High-Quality Interactions with Faculty*
- *Academic Challenge and High Expectations*
- *Diversity Experiences*
- *Frequency of Interacting with Faculty and Staff*
- *Interactions with Peers*
- *Cooperative Learning*

Three of these scales predicted student growth on a wide range of outcomes.³ The relationship between the remaining three scales and growth on the outcomes was much more limited.

The three scales that predicted growth on many of our outcome measures are as follows:

1. *Good Teaching and High-Quality Interactions with Faculty* – Includes 23 questions asking students about their experiences with the following:
 - a. Faculty interest in teaching and student development
 - Example – To what extent do you agree that most faculty with whom you have had contact are genuinely interested in students?
 - Example – To what extent do you agree that most faculty with whom you have had contact are genuinely interested in teaching?

³ We refer to these three clusters as “good practices scales” elsewhere in our reports. To learn more about the scales, read “Effective Practices and Experiences from the Wabash National Study” at http://www.liberalarts.wabash.edu/storage/Effective_Practices_and_Experiences.pdf

- b. Prompt feedback
 - Example – How often have faculty informed you of your level of performance in a timely manner?
 - c. Quality of nonclassroom interactions with faculty
 - Example – To what extent do you agree that your nonclassroom interactions with faculty have had a positive influence on your intellectual growth and interest in ideas?
 - d. Teaching clarity and organization
 - Example – How often have faculty given clear explanations?
 - Example – How often have faculty made good use of examples and illustrations to explain difficult points?
2. *Academic Challenge and High Expectations* – Includes 31 questions asking students about their experiences with the following:
- a. Academic challenge and effort
 - Example – In your experience at your institution during the current school year, how often have you worked harder than you thought you could to meet an instructor's standards or expectations?
 - b. Frequency of higher-order exams and assignments
 - Example – How often have exams or assignments required you to argue for or against a particular point of view and defend your argument?
 - c. Challenging classes and high faculty expectations
 - Example – How often have faculty asked you to point out any fallacies in basic ideas, principles, or points of view presented in the course?
 - d. Integrating ideas, information, and experiences
 - Example – To what extent do you agree that courses have helped you understand the historical, political, and social connections of past events?
 - Example – In your experience at your institution during the current school year, how often have you worked on a paper or project that required integrating ideas or information from various sources?
3. *Diversity Experiences* – Includes 9 questions asking students about their experiences with the following:
- a. Diversity experiences
 - Example – How often have you attended a debate or lecture on a current political/social issue during this academic year?
 - Example – In your experience at your institution during the current school year, how often have you had serious conversations with students who are very different from you in terms of their religious beliefs, political opinions, or personal values?
 - b. Meaningful discussions with diverse peers
 - Example – How often have you had discussions regarding inter-group relations with diverse students (e.g., students differing from you in race, national origin, values, religion, political views) while attending this college?

After controlling for a wide array of qualities that students bring with them to college, higher levels of the experiences identified in these three scales predict growth on nearly every outcome measure that we have in the study.^{4,5} (See Table 5 below)

Table 5: Correlation between Good Practices and Outcomes (+/- indicates significant positive or negative relationship)

	Good Teaching	Academic Challenge	Diversity Experiences
CAAP Critical Thinking Test			
Critical Thinking score	+		+
Defining Issues Test 2			
P-score		+	
N2-score	+	+	
Ryff Scales of Psychological Well Being			
Self-Acceptance	+	+	-
Positive Relations With Others	+	+	
Autonomy	+	+	+
Environmental Mastery	+	+	
Purpose in Life	+	+	-
Personal Growth	+	+	+
Socially Responsible Leadership Scale			
Consciousness of Self	+	+	
Congruence	+	+	+
Commitment	+	+	
Collaboration	+	+	
Common Purpose	+	+	+
Controversy w/ Civility	+	+	+
Citizenship	+	+	+
Change		+	+
Miville-Guzman Universality-Diversity Scale			
Full Score	+	+	+
Diversity of Contact	+		+
Relativistic Appreciation	+	+	+
Comfort with Differences	+	+	+
Orientation toward learning scales			
Openness to Diversity	+	+	+

⁴ We statistically controlled for an incoming student’s score on each outcome measure; for gender, ethnicity, parental education, the student’s age, whether or not the student had dependents, the ethnic/racial composition of the student’s high school, the student’s high school involvement, the student’s educational goals, full-time enrollment, whether or not the institution was the student’s first choice, the student’s incoming academic motivation, the type of institution the student attends; and for whether or not the student is an athlete, belongs to a fraternity or sorority, lives on campus, and is working.

⁵ To determine whether each factor made a unique contribution to change in the outcome, we also entered all of the factors into the analysis simultaneously.

Life goals and orientation toward learning scales			
Contribution to the Arts		–	+
Contribution to the Sciences			
Political Social Involvement	+	+	+
Professional Success		+	
Academic Motivation	+	+	–
Positive Attitude Toward Literacy	+	+	+
Need for Cognition Scale			
Need for Cognition score	+	+	+

As you can see, student responses to the questions on the *Good Teaching and High-Quality Interactions with Faculty*, *Academic Challenge and High Expectations*, and *Diversity Experiences* scales correlate significantly with growth on most of the outcomes that we measured. The *Good Teaching* questions had a significant positive impact on 24 of the 29 outcomes, the questions on *Academic Challenge* had a significant positive effect on 25 of the 29 outcomes, and the questions on *Diversity Experiences* had a significant positive impact on 17 of the 29 outcomes.

We should point out that there were also a few negative effects. For example, students who reported experiencing high levels of *Academic Challenge and High Expectations* were more likely to decline on the extent to which they valued making a contribution to the arts, and students with high levels of *Diversity Experiences* were more likely to have lower scores on two of our well-being scales and on academic motivation. However, overall, *Good Teaching and High-Quality Interactions with Faculty*, *Academic Challenge and High Expectations*, and *Diversity Experiences*, as defined by the questions in these scales, all have a positive impact on outcomes, ranging from critical thinking and interest in doing difficult intellectual work to well-being and leadership.

It is important to note that we did not begin our research with the idea that particular practices or conditions were more effective than others. Rather, we identified the experiences in these three high-impact scales based on evidence from this study. These findings are especially interesting in light of the fact that they point to the impact of well-known good practices that *are not* contingent, as far as we know, on a specific institutional context or academic discipline.

As we indicated earlier, there were three scales measuring familiar teaching practices and conditions that had a much weaker or more mixed relationship with the outcomes that we examined:

- *Frequency of Interacting with Faculty and Staff* – While student reports that faculty were genuinely interested in their development or had a positive influence on their personal growth was correlated with student development on the outcomes,⁶ the *frequency* with which students interacted with faculty and staff influenced only a few outcome measures, including academic motivation, the desire to contribute to the sciences and the arts, and the desire for professional success. Although these are important outcomes, the frequency of interacting with faculty and staff also had a *negative* relationship with the development

⁶ These two questions are among the questions in the *Good Teaching and High-Quality Interactions with Faculty* scale.

of moral reasoning and critical thinking. This last result may occur because many students who interact frequently with faculty may be doing so because they are having academic difficulty. In any case, the small impact of this cluster of questions may indicate that the quality of students' interactions with faculty and staff matters more than the quantity of these interactions.

- *Interactions with Peers* – The degree of positive peer interactions and amount of cocurricular involvement were positively related to growth on our measures of leadership and well-being. However, they were negatively related to academic motivation, the desire to engage in challenging intellectual work, and the desire to be politically and socially involved in the community.
- *Cooperative Learning* – Working in study groups or with peers on projects in and out of class was positively related to growth on some of our measures of leadership, to students' desire to contribute to the sciences, and to students' desire for professional success, but little else. Cooperative learning is a familiar and important pedagogy, but based on our analyses, once we take into account the kinds of basic good practices that are captured in the *Good Teaching and High-Quality Interactions with Faculty*, *Academic Challenge and High Expectations*, and *Diversity Experiences* scales, cooperative learning does not add much in predicting student growth. One possibility that could account for this surprising finding is that faculty and staff who are likely to utilize this pedagogy are also people who are already engaging in practices included in our *Good Teaching*, *Academic Challenge*, and *Diversity* scales, so engaging in cooperative learning as a specific pedagogy is not contributing an additional impact beyond those practices.

We were surprised by the fact that these three scales did not correlate with many of our outcome measures, but that is what the data indicate to this point.

III – How frequently are students experiencing these effective teaching practices and conditions?

Given our finding that three core sets of student experiences are correlated with student growth on our outcomes, the next obvious question is, “How often are students experiencing these important practices and conditions?” To get at this question, we created three simple categories to describe the extent to which students are getting high scores on these scales of good teaching practices and institutional conditions.

- **Strong** for students who typically responded “frequently” or “often” to questions about the good practices and conditions
- **Moderate** for students who typically responded “sometimes” or “occasionally” to questions about the good practices and conditions
- **Weak** for students who typically responded “never” or “rarely” to questions about the good practices and conditions⁷

There is both good and challenging news in the results of this analysis. As shown in Table 6 below, a substantial portion of students at all institutions report high levels of the experiences in the *Good Teaching and High-Quality Interactions with Faculty* scale, and a small proportion of

⁷ To learn more about how we developed these three categories, read “Appendix: Classification System for the Good Practice Scale Scores” at http://liberalarts.wabash.edu/storage/GP_scales_Appendix.pdf

students report very low levels of these conditions. On the other hand, the majority of students report that they are experiencing these good practices and conditions only “sometimes” or “occasionally.”

Table 6: Good Teaching and High-Quality Interactions with Faculty

	Strong	Moderate	Weak
Small Institutions	44%	56%	<1%
Large Institutions	28%	72%	1%

The same pattern is clear for the experiences measured by the *Academic Challenge and High Expectations* scale. However, the proportion of students in the strong category is lower, while the proportion in the middle category is even greater, regardless of the type of institution that students attend.

Table 7: Academic Challenge and High Expectations

	Strong	Moderate	Weak
Small Institutions	26%	74%	<1%
Large Institutions	18%	82%	0

Table 8 shows a different pattern for *Diversity Experiences*. A considerable portion of students at all institutions report “never” or “rarely” having these kinds of experiences, and there are very few students who report having the highest levels of the teaching practices and institutional conditions identified in this scale.

Table 8: Diversity Experiences

	Strong	Moderate	Weak
Small Institutions	5%	84%	12%
Large Institutions	3%	83%	14%

A critical point to bear in mind in reviewing these tables is that the different categories of experience—strong, moderate, and weak—predict different levels of growth on the outcomes. Students who report higher levels of the experiences measured by these three scales are more likely to grow on the outcomes than students who have moderate levels of these experiences. Students with moderate levels of these experiences are also more likely to grow than students with the weakest level of these experiences. Given concerns about the validity of students’ self-reports, we feel it is important to emphasize this point. Although students’ reports about their college experiences may not always be entirely accurate, these three categories predict different levels of growth on our outcome measures.

This brings us back full circle to our finding that, on the whole, students change very little on the outcomes that we have measured over their first year in college. Despite the fact that students do not change, we have also found a set of teaching practices and conditions that predict student growth on the outcomes. This seems to suggest that across the institutions in this study, these effective teaching practices and institutional conditions are not prevalent enough to produce widespread change. Indeed, the data from the first year of the study also indicate that most of our students are experiencing moderate levels of these important practices and conditions. Thus, one hope for improving student growth on the outcomes is to expand the degree to which students encounter these supportive conditions and practices. It is, of course, true that students themselves

bear some responsibility to seek out these conditions and practices, and as part of the course of our work over the next year or so, we will talk with students about the factors that may impede many of them from more deeply engaging with their education. At the same time, we bear an educational responsibility to find ways to strengthen the impact of their education at our institutions.

Conclusion

We have only begun to analyze the data from the first year. Among the important pieces that are forthcoming is information about whether the conditions and practices differentially impact students from different backgrounds, about the impact of alcohol consumption, and many other analyses. We will also work with our colleagues at study institutions to help them translate this information into positive action on their campuses.

Please do not hesitate to contact us if you have any questions about these analyses or the study, or if you would like to talk about ways that we can help you use this information to improve student learning at your institution.

Charles Blaich
Director of Inquiries

Kathy Wise
Associate Director of Inquiries