

Vassar College Assessment Committee

End of Year Report, June 28, 2013

This report is organized as follows. Part I is a narrative, detailing the subject matter discussed at our 12 meetings held during Academic Year 2012-13. Part II will be a summary of the statistical analysis of the Wabash Study conducted this year to be provided by David Davis-Van Atta. Part II will include numerous tables and charts prepared by Mr. Davis-Van Atta, Director, Office of Institutional Research, Vassar College, to whom we express our sincere thanks for his major contributions to the Committee's deliberations this year. Throughout the narrative, references to specific tables and appendix documents will be provided.

The committee consisted of three faculty members, Mark Andrews, Shirley Johnson-Lans, and Alison Keimowitz; Dean Eve Dunbar; Director of Institutional Research, David Davis-Van Atta; and Susan Ward, Administrative Assistant LTR, who assisted us greatly by taking minutes at the meetings. David Davis-Van Atta, as mentioned above, provided invaluable assistance by leading us through the complex Wabash National Study of Liberal Arts Education (WNS) and undertaking extensive statistical analysis of the Vassar WNS data. At the end of the year, he also provided a comparative study of two peer colleges within the Wabash Study, Connecticut and Hamilton colleges. The Committee had fewer members than has been the case in some recent previous years, and it was further diminished after Spring Break when Alison Keimowitz unfortunately had to withdraw from the committee for medical reasons.

The first meeting of the Committee was devoted to organizing ourselves and our work. Shirley Johnson-Lans was chosen to chair the Committee. Dean Jonathan Chenette gave the committee the task of analyzing the findings of the Wabash National Study of Liberal Arts Education (the WNS) in order to see what inferences could be drawn from that study about Vassar students' academic achievement and motivation that would help in the assessment of our educational process that was mandated by the Middle States Committee at the time of the College's Re-accreditation in 2009 and which we need to undertake in any event in order to evaluate the quality of education that Vassar is providing to students.

Our next meeting was held on October 24th at which time Mr. Davis-Van Atta provided the other members of the Committee with binders providing an overview of the Wabash study including sample questionnaires administered in the several parts of the study. He walked us through what the study is about, what universities and colleges participated in the Wabash project, and the process involved at Vassar where students of the Class of 2011 were given the two standardized tests and several surveys during their freshman orientation in Fall 2007 and again at end of their senior year. Although Mr. Davis-Van Atta had only received the data in September, 2012 (about six weeks before the meeting) he had already done a good deal of preliminary analysis and was able to describe details of the way the tests were administered at Vassar. We learned that the freshmen class consisting of 677 students was assembled to answer the questionnaires, but during the test a student asked whether doing so was required and was informed that participation was voluntary. At that time some students left, so only 479 students completed the Wabash questionnaires as entering students. At many other schools, students were re-examined at the end of their first year. At Vassar this failed because only a handful of students (13) were willing to participate. Vassar re-administered the Wabash surveys to seniors in the Spring of 2011 at which time 390 students completed the exercise. In order to get this number, Vassar provided approximately \$20,000 in a group subsidy (based on number of students participating) for Senior Week activities. The Vassar subsample of 283 students is composed of those who had participated both upon entering Vassar and at the end of the four years.

The Committee members immediately raised the question of possible sample bias. Mr. Davis-Van Atta agreed to provide an analysis comparing the personal characteristics (including academic achievement level, division of major, race/ethnicity, gender, U.S. vs. international student status, and socio-economic status) of the subsample of students in the WNS sample vs. the Class of 2011 as a whole.

Three more committee meetings were held in the first semester. At the first of those meetings Mr. Davis-Van Atta reported that he had found no significant differences between Vassar Class of 2011 participants and non-participants in the Wabash study with respect to socio-economic status, ethnicity, gender, academic achievement (GPA at time of graduation) or division of the curriculum in which students had majored. *[Please see first two pages of statistical appendix entitled, "Comparisons of Various Key Populations and Characteristics"]*

There remains, of course, the possibility of some sample bias based on unmeasured attitudinal characteristics of the participants, "the WNS cohort". [Please see the Appendix Report on WNS sampling.] However, we decided to proceed on the assumption that the sample fairly represented the Vassar Class of 2011.

The Committee realized that analyzing the Wabash Study would be a formidable task and that it would be necessary to focus on a limited number of questions. We decided that our main concern should be with academic dimensions of the student experience at Vassar and that we should focus on changes between freshman and senior responses. We also decided that it would be more fruitful to first analyze the CAAP and DIT-2 tests which are more about reasoning ability (and are scored based on questions that have wrong and right answers) and to postpone looking at the “attitudinal” sections of the Wabash Study which focus on academic motivation, and in particular changes in motivation between entering and completing the Vassar undergraduate education.

At our next meeting, we took a slight detour and examined Vassar students responses in the COFHE (21 peer colleges) Senior Survey, 2010, which provide information that was deemed complementary with the Wabash Study. This is a data set that last year’s Assessment Committee considered in its study of the Quantitative Requirement. This survey includes self-reported outcomes in quantitative areas. Mr. Davis-Van Atta provided us with comparative outcomes on the items on Using Quantitative Tools, Understanding Scientific Method, and Evaluating the Role of Science and Technology. In these areas, Vassar students (sample size 450) were found to be at the low ends of these scales, compared with the other schools in the study, with over half of our student participants rating quantitative skills to be of low or no importance. We are also at the low end of the distribution with respect to the proportion of our students majoring in the sciences.

Returning to our analysis of the Wabash study at our next meeting, the Committee members turned their attention to familiarizing ourselves with the actual outcome measures included in the surveys. We reaffirmed that the areas in which we could most profitably devote our time would be the outcomes in the Collegiate Assessment of Academic Proficiency (CAAP) Critical Thinking Test and in the Orientation toward Learning (academic motivation). However, there was also some interest in the Defining Issues Test (DIT). About half of our participating Vassar students were given the CAAP test and the other half were given the DIT.

During the second semester we decided to focus on Academic Motivation and Academic Achievement and to begin by examining the CAAP (Collegiate Assessment of Academic Proficiency) and DIT-2 (Defining Issues Test), since they were the two most objective measured learning outcomes available. We first looked at the Vassar outcomes on the CAAP, based on analyses we requested from Mr. Davis-Van Atta. Our first set of questions that we hoped the data analysis could inform, were whether there were systematic differences between students based on their academic interests (division of major, for instance), academic achievement level (GPA) or demographic characteristics. We began by looking at statistical data consisting of simple (two-way) correlations between personal characteristics of the Vassar sample and changes in CAAP scores between freshman and senior years, as well as correlations between initial standings in freshman year and changes over time. The Committee also hoped to gain insights from looking at results of multivariate regression analyses of outcomes, and this was to be the next step in our attempts to gain a handle on what implications we could draw from the CAAP results. We focused on a few personal characteristics such as senior year GPA, Race/Ethnicity, Gender, Division of Major, and comparison between initial CAAP score and change in score from freshman to senior year. Because of limitations imposed by sample size, additional variables of interest, such as U.S. citizen or resident/ international student status, were omitted from our study. By the end of February, we had

opined that there were no observable systematic relationships between personal characteristics of the students (demographic or academic) and their changes in scores on the CAAP between entering freshmen and end of the senior year tests. There were overall slight improvements in scores, and the improvements were greater for those scoring lower on the initial test. Thus it appeared that there was a simple “reversion to the mean”. However, the improvement in average score was very modest (only about one point.) This was what one would expect since the test is probably more a measure of aptitude than of the kind of achievement that would be likely to change as a result of more years of education. Using the average scores in the Wabash study as a whole, we noted less improvement at Vassar than in the overall averages for all participating schools but also much higher initial scores at Vassar. This was not surprising. As the scoring involves a “ceiling,” there is less room for improvement when initial scores are in the higher part of the distribution. *[Please see Table 4 in the attached statistical appendix]*

By late February, the Committee was able to look at the results of the multivariate analysis. There were no significant regression coefficients (partial effects of any of the explanatory variables) nor were any of the R^2 s of reasonable magnitude. In other words, moving from the simple correlations to more elaborate multiple regression models provided no additional insight into the CAAP scores or the change in CAAP scores of the Vassar students. *[Please see second half of page 2 of Table 4]*

The Committee members concluded that we should report out that any changes in the CAAP scores are small and should not be the focus of much attention, particularly as they have not been identified as associated with any systematic differences between students of different demographic or academic achievement groups.

In March we turned to an analysis of the DIT-2 scores and changes in scores between freshman and senior year. The results were similar to what we found when we analyzed the CAAP scores. The DIT was deemed by us to include more nuanced or in-depth questions. The positive change in average scores from entering freshman to end of senior year results were a little larger in the DIT than in the CAAP (about 6.5%), however, again socio-economic status (SES), academic measures (SAT scores, senior GPA), and other demographic variables were not significantly correlated with any DIT-2 outcomes that we studied. *[Please see Table 5 in the attached statistical appendix]*

At this time, we turned our attention to the Academic Motivation Index. The “Academic Motivation index” (AMI) is an ad-hoc WNS measure formed as a composite of eight individual behaviors and attitudes related to academic work (e.g. working for grades vs. interest in learning, frequency of coming to class prepared or not, reading more than just what is assigned, or not, enjoying working on hard problems, etc.). Vassar’s average AMI dropped from beginning of freshman year (which we interpret as measuring behaviors and attitudes in high school) to end of senior year. About 70 percent of the WNS participants showed a decline; 25 showed a gain, and 5 percent showed no change. There was found to be an inverse relationship between the magnitude of the decline and the initial AMI level at entering freshman year. There were no statistically significant correlations found with race/ethnicity, gender, or foreign citizenship. There did however appear to be a greater than average decline among those who had majored in the arts/humanities divisions of the Vassar curriculum. Changes in AMI scores at Vassar were also correlated with level of faculty/student interaction, quality of relationship with faculty, frequency of working harder on an academic assignment or subject than was necessary, frequency

of coming to class prepared, hours spent on homework, frequency of making a class presentation, frequency of discussing work or grades with a faculty member, and level of prompt feedback on homework, etc. The degree of decline was found to be negatively correlated with hours per week spent socializing with peers and hours per week spent in extra-curricular activities. [Please see Table 3 in the attached statistical appendix and “Summary of Selected Bivariate Correlations between Change in Academic Motivation scores and various activities/behaviors a (as seniors)” which includes multi-color charts and is placed in the Appendix following Table 3.]

Although the decline from freshman to senior year in AMI scores was characteristic of the colleges that participated in the WNS, Vassar’s average AMI score declined more than the average for the WNS college group as a whole, although each participating college showed a decline in this index from beginning of freshman year to senior year.

Please note that the freshman were administered these tests and surveys before they had even begun Vassar classes, and thus the Committee felt comfortable in interpreting these as reflecting attitudes and behaviors that students had at the time of high-school graduation. Although no formal statistical tests confirmed this, Mr. Davis-Van Atta reported to the Committee that it appears that “the stronger the academic rigor of a college and/or the greater its selectivity in admission, the greater the freshman-to-senior decline in AMI.”

The Committee next studied the findings on “Good Practice Measures”, within the Academic Motivation part of the Wabash surveys. We discussed at length how one might interpret Vassar’s apparent “worse scores” on several of these measures. In addition to the obvious hypothesis that something is amiss or lacking in the Vassar educational process, an alternative hypothesis was proposed that Vassar seniors have, as a result of their four years of education, developed more self-sufficiency and confidence in studying and therefore value interaction with faculty less than do students at many of the other colleges in the Wabash Study.

It was decided that Factor Analysis would be a very good technique to further analyze the AMI findings. Student-Faculty Interaction was one dimension that was studied using this technique. However, no systematic relationships were uncovered between the student answers to a composite measure of value of student-faculty interaction and other variables of interest current available in the Wabash Data.

Our speculations led us to want to look at correlations between student responses to this set of questions and several other measures of intellectual achievement and/or maturity as scholars. For instance, we think it would be useful to look at whether students are elected to Phi Beta Kappa or not (create a dummy, 0/1, variable for this) and also to look at whether students received departmental honors in their major. We also thought it would be useful to look at those who double majored and to single out those who had experienced some independent work with a faculty member, whether in a summer program such as the URSI or Ford programs or in connection with a senior thesis or science experiment. If possible we would also like to look at AMI outcomes by individual major field as well as by curricular division. Such data are available at Vassar and can be merged with the Wabash data. Mr. Davis-Van Atta explained that this can be done over the coming summer.

We reiterate that we have, given the data currently available, been unable to unearth any dynamics in the Vassar educational process that could explain the findings on changes in AMI responses from time of entering Vassar to end of senior year.

The final part of our analysis of the Wabash Study was made possible when Connecticut College and Hamilton College agreed to share some of their summary findings from the Wabash National Surveys administered at their colleges. However, given the agreement of confidentiality, we are not reporting out any of the findings from these two other schools. However, we did find them reassuring in that these in more similar schools, in terms of academic selectivity, overall exhibit outcomes closer to ours. [*Please see, "Summary of Findings for Vassar Compared to Two Peer Colleges" which follows Table 5 in the Appendix and which was provided by Mr. Davis Van-Atta to supplement this report.*]

After Spring Break the Committee also considered recommendations to next year's Committee and also recommendations to FPCC about the Assessment Committee Structure. There is unanimity in the opinion that it would be better to have longer terms for elected members of the Committee, given the length of the learning curve and probability that the Committee is often involved in successive year studies of related issues. A two-year term was recommended as we thought longer terms would have a discouraging effect on the willingness of faculty members to serve on the Committee. We also thought it would make sense to have staggered terms if possible, and we recommend that the composition of the faculty members be: two tenured and one un-tenured member, with representation from three different academic divisions.

We recommend that next year's Committee continue the analysis of the Wabash Study and try to learn more about its implications for policy recommendations relating to Vassar's educational mission. We hope that next year's Committee can increase the understanding of our students' seeming decline from freshman to senior year in the "Good Practices" part of the AMI index. We recommend studying possible correlations (or lack of them) between the responses to the "Good Practices" questions and measures of learning outcomes. This analysis could throw some light on the validity of the "Good Practices" questions, which seemed to us to contain a good deal of hard-to-interpret subjective information.

Finally, our bottom line conclusion, after a year of working to understand the Wabash data and draw what conclusions we can from it, is that we do not yet have a handle on understanding the implications of the results of the Vassar student responses to the WNS and in particular to their responses to the Academic Motivation Index. However, we think that our finding of no correlation between our student respondents' demographic or academic characteristics (measured by SAT scores and GPA at end of senior year) and their responses to the CAAP, DIT or AMI is in itself an important finding.

Respectively submitted,

Shirley Johnson-Lans

Chair, Assessment Committee

Academic Year, 2012-13.

