Running Head: MAJORS MATTER

Majors Matter: Differential Performance on a Test of General College Outcomes

Jeffrey T. Steedle, Council for Aid to Education

Michael Bradley, New York University

Paper presented at the 2012 Annual Meeting of the American Educational Research Association in Vancouver, British Columbia

Please send correspondence to:
Jeffrey T. Steedle
Council for Aid to Education
215 Lexington Ave., 21st Floor
New York, NY 10016
212-217-0728
jsteedle@cae.org

Abstract

The study described here investigated differential performance on the Collegiate Learning Assessment (CLA), an open-ended test of critical thinking and writing skills. Analyses revealed significant differences in performance between seniors at four-year colleges in different fields of study. Specifically, students studying natural sciences, social sciences, and humanities and languages scored the highest, and students studying business and education scored the lowest. Significant differences persisted after controlling for entering academic ability, sex, race, and language spoken at home, which is consistent with the thesis that students in certain fields of study gain more critical thinking and writing skills during college. There was no significant interaction between students' fields of study and the content of the CLA tasks, which suggests that general college outcomes may be measured without great concern for the confounding influence of content knowledge on task performance.

Majors Matter: Differential Performance on a Test of General College Outcomes

The practice of setting goals for general college outcomes in domains such as critical thinking and writing has become widespread, with 78% of Association of American Colleges and Universities member institutions reporting that "they have a common set of intended learning outcomes for all their undergraduate students" (Hart Research Associates, 2009, p. 1). These so-called higher-order, 21st century skills are said to cut across academic disciplines, but students practice them most in their chosen field. Thus, instructors in all content areas share responsibility for teaching these skills. On this issue, W. Robert Connor, President of the Teagle Foundation, recently posed the question, "Do majors matter?" (2011). Put another way, do students graduating with degrees in different fields of study demonstrate different levels of competency on general college outcomes?

The research reported here addresses this question by investigating differential performance on the Collegiate Learning Assessment (CLA), a test that is said to measure the critical thinking and writing skills espoused in general learning outcomes statements. The CLA performance of graduating seniors in 7 fields of study was compared before and after controlling for four student characteristics: entering academic ability, sex, race, and language spoken at home. An additional component of the analysis explored whether academic domain knowledge interacted with CLA task content. For instance, science majors may perform better on tasks requiring them to analyze scientific evidence and claims. The analysis employed two-way ANOVA including factors for field of study, task, and their interaction.

The results of this research illuminate differences in general college outcomes between fields of study. This provides guidance for future research into the specific causes for such differences and could lead to improvements in postsecondary educational programs.

Background

This work follows several studies that examined differential performance on the CLA. The first such study revealed significant differences in CLA performance between fields of study among seniors in the spring of 2006 at Kalamazoo College (Sotherland, Dueweke, Cunningham, & Grossman, 2007). After controlling for students' SAT scores, foreign language, humanities, and social science majors had higher CLA scores than natural science majors (these results were replicated in Kalamazoo College's 2007 CLA results, Sotherland, 2009).

In a follow-up analysis, Sotherland (2009) examined differences in freshman year (fall 2005) CLA scores based on students' majors upon graduation in spring 2009. While senior CLA scores in spring 2009 displayed differential performance by field of study, freshmen who were "good at the CLA" did not gravitate toward majors with exceptional CLA performance at senior year. This eliminated self-selection as an explanation for differential CLA performance in the senior year at Kalamazoo College.

To investigate the possible interaction between students' fields of study and CLA task content, Shavelson (2009) analyzed the performance of three fields of study on CLA tasks categorized into one of three content areas: social sciences, science/engineering, or humanities. Students majoring in the social sciences scored higher than all other fields of study. However, the performance differences between social sciences and science/engineering majors and between social sciences and humanities majors were small on tasks involving science/engineering and humanities content, respectively.

Among other investigations described in the highly publicized book *Academically Adrift,* Arum and Roksa (2011) examined relative performance on the CLA of different fields of study. After controlling for freshman CLA performance, students studying science/mathematics or

humanities/social sciences performed the best as college sophomores, and students studying business or education performed the worst. Differences among fields of study were reduced substantially, but not eliminated, after controlling for academic demands, social background, academic preparation, and institutions attended. Follow up analyses revealed that four-year gains on the CLA were greater among students majoring in the social sciences, humanities, natural sciences, and mathematics than students majoring in business, education, social work, and communications (Arum, Roksa, & Cho, 2011). Some component of these differences was attributable to different reading and writing requirements in different majors.

With a greater sample size, more fields of study, and tasks treated individually (rather than rough groupings), the research described here provides analysis at a finer grain than previously attained. The research reported here addresses two questions:

1. After four years of college, do students in different fields of study perform differently on CLA Performance Tasks before and after controlling for entering academic ability?
2. Are students in certain fields of study significantly advantaged or disadvantaged on CLA Performance Tasks because of task content?

Method

Measures

In the CLA, one 90-minute *Performance Task* is randomly assigned to approximately half of participating students.[1] Students must solve a problem and propose a course of action after analyzing a "document library" containing a mixture of trustworthy and unreliable information. Although critical thinking and writing skills may transcend academic disciplines,

[1] Other students are randomly assigned a 75-minute *Analytic Writing Task*, which entails writing a persuasive essay and critiquing the arguments of others. Performance Tasks are the focus of this analysis because, unlike Analytic Writing Tasks, they place students in real-world contexts, which is necessary for studying the interaction between fields of study and task contents.

authentic measurement requires some real-world context in which students demonstrate those skills. Table 1 describes six Performance Tasks and the content areas relevant to these tasks.

Table 1
CLA Performance Task descriptions

Task	Description	Content
1	Determine the cause of a widespread medical problem on a college campus.	Health, Natural Science, Social Science, Education
2	Determine the cause of an unusual deformity discovered in local wildlife.	Health, Natural Science
3	Make a decision with serious consequences for the residents of a city.	Business, Engineering
4	Classify writings and artwork as representative of different themes.	Humanities
5	Determine the cause of a recent accident involving a young student.	Health, Business, Social Science, Education
6	Make a decision about the relative effectiveness and value of extracurricular programs.	Health, Business, Education, Social Science

Performance Task scoring was carried out by a distributed network of calibrated human scorers. Responses to the Performance Tasks are evaluated on four criteria: *Analytic Reasoning and Evaluation* (identifying and interpreting relevant information, evaluating the credibility of information), *Problem Solving* (synthesizing information, making a decision, recognizing where matters are left uncertain), *Writing Effectiveness* (constructing an organized and cohesive essay with support for positions), and *Writing Mechanics* (demonstrating command of Standard Written English).

Subjects

The sample comprised 12,632 graduating seniors from 236 4-year institutions in the United States that participated in the CLA in either spring 2007 or 2008. The sample was 62% female and 74% White/non-Hispanic with an average age of 22.4 years. Most students (91%) reported English as the language spoken at home. Because of the need to control for entering

academic ability (SAT or ACT converted to the SAT score scale), sex, race, and language spoken at home, only students with those data were included. Based on their reported first majors, students were classified into seven fields of study (Table 2).

Table 2
Academic fields of studies and majors

Field of Study	Majors	N
Natural Sciences	Agriculture, Biological/Life Sciences, Physical Sciences	1652
Social Sciences	Anthropology, Economics, Ethnic/Cultural Studies, History, Law Enforcement, Multi/Interdisciplinary Studies, Political Science, Psychology, Sociology	2678
Humanities and Languages	Communications, English & Literature, Foreign Languages & Literature, Liberal/General Studies, Philosophy, Religion, Visual and Performing Arts.	2348
Business	Business, Public Administration	2629
Technology, Engineering, and Math	Architecture, Computer and Information Systems, Engineering & Technology, Mathematics	1160
Education	Education, Physical Education	1145
Health	Health-related Fields, Nursing & Physical Therapy	1020

Note: Students in unclassifiable majors were excluded (N=1060).

Analysis

Two-way ANOVA was employed to investigate the relationship between students' fields of study and CLA Performance Task scores before and after controlling for student characteristics. Both analyses included field of study, task, and their interaction as factors. The first analysis used CLA scores as the outcome, and the second analysis used CLA scores adjusted for entering academic ability, sex, race, and language spoken at home. These adjusted scores were obtained by regressing CLA scores on the four aforementioned variables and recording the residuals.

Results

The first ANOVA revealed a significant main effect for field of study, which indicated significant differences in average performance between fields of study (Table 3).[2] There was no significant interaction between field of study and task, meaning that students in each field of study did not perform significantly better or worse on tasks that that varied in their content. The factors included in this analysis accounted for 3.1% of the variance in CLA scores, with field of study alone accounting for 2.6%.

Table 3
Two-way ANOVA with CLA scale scores as the outcome

Source	SS	*df*	MS	*F*	*p*
Task (T)	990877	5	198115	5.6540	.000
Field of Study (FS)	11959000	6	1993167	56.8825	.000
T × FS	1094596	30	36487	1.0413	.4039

Figure 1 illustrates the significant main effect for field of study, with average scores for Natural Sciences, Technology, Engineering, and Math, Social Sciences, and Humanities and Languages clustered above Health, Business, and Education. The lack of interaction between field of study and task is revealed in Figure 1 by the lack of crossings of lines representing fields of study with very different overall averages. Table 4 shows which fields of study had significant average differences according to Tukey's Honestly Significant Difference post hoc test. For example, Health, Business, and Education scored significantly lower than the other fields of study, but not significantly different from one another. Students studying Natural Sciences scored significantly higher on average than all other fields of study.

[2] There was also a significant main effect for task, which was somewhat unexpected due to the scaling procedures employed by CAE to adjust for differences in the difficulty of tasks. Of course, with such large sample sizes, one would expect even small differences to be significant. The significance of this factor does not jeopardize conclusions or interpretations of the other factors, which are the main focus of this research.

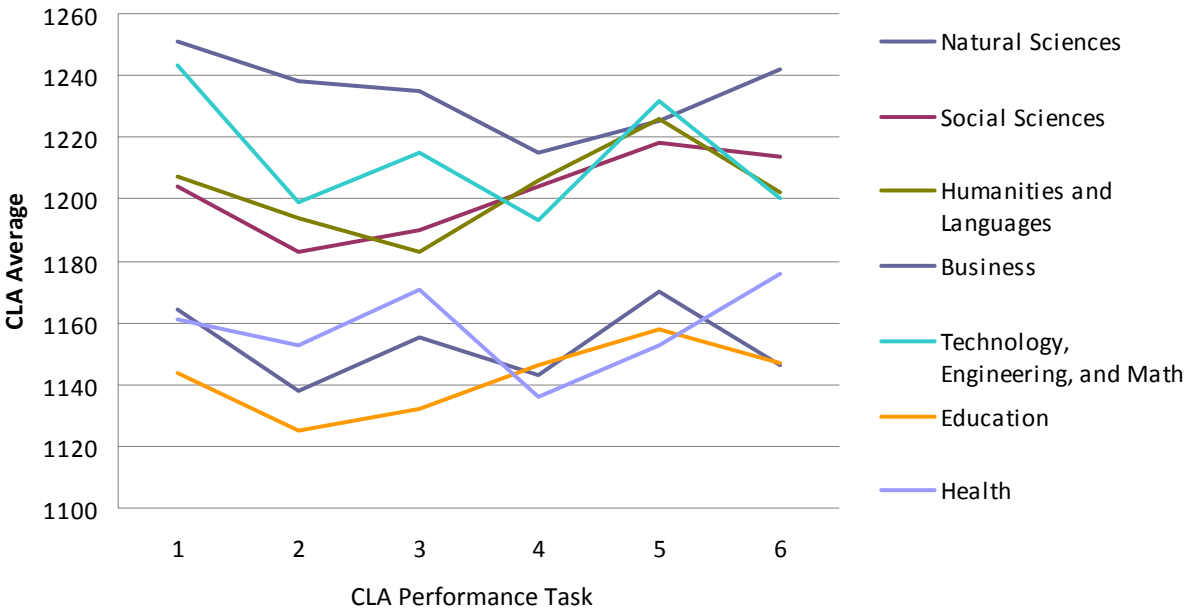

Figure 1. Average performance on 6 CLA Performance Tasks.

Table 4
Average CLA performance of different fields of study showing homogeneous subsets based on post hoc tests

	Subset		
Field of Study	1	2	3
Natural Sciences	1235		
Technology, Engineering, and Math		1213	
Social Sciences		1202	
Humanities and Languages		1202	
Health			1159
Business			1152
Education			1142

Next, ordinary least squares regression was conducted to adjust CLA scores for entering academic ability, sex, race, and language spoken at home. Entering academic ability alone accounted for 27.2% of the variance in CLA scores. Including the other three variables increased the variance accounted for by 0.3%. This small, statistically significant increase was accounted for by higher performance among female test takers and achievement gaps between racial/ethnic groups. Language spoken at home was not a significant predictor of CLA scores when the other

variables were also controlled for. The resulting raw residuals reflected student performance on the CLA relative to expectations based on entering academic ability, race, sex, and language spoken at home. These residuals were used as the outcome in the second ANOVA. The significance of factors was similar to the first ANOVA, with significant main effects for field of study and task and no significant interaction (Table 5). The factors included here accounted for only 1.03% of the variance in CLA scores. This reduction from the first ANOVA was not surprising given that choice of major is associated with entering academic ability and possibly other student characteristics.

Table 5
Two-way ANOVA with adjusted CLA scale scores as the outcome

Source	SS	*df*	MS	*F*	*p*
Task (T)	883633	5	176727	6.8153	.000
Field of Study (FS)	1493284	6	248881	9.5979	.000
T × FS	1017321	30	33911	1.3077	.121

The scores plotted in Figure 2 and listed in Table 6 reflect the average performance of fields of study relative to expectations based on entering academic ability, sex, race, and language spoken at home (positive scores indicate performance above expected). These values are indicative of the relative gains in critical thinking and writing skills that students attain during college. The main effect for field of study is less apparent in Figure 2 than it was in Figure 1, but it is still clear that some fields of study perform consistently better than others on this metric.

Differences between average adjusted scores in Table 6 may seem small, but considering that within-school standard deviations are about 175 and freshman-to-senior year longitudinal effect sizes are only around 0.50 (Arum, et al., 2011), some of these differences could be

considered quite large and indicative of substantial inequities in the acquisition of critical thinking and writing skills (e.g., the difference between Natural Sciences and Business reflects 0.16 standard deviations).

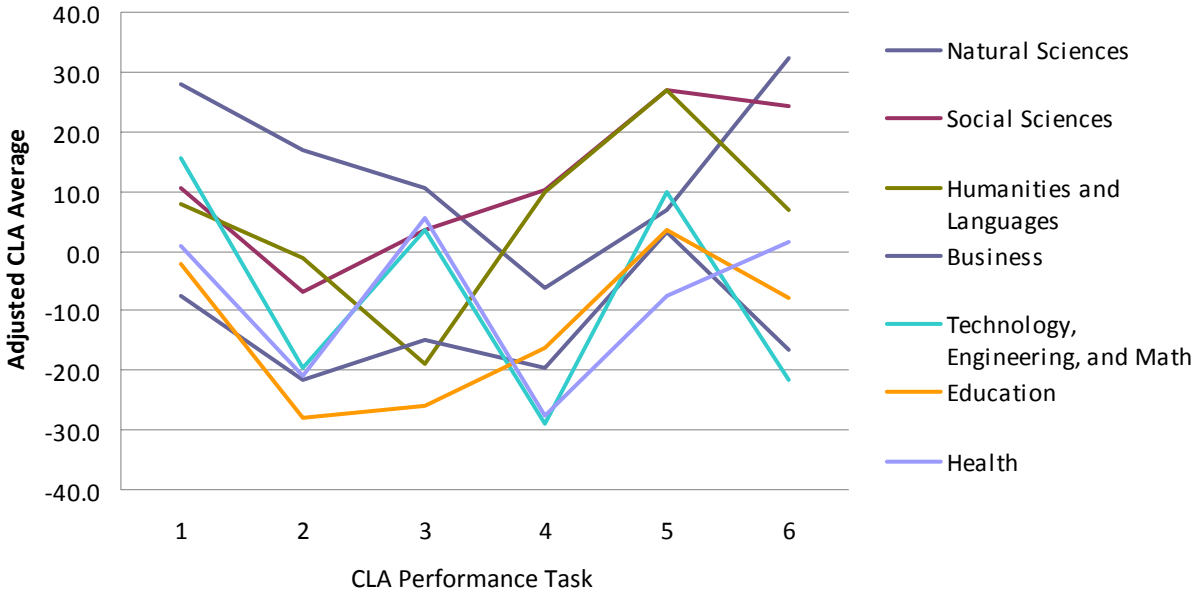

Figure 2. Average adjusted performance on 6 CLA Performance Tasks.

Table 6
Average adjusted CLA performance of different fields of study showing homogeneous subsets based on post hoc tests

	Subset		
Field of Study	1	2	3
Natural Sciences	14.6		
Social Sciences	11.0		
Humanities and Languages	4.8	4.8	
Technology, Engineering, and Math		-7.0	-7.0
Health		-7.6	-7.6
Education			-12.6
Business			-13.2

Natural Sciences, Social Sciences, and Humanities and Languages still topped the rankings after adjusting, and Business and Education were still ranked lowest. One notable change was that the standing of Health majors improved after adjusting since their average adjusted CLA score was no longer significantly different from Humanities and Languages and Technology, Engineering, and Math. This suggests that their low unadjusted average largely reflected student characteristics—most likely a lack of skills they brought to college (i.e., low SAT or ACT scores)—not a lack of learning during college. In contrast, Technology, Engineering, and Math had high unadjusted performance, but performance below expected after adjusting. That is, their high unadjusted average probably reflected the skills they brought to college, not those acquired during college.

Although there was no significant interaction between field of study and task, there is still some evidence that students perform better on tasks that include content related to their respective fields of study. For example, on Task 4 (a task that focuses heavily on humanities content), Humanities and Languages performed above its average, but Natural Sciences and Technology, Engineering, and Math performed below their respective averages. On task 5 (a task related to health, business, social sciences, and education), Business, Social Sciences, and Education performed above their respective averages, and Health performed near its average.

Discussion and Conclusions

This study investigated differential performance of fields of study on the CLA. In terms of overall performance, Natural Sciences, Technology, Engineering, and Math, Social Sciences, and Humanities and Languages performed better than Health, Education, and Business. CLA scores were then adjusted for student characteristics in order to reveal possible differences in the acquisition of critical thinking and writing skills during college. After adjusting, the order

remained fairly consistent except that Health improved in the rankings, and Technology, Engineering, and Math declined. Finding significant differences among the fields of study is consistent with previous research indicating that students in certain fields of study gain more critical thinking and writing skills during college (Arum & Roksa, 2011; Arum, et al., 2011).

Critics of standardized achievement testing in postsecondary education question whether skills like critical thinking and writing can be assessed apart from the domain in which those skills were learned. In neither analysis reported here was there a significant interaction between field of study and task. This finding suggests that critical thinking and writing skills can be assessed reasonably using complex, authentic performance assessments without great concern for the confounding effects of content knowledge on performance on specific tasks.

The results of this research are limited because they provide only descriptive, correlational information; they cannot reveal why fields of study perform differently on the CLA. Though the evidence presented here is consistent with the notion that students in certain fields of study gain more of the skills measured by the CLA, uncontrolled-for variables could further reduce these apparent differences (e.g., CLA pre-test scores). Future studies would benefit from additional information about differences between fields of study in terms of academic demands and expectations. Such variables may help explain significant differences between fields of study in senior performance on a standardized test of critical thinking and writing skills. Knowledge of the reasons for differences in general college outcomes could lead to important improvements in academic programs.

References

Arum, R., & Roksa, J. (2011). *Academically adrift: Limited learning on college campuses*. Chicago, IL: University of Chicago Press.

Arum, R., Roksa, J., & Cho, E. (2011). *Improving undergraduate learning: Findings and policy recommendations from the ssrc-CLA longitudinal project*. Brooklyn, NY: Social Science Research Council. Retrieved from http://highered.ssrc.org/files/SSRC_Report.pdf

Conner, R. W. (2011). Do majors matter. *Inside Higher Ed*. Retrieved from http://www.insidehighered.com/views/2011/06/16/connor_essay_on_why_majors_matter_in_how_much_college_students_learn

Hart Research Associates. (2009). *Learning and assessment: Trends in undergraduate education - a survey among members of the association of american colleges and universities*. Washington, DC: Hart Research Associates. Retrived from http://www.aacu.org/membership/documents/2009MemberSurvey_Part1.pdf

Shavelson, R. J. (2009). *Measuring college learning responsibly: Accountability in a new era*. Stanford, CA: Stanford University Press.

Sotherland, P. (2009). Cogitating aggravating disaggregating. *Asessment Notes*. Retrieved from http://www.liberalarts.wabash.edu/assessment-notes-cla-data/

Sotherland, P., Dueweke, A., Cunningham, K., & Grossman, B. (2007). Multiple drafts of a college's narrative. *Peer Review*, *9*(2), 20-23.