



## Measuring Learning Outcomes in Higher Education: Motivation Matters

### Executive Summary<sup>i</sup>

Ou Lydia Liu, Brent Bridgeman, & Rachel Adler

Educational Testing Service

**Background/Context:** With the pressing need for accountability in higher education, student learning outcomes assessments have been widely used to evaluate learning and inform policy. Researchers are making conclusions about the quality of U.S. higher education based on learning outcomes assessment data. For example, in the book *Academically Adrift* (Arum & Roksa, 2011), the authors found that 45% of the college students included in their sample did not make any progress on skills measured by the Collegiate Learning Assessment (CLA; Klein, Benjamin, Shavelson, & Bolus, 2007) during two years of college time. Although the outcomes assessment results may have significant implications for institutions, they usually bear no direct consequences for individual students. Therefore, student motivation in taking the low-stakes assessments is in question.

**Purposes:** This study addressed three research questions: What is the relationship between student motivation and test scores? Do motivational instructions that affect personal stakes in the outcomes affect student motivation and test performance? Do conclusions drawn about college learning gains change with test format (i.e., multiple-choice vs. essay) and motivational instructions?

**Intervention:** There were three motivational conditions: a control condition, an institutional condition, and a personal condition. In the control condition, students were told that their test scores would only be used for research purposes. In the

institutional condition, students were told that their test scores would be averaged with the scores of the other students at their institution in evaluating the instructional effectiveness of their institution, and that this might affect how their institution is viewed and therefore affect the value of their diploma. In the personal condition, the students were told that their test scores might be released to faculty at their institution or to potential employers to evaluate their academic ability.

**Research Design:** The motivational conditions were embedded in regular consent forms. Students were randomly assigned to the conditions within a testing session. Approval was obtained from the institutional review board and students were debriefed after the completion of the data collection.

**Sample:** Students (n=757) from three higher education institutions were recruited, including one research institution (n=340), one master's institution (n=299), and one community college (n=118).

**Instruments:** Students took the abbreviated version of the multiple-choice *ETS*® Proficiency Profile, an essay which is an optional part of the *ETS* Proficiency Profile, and a ten-item survey measuring self-reported motivation.

**Findings:** Motivational conditions had a statistically significant and substantial impact on student performance on both the *ETS* Proficiency Profile and the essay scores, and on students' self-report motivation (Table 1). The performance difference was as large as .68 standard deviations between students in the control and personal conditions. Most strikingly, after controlling for SAT® scores, the score differences between sophomores and seniors, namely the value-added learning that could be attributed to students' college education, varied dramatically from negative gain (-0.23 SD) to substantial gain (0.72 SD) across motivational conditions and tests (multiple-choice vs. essay: Figure 1). In addition, students' self-reported motivation was a significant predictor of both the *ETS* Proficiency Profile and essay scores, after

controlling for SAT scores for university students and placement test scores for community college students.

**Conclusions/Recommendations:** Limited college learning reported in prior research is likely an underestimate of student learning, if students received test instructions similar to the instruction used in our control condition. There are practical strategies that institutions can use effectively to enhance students’ test-taking motivation. Institutions that employ different motivational strategies should be compared with each other with great caution, especially when the comparison is for accountability purposes.

References:

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

Klein, S., Benjamin, R., Shavelson, R., & Bolus, R. (2007). The Collegiate Learning Assessment: Facts and fantasies. *Evaluation Review*, 31(5), 415–439.

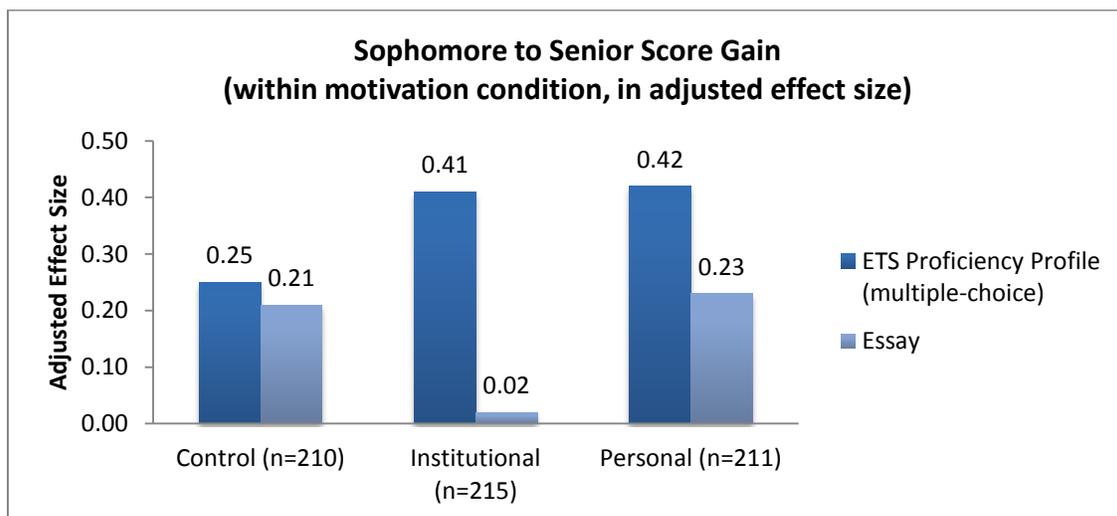


Figure 1: Sophomore to senior score gain (value-added) in effect size adjusted for SAT scores. Note: A general linear model analysis (GLM) was used to calculate the sophomore to senior score gain. In the GLM, the ETS Proficiency Profile and essay scores were used as separate outcomes variables, with motivational condition and class status being fixed factors, and SAT scores as a covariate. The adjusted effect size was the standardized mean difference between the estimated mean scores between sophomores and seniors in each condition.

Table 1

*Comparison by Motivational Condition and by Institution*

| <b>Self-Report Motivation Score</b> |         |      |     |             |      |     |          |      |     |          |          |          |       |       |
|-------------------------------------|---------|------|-----|-------------|------|-----|----------|------|-----|----------|----------|----------|-------|-------|
|                                     | Control |      |     | Institution |      |     | Personal |      |     | $d_{CI}$ | $d_{CP}$ | $d_{IP}$ | $F$   | $p$   |
|                                     | n       | Mean | SD  | n           | Mean | SD  | n        | Mean | SD  |          |          |          |       |       |
| RI                                  | 111     | 3.65 | .59 | 116         | 3.80 | .59 | 113      | 3.88 | .64 | .25      | .37*     | .13      | 4.43  | .010  |
| MI                                  | 99      | 3.59 | .60 | 99          | 3.76 | .60 | 98       | 3.88 | .61 | .28      | .48**    | .20      | 5.81  | .003  |
| CC                                  | 40      | 3.57 | .69 | 42          | 3.93 | .65 | 36       | 3.95 | .65 | .54*     | .57*     | .03      | 4.06  | .02   |
| Total                               | 250     | 3.61 | .63 | 257         | 3.81 | .60 | 247      | 3.89 | .63 | .31**    | .43***   | .14      | 13.68 | <.001 |

| <b>ETS Proficiency Profile Score</b> |         |      |       |             |      |       |          |      |       |          |          |          |       |       |
|--------------------------------------|---------|------|-------|-------------|------|-------|----------|------|-------|----------|----------|----------|-------|-------|
|                                      | Control |      |       | Institution |      |       | Personal |      |       | $d_{CI}$ | $d_{CP}$ | $d_{IP}$ | $F$   | $p$   |
|                                      | n       | Mean | SD    | n           | Mean | SD    | n        | Mean | SD    |          |          |          |       |       |
| RI                                   | 111     | 453  | 18.13 | 116         | 460  | 20.66 | 113      | 461  | 21.79 | .37*     | .40**    | .04      | 5.37  | .005  |
| MI                                   | 99      | 460  | 20.19 | 99          | 462  | 19.27 | 98       | 467  | 19.64 | .13      | .37*     | .25      | 3.5   | .032  |
| CC                                   | 40      | 435  | 20.74 | 42          | 443  | 18.48 | 36       | 450  | 21.08 | .37      | .68**    | .35      | 4.79  | .010  |
| Total                                | 250     | 453  | 21.11 | 257         | 458  | 20.84 | 247      | 462  | 21.62 | .26*     | .41***   | .16      | 11.19 | <.001 |

| <b>Essay Score</b> |         |      |      |             |      |     |          |      |      |          |          |          |       |       |
|--------------------|---------|------|------|-------------|------|-----|----------|------|------|----------|----------|----------|-------|-------|
|                    | Control |      |      | Institution |      |     | Personal |      |      | $d_{CI}$ | $d_{CP}$ | $d_{IP}$ | $F$   | $p$   |
|                    | n       | Mean | SD   | n           | Mean | SD  | n        | Mean | SD   |          |          |          |       |       |
| RI                 | 111     | 4.20 | .84  | 116         | 4.46 | .82 | 113      | 4.60 | .93  | .31      | .45*     | .16      | 6.24  | .002  |
| MI                 | 99      | 4.19 | .88  | 99          | 4.30 | .93 | 98       | 4.53 | .83  | .12      | .39*     | .26      | 3.73  | .025  |
| CC                 | 40      | 3.30 | 1.18 | 42          | 3.81 | .99 | 36       | 3.97 | 1.08 | .47      | .59*     | .15      | 4.04  | .020  |
| Total              | 250     | 4.07 | .96  | 257         | 4.29 | .93 | 247      | 4.46 | .95  | .23*     | .41***   | .18      | 12.93 | <.001 |

Note: RI= research university; MI= master's university; CC = community college.  
 $d_{CI}$  = standardized mean difference ( $d$ ) between the *Control* and *Institutional* conditions.  
 $d_{CP}$  = standardized mean difference ( $d$ ) between the *Control* and *Personal* conditions.  
 $d_{IP}$  = standardized mean difference ( $d$ ) between the *Institutional* and *Personal* conditions.  
 \*  $p < .05$ . \*\*  $p < .01$ . \*\*\*  $p < .001$ .

<sup>i</sup> This executive summary is a summary of the following paper:  
 Liu, O. L., Bridgeman, B., & Adler, R. (in press). Measuring learning outcomes in higher education: Motivation matters. *Educational Researcher*.