

## Early Validity Evidence for the Writing Assessment

### Use of the GRE Writing Assessment (GRE-W)

Many graduate faculty have decided that the skills and abilities required in the two types of GRE Writing Assessment tasks (“Present Your Perspective on an Issue” and “Analyze an Argument”) are important for success in their graduate courses, and they want to know the relationship of GRE-W scores to other indicators of success. Unfortunately, the best information will not be available until a fairly large number of students have taken the test and progressed through at least part of their graduate training. Only then will indicators of student success (such as grades, faculty rankings of students, and scores on papers, projects, and comprehensive examinations) exist.

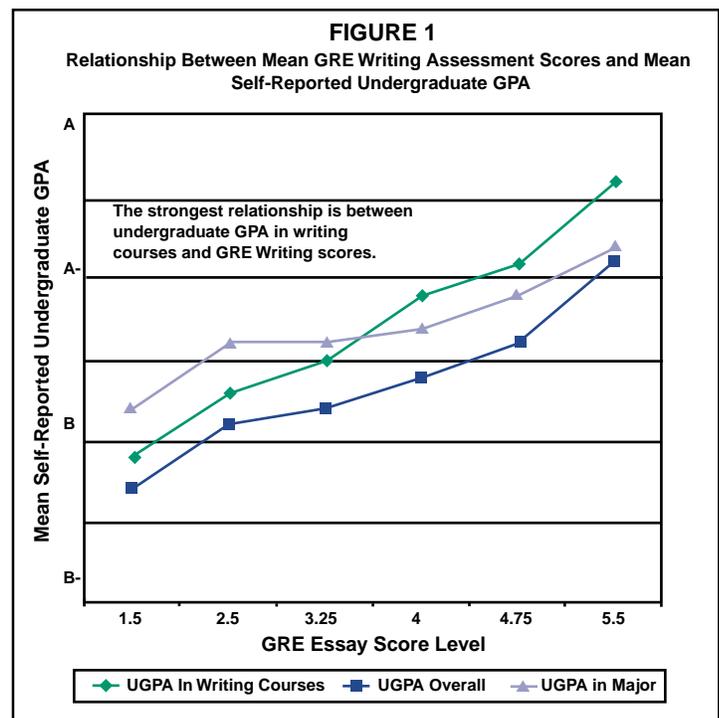
### A Look at the Data

Until these kinds of data are available, faculty can review ETS research data from a national study of more than 2,000 juniors, seniors, and first-year graduate students representing 26 geographically diverse colleges and universities.<sup>1</sup> In this study, the relationships between GRE-W scores, overall undergraduate GPA (UGPA), UGPA

in writing courses, and UGPA in the major field were analyzed. As shown in Figure 1, the GRE-W score has the strongest relationship with UGPA in writing courses. There are also positive relationships between the GRE-W score and overall UGPA and UGPA in the major field.

Figure 2 shows the relationship between GRE-W scores and course-related writing samples supplied by the test takers. One sample was identified by the student as a “typical” paper; the other was “not as well written.” These data show a positive relationship between the quality of the writing samples and the GRE-W score.

As part of the validation effort, two kinds of reliability were estimated for the GRE-W



<sup>1</sup> Powers, D.E., Fowles, M.E., and Welsh, C.K. (1999). *Further validation of a writing assessment for graduate admissions* (GRE Report 96-13). Princeton, NJ: ETS.

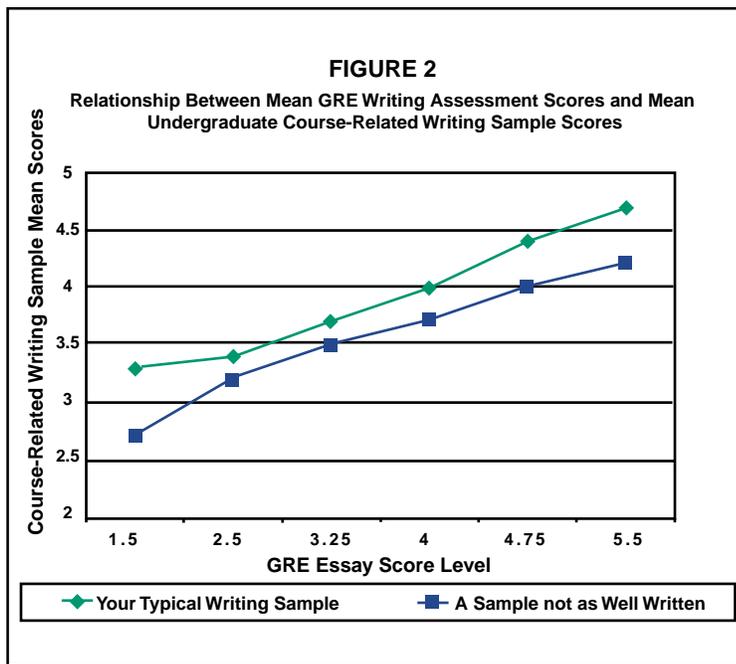
based on research data: one related to the topics (test reliability<sup>2</sup>) and one related to readers (interreader reliability<sup>3</sup>). The test reliability of the GRE-W is .72, a value that compares favorably with those of other free-response tests for which the reported scores are obtained by averaging the performance on two writing tasks.

Because a successful writing assessment depends on consistent judgments across individual essay readers, GRE researchers also examined interreader reliability. Figure 3 gives the percentages of reader agreement for each type of essay writing task. As Figure 3 shows, there were almost no score discrepancies of two or more points on the six-point score scale, very likely because of extensive reader training and monitoring.<sup>4</sup>

### Future Validity Research Plans

The information presented here gives some initial evidence related to the validity and

reliability of the GRE-W. As operational use of the GRE-W increases, the GRE Program will collect additional validity evidence, and we will be happy to collaborate with departments or programs wishing to conduct their own validity studies and to provide guidance about how best to approach this task. For more information about the GRE-W, send an e-mail to [NewGeneralTestInfo@ets.org](mailto:NewGeneralTestInfo@ets.org).



**FIGURE 3**

Agreement Between Two Readers of the Same Essay		
	Issue Task	Argument Task
Exact agreement	68.4%	68.1%
Agreement within 1 score point	29.9%	30.8%
Agreement within 2 score points	1.6%	1.0%
Agreement within 3 score points	0.1%	0.1%

<sup>2</sup> This is an internal consistency estimate of reliability.

<sup>3</sup> Interreader reliability is defined as the degree of agreement between two readers.

<sup>4</sup> See *An Introduction to the GRE Writing Assessment* (available from the GRE Web site at [www.gre.org](http://www.gre.org)) for copies of the scoring guides.

