

2011

The Assurance of Learning Tool as Predictor and Criterion in Business School Admissions Decisions: New Use for an Old Standard?

Bryan J. Pesta

Cleveland State University, b.pesta@csuohio.edu

Robert F. Scherer

Cleveland State University, r.scherer@csuohio.edu

Follow this and additional works at: https://engagedscholarship.csuohio.edu/bus_facpub

 Part of the [Business Administration, Management, and Operations Commons](#)

How does access to this work benefit you? Let us know!

Publisher's Statement

This is an Author's Accepted Manuscript of an article published in Journal of Education for Business, 2011, available online: <http://www.tandfonline.com/10.1080/08832323.2010.492051>.

Original Published Citation

Pesta, B. J., Scherer, R. F. (2011). "The Assurance of Learning Tool as Predictor and Criterion in Business School Admissions Decisions: New Use for an Old Standard?". Journal Of Education for Business, 86, pp. 163–170.

This Article is brought to you for free and open access by the Monte Ahuja College of Business at EngagedScholarship@CSU. It has been accepted for inclusion in Business Faculty Publications by an authorized administrator of EngagedScholarship@CSU. For more information, please contact library.es@csuohio.edu.

The Assurance of Learning Tool as Predictor and Criterion in Business School Admissions Decisions: New Use for an Old Standard?

Bryan J. Pesta and Robert F. Scherer

Cleveland State University, Cleveland, Ohio, USA

The Association to Advance Collegiate Schools of Business incorporates program assessment as an integral part of the accreditation process. Assessment tools created to meet assurance of learning standards, however, must go beyond grades and measure student learning directly. The author shows that an in-house assessment tool predicted student learning and correlated well with admissions criteria used to select students into an MBA program. Specifically, assessment exam scores from 182 MBA students correlated .47 with their final MBA grades. The assessment exam scores themselves were also well predicted by student GMAT scores and undergraduate grades. The results show that assurance of learning assessment tools can be useful for more than just accreditation decisions.

Keywords: AACSB, assurance of learning, business education, GMAT

As part of delivering an education, business schools must first decide which students to admit and then ultimately assess what those students have learned. Traditionally, admissions decisions have relied heavily on the predictive validity of undergraduate grades (UGPA) and scores on the Graduate Management Admissions Test (GMAT). The obvious criterion to be predicted has been the MBA student's final GPA. My purpose here is first to argue that grades are useful but deficient measures of student learning. Problems with grades—reviewed subsequently—suggest that business schools should develop alternate criteria with which to validate their admissions decisions.

The present literature on the GMAT's validity also supports this notion. Kuncel, Crede, and Thomas (2007) recently suggested that researchers "expand the criterion space" (p. 64) by going beyond grades as the outcome measure (see also Sireci & Talento-Miller, 2006). Likewise, accreditation standards from the Association to Advance Collegiate Schools of Business (AACSB; 2009) require that institutions develop assurance of learning assessment tools (other than class grades) that measure student learning directly. My contention is that program assessment tools could serve as ideal complements to GPAs. Combining grades with assessment scores creates

a composite outcome measure, which may help business schools as they validate tools used in admissions decisions.

The present article proceeds in two stages. First, I show that program assessments are valid outcome measures because they correlate well with final MBA GPAs (i.e., they possess convergent validity). Next, I show that the validity of traditional predictors used in business school admissions decisions (UGPA and GMAT scores) generalizes well when predicting performance on this new criterion. Framing my investigation in this way allows to also address an unresolved issue in the admissions decision literature; namely, whether the GMAT subtests (especially the Analytical writing Assessment [AWA]) possess incremental validity.

Problems With Grades as Outcome Measures

MBA grades are a reasonable and convenient outcome measure in research on the validity of the GMAT. Indeed, a large body of research has established that the combination of GMAT and UGPA scores predicts success in MBA programs (Groppe, 2007; Kuncel et al., (2007); Oh, Schmidt, Shaffer, & Le, 2008; Sireci & Talento-Miller, 2006; Talento-Miller & Rudner, 2008). With success almost always defined as MBA GPA (however, see Dobson, Krapljan-Barr, & Vielba, 1999), typically 25% or more of the variance in first-year MBA grades is explained by some linear combination of GMAT and UGPA scores (Graduate Management Admissions Council, 2009).

Correspondence should be addressed to Bryan J. Pesta, Cleveland State University, Department of Management, 2121 Euclid Avenue, BU-439, Cleveland, OH 44115, USA. E-mail: b.pest@csuohio.edu

Although useful, grades alone are deficient criteria for measuring success in an MBA program (unless student success is defined solely by graduation rates). This is especially true when grades are used as proxies for what students have learned (Pfeffer & Fong, 2002). A host of factors beside knowledge of business management likely contribute to an MBA student's GPA. Examples include motivation (Fulton & Turner, 2008; for a meta-analytic review, see Robbins et al., 2004), intelligence (Gottfredson, 2004; Pesta & Poznanski, 2009), personality (Duff, Boyle, Dunleavy, & Ferguson, 2002; Rothstein, Paunonen, Rush, & King, 1994), study habits (Plant, Ericson, Hill, & Asberg, 2005), and years of work experience (Adams & Hancock, 2000; Stinebrickner & Stinebrickner, 2003). Further, as outcome measures, grades suffer from both inflation and range restriction (Gottfredson).

Grades are not invalid per se, but they are determined by multiple factors. What business schools need are unidimensional measures of student learning to complement grades as indicators of student success. Program assessment tools, created to meet assurance of learning standards, could be ideal as criterion surrogates for grades.

Program Assessments as Alternative Outcome Measures

As of December 2009, the AACSB had accredited 560 institutions worldwide. All schools seeking accreditation must develop assessment tools that measure the effectiveness of their curriculum, as outlined in the AACSB, assurance of learning standards (AACSB, 2009). Schools are free to develop whatever assessments seem appropriate, but these must include direct measures of learning. Course grades are not program assessment measures (AACSB, 2009).

My college has worked to meet the Assurance of Learning standards by developing several assessment tools. These tools range from content-valid multiple-choice tests to rubrics that evaluate soft skills (e.g., communication ability), as evaluated by business presentations students make in class. My focus here is on the former: a content valid assessment exam created to measure management knowledge. To develop the exam, department faculty (i.e., subject matter experts) decided which course material was most critical and then wrote test items to cover that material. The items then underwent a series of revisions based on statistical analyses of student performance in pilot studies. The result was a well-developed measure of what students have learned about business management upon graduation from my program.

My first hypothesis therefore relates to the validity of the assessment exam as an outcome measure for MBA student learning. One way to establish the validity of a measure is to show that it correlates with other validated measures of the same construct (i.e., convergent validity; see Anastasi & Urbina, 1997). Hence, my first hypothesis is the following:

Hypothesis 1 (H₁): Scores on the program assessment would correlate significantly with MBA grades, suggesting that the assessment possesses convergent validity.

After demonstrating the validity of the program assessment as an outcome measure, I then shift the focus from using the assessment as a predictor, to using it as a criterion. The shift allows to expand the criterion space for testing the validity of the GMAT (Kuncel et al., (2007)). By operationalizing student performance both via MBA GPAs and assessment exam scores, I can perhaps achieve a richer picture of the GMAT's validity and utility. I can also address some of the unresolved issues in the literature on the validity of business school admissions decisions.

Loose Ends in the Validity Literature

Despite recent meta-analytic investigations of the GMAT's validity (Kuncel et al., (2007); Oh et al., 2008; Talento-Miller & Rudner, 2008), more data are needed on whether GMAT verbal and quantitative subtests possess incremental validity. Because of how students typically are selected into business programs (weak scores on one subtest can be offset by strong scores on the other), some studies actually show a negative relationship between verbal and quantitative scores, whereas others show a fair degree of multicollinearity (Talento-Miller & Rudner). Differences in whether the subtest scores correlate within studies then affects meta-analytic conclusions about incremental validity across studies. Here I test whether the GMAT subtests show incremental validity for both MBA grades and assessment examination performance.

Second, I further explore whether the newest component of the GMAT—the Analytical writing Assessment—adds anything to prediction accuracy over GMAT verbal and quantitative scores. Two recent studies (Talento-Miller, 2008; Talento-Miller & Rudner, 2008) showed mixed evidence of incremental validity for the writing subtest, and suggested that more data are needed on this issue. I provide these data, both with MBA grades and assessment exam performance as the criteria. The issue of incremental validity was tested via the following two hypotheses:

Hypothesis 2 (H₂): GMAT verbal and quantitative scores plus UGPA would show incremental validity for predicting scores on the program assessment (and final MBA GPAs).

Hypothesis 3 (H₃): GMAT writing scores would show incremental validity (over verbal and quantitative scores, plus undergraduate grades) for predicting scores on the program assessment (and final MBA GPAs).

In sum, I attempt to show that program assessment exams are valid as outcome measures for MBA student learning. Once establishing that assessment performance predicts MBA grades, I then use the assessment exam as a criterion to explore the validity of both GMAT scores

TABLE 1
Descriptive Statistics and Simple Correlations for the Study Variables

Variable	<i>M</i>	<i>SD</i>	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.
1. Sex	0.36	0.48	—										
2. Age	27.14	5.08	-.09	—									
3. BBA	0.57	0.50	-.11	-.20	—								
4. International	0.81	0.40	.07	.01	.05	—							
5. UGPA	3.13	0.42	.13	-.05	.03	.12	—						
6. GMAT total	508.6	82.7	-.17	-.03	-.15	-.13	.04	—					
7. Verbal %	48.9	22.5	-.11	.06	-.05	.35	.14	.64	—				
8. Quantitative %	37.2	23.2	-.14	-.07	-.18	-.48	-.08	.77	.12	—			
9. Writing %	47.8	26.0	.11	-.13	-.13	.30	.11	.17	.34	-.06	—		
10. MBA GPA	3.54	0.24	-.07	.19	-.10	.02	.42	.43	.36	.27	.13	—	
11. MBA exam %	58.6	10.2	-.12	.20	-.02	.24	.20	.47	.55	.15	.26	.47	—

Note. Sex was coded 0 for males and 1 for females. BBA was coded 0 for nonbusiness majors and 1 for business majors. International was coded 0 for nonnative U.S. students and 1 for U.S. students. UGPA = undergraduate grade point average. All correlations $> .14$ were significant ($p < .05$).

and UGPA. Doing so allows me to also contribute to the graduate business school literature on admissions decisions, as I provide a novel criterion—scores on the program assessment—and focus analyses on several unresolved issues in the literature. Finally, to replicate prior studies, and to serve as a basis of comparison for data on the assessment exam, I also present analyses using final MBA grades as a criterion.

METHOD

Participants

Participants were 182 MBA students admitted to my program between the summer 2003 and fall 2005 semesters. All students graduated by the summer 2007 semester. Table 1 shows the demographic characteristics of the sample. Many of the variables in the table were dichotomized for use in regression analysis. For reference, the sample comprised 117 (64.3%) men and 65 (35.7%) women, with a mean age of 27.14 years ($SD =$, range = 21–45 years). My business college has a reasonable mix of students continuing directly from their undergraduate education, and of those returning to school after working for some time (participants 30 years of age or older represented 26.9% of the sample). Most of the students (56.6%) had undergraduate majors in business administration, and most were U.S. natives (80.8%).

Also noteworthy is the wide range of GMAT scores and UGPA represented in the sample. Verbal scores ranged from 6% to 94%, spanning almost the entire range of possible scores. Likewise, quantitative scores spanned the entire range, from 1% to 99%. Undergraduate GPAs in the sample were as low as 2.00 and as high as 4.00. Hence, the validity coefficients reported here are not as range-restricted as those seen in the broader literature (e.g., Kuncel et al., 2007).

Materials and Procedure

The following variables in Table 1 were coded from student transcripts: (a) sex (coded 0 for male and 1 for female), (b) age in years, (c) undergraduate major (coded 0 for nonbusiness majors and 1 for business majors), (d) international status (coded 0 for nonnative students and 1 for U.S. natives), (e) UGPA, and (f) final MBA GPA. I coded GMAT total, quantitative, verbal, and writing scores from each student's application form.

The assessment examination was developed in house, for use in program evaluation consistent with the AACSB's assurance of learning standards. My college designed the exam to be a content-valid measure of what students should know upon graduating from an MBA program. Faculty from each academic department determined which content was most important, and wrote test items to cover that content. This resulted in an 81-item, multiple-choice exam, covering all functional areas of business (management, finance, accounting, marketing, operations management, and information science).

Students completed the exam during a class period of a capstone MBA course, which is the last class required before graduation. They had 90 min to complete the exam and received extra credit for participation. All students present in class on administration days completed the exam. To motivate students, the extra credit was scaled such that higher scores received more points. Across administrations, the exam produced internal consistency reliabilities in the upper .70s. The exam also produced a wide range of scores (from 22% to 79%), suggesting substantial differences in learning across graduates.

Analytical Approach

I tested Hypotheses 1–3 via simple correlations, and then by using hierarchical linear regression. The former analyses

establish criterion validity whereas the latter establish incremental validity.

RESULTS

Table 1 shows descriptive statistics and simple correlations for the study variables. I used $p < .05$ as the level of significance for all statistical tests. Looking first at MBA grades as the criterion, the pattern of correlations replicates Kuncel et al.'s (2007) meta-analysis. Both UGPA ($r = .42$) and GMAT total scores ($r = .43$) significantly predicted MBA GPAs. Subtest scores for the GMAT had nominally lower criterion validity, with verbal scores ($r = .36$) predicting better than did quantitative scores ($r = .27$). The writing scores alone did not significantly predict MBA grades ($r = .13$).

Assessment exam scores also correlated with MBA grades ($r = .47$), clearly supporting H_1 . As a validity coefficient, the .47 value is about as large as those seen in the human resources literature for the most effective selection methods predicting job performance (e.g., Anderson & Hulsheger, 2008; Schmidt & Hunter, 1998, 2004). Program assessments therefore possess convergent validity with MBA grades, suggesting that they may compliment grades as outcome measures in research on the validity of admissions decisions.

I next turn from using the program assessment as a predictor of student outcomes (i.e., grades), to using it as a criterion for validating GMAT scores and UGPA. From Table 1, all of the traditional variables used in admissions decisions significantly predicted assessment exam scores as an outcome measure. The validity coefficients, however, varied considerably in size. The best predictors were GMAT verbal ($r = .55$) and GMAT total scores ($r = .47$). GMAT writing scores also predicted exam performance ($r = .26$), and the weakest predictors were UGPA ($r = .20$), and GMAT quantitative scores ($r = .15$). The simple correlations show that program assessment tools are viable criteria for research aimed at validating business school admissions decisions. They also support the validity of the GMAT and undergraduate grades as predictors of outcome variables other than MBA grades.

Of further interest in Table 1 is that (a) among the demographic variables, only age significantly predicted both MBA grades ($r = .19$), and assessment exam scores ($r = .20$); and (b) nonnative students (coded 0) scored higher than U.S. natives (coded 1) on the quantitative subtest ($r = -.48$), but they scored lower on both the verbal ($r = .35$) and writing ($r = .30$) subtests. The latter result is perhaps not surprising, as English is a second language for my international students (Talento-Miller [2008] reported data on the GMAT's validity for students whose native language is not English). The mix of nonnative and U.S. students in the same sample may also explain why the verbal and quantitative scores themselves were not significantly correlated within my sample ($r = .12$; but, after controlling for native

status, the correlation increased to $r = .35$, which is similar to that reported by Talento-Miller & Rudner [2008]).

The advantage to nonnative students on the quantitative subtest could be an artifact. Nonnative students with low verbal scores can only be accepted into the program when those scores are offset by correspondingly high quantitative scores (i.e., students scoring low on both the verbal and quantitative subtests would not meet my admission criteria, and so are not represented in this sample). I dealt with this here by controlling for student status in the regression analyses presented subsequently.

Controlling for student status also provides a fairer test of the incremental validity of the GMAT subtests. The control allows the verbal and quantitative subtests to be correlated, which they should be in a nonbiased sample of participants, given the positive manifold (i.e., the ubiquitous finding that scores on a variety of mental tests are all positively correlated; e.g., Jensen, 1998; Johnson, te Nijenhuis, & Bouchard, 2008).

It is also possible that the mix of U.S. native and nonnative students in my sample attenuated the true validity of the GMAT writing scores (and verbal scores, as they correlated significantly with writing scores, $r = .34$). To test this possibility, I conducted an additional analysis where student status (U.S. native vs. nonnative) was partialled out from all simple correlations in Table 1. Nonetheless, controlling for student status did little to change the validity coefficients for either the writing scores ($r = .13$ and $.20$, for MBA grades and assessment scores, respectively) or the verbal scores ($r = .37$ and $.52$, respectively). The predictive validity of the quantitative scores, however, was improved after controlling for student status ($r = .32$ and $.31$, respectively). At any rate, the student status data have important implications for programs with a large international student base. I return to this issue in the discussion.

The Table 1 data are only simple correlations, which establish criterion but not incremental validity. At a practical level, incremental validity is most important. It would be unwise for an admissions committee to spend time weighting selection criteria that are redundant with those they already have. Tests of incremental validity are therefore presented in Table 2. Here, I conducted a series of hierarchical regressions using MBA grades and then assessment exam scores as the criterion. For each regression, I entered age, student status (U.S. native vs. nonnative), and UGPA in Step 1. In Step 2, I entered GMAT verbal and quantitative scores, followed by GMAT writing scores in Step 3.

I entered the writing scores separately at Step 3 because they are a relatively new addition to the GMAT. Less research has been done on the validity of the writing scale, and so I thought a specific test of its incremental validity would be informative. Note that no conclusions would change by instead entering the writing scores together with the other variables at Step 2 (however, information on the before and after validity of the writing scale would be lost).

TABLE 2
Regressions Testing the Incremental Validity of GMAT Scores for MBA Grades and MBA Assessment Exam Scores

Step	MBA grades			MBA exam		
	<i>B</i>	<i>SE</i>	β	<i>B</i>	<i>SE</i>	β
Step 1						
Age	0.010	.003	.213*	0.334	.114	.206*
International status	-0.021	.040	-.035	4.54	1.47	.217*
Undergraduate GPA	0.248	.038	.440*	3.69	1.39	.187*
Step 2						
Age	0.010	.003	.216*	0.307	.097	.189*
International Status	0.018	.046	.030	3.38	1.60	.162*
Undergraduate GPA	0.239	.034	.424*	2.89	1.19	.147*
GMAT Verbal	0.002	.001	.235*	0.161	.025	.439*
GMAT Quantitative	0.003	.001	.305*	0.070	.026	.198*
Step 3						
Age	0.010	.003	.224*	0.331	.098	.204*
International Status	0.012	.046	.019	2.95	1.62	.141
Undergraduate GPA	0.238	.034	.421*	2.81	1.19	.142*
GMAT Verbal	0.002	.001	.222*	0.151	.026	.412*
GMAT Quantitative	0.003	.001	.304*	0.070	.026	.198*
GMAT Writing	0.000	.001	.050	0.032	.021	.102

Note. The overall R^2 values were .39 for both MBA grades and MBA exam scores. MBA grades: $R^2 = .23$ for Step 1; $\Delta R^2 = .16$ ($p < .05$) for Step 2; $\Delta R^2 = .00$ (ns) for Step 3. MBA exam: $R^2 = .13$ for Step 1; $\Delta R^2 = .25$ ($p < .05$) for Step 2; $\Delta R^2 = .01$ (ns) for Step 3.
* $p < .05$.

Looking first at MBA grades as the criterion to replicate prior research, both age ($\beta = .213$) and UGPA ($\beta = .440$) were significant at Step 1. Being older was associated with higher grades in my sample, and students with higher UGPAs tended to have higher MBA GPAs. At Step 2, both GMAT verbal ($\beta = .235$) and GMAT quantitative ($\beta = .305$) scores explained unique variance (i.e., were incrementally valid) in MBA grades. Note also that age ($\beta = .216$) and UGPA ($\beta = .424$) remained significant as predictors here. Steps 1 and 2 support H_2 , and replicate results from prior studies, including the recent meta-analysis by Kuncel et al. (2007; see also Talento-Miller & Rudner, 2008). The combination of GMAT scores and UGPA (plus age) in this sample explained 39% of the variance in MBA grades.

Step 3, however, shows that the GMAT writing scores ($\beta = .050$) added no incremental validity to the prediction of MBA student grades. The percentage of variance explained at Step 3 (39%) was the same as that for Step 2. Alternatively, all the predictors that were significant at Step 2 remained significant at Step 3: age ($\beta = .224$), UGPA ($\beta = .421$), verbal scores ($\beta = .222$), and quantitative scores ($\beta = .304$). H_3 was therefore not supported with MBA grades as the criterion.

Table 2 also presents the results of the regression analysis with the program assessment exam as the criterion variable. All three predictors were significant at Step 1. First, increased age was associated with better exam performance ($\beta = .206$); nonnative students (coded 0) scored lower than did native U.S. students (coded 1; $\beta = .217$), and students with higher

UGPAs tended to score higher on the MBA assessment exam ($\beta = .187$). At Step 2, all predictors were again significant. Both GMAT verbal ($\beta = .439$) and GMAT quantitative ($\beta = .198$) scores explained incremental variance in the MBA exam, even when controlling for age, student status, and UGPA (thus supporting H_2). The percentage of variance explained by the Step 2 variables was 38%.

At Step 3, GMAT writing scores failed to predict unique variance in assessment exam scores ($\beta = .102$). The variance explained at Step 3 was 39%, representing only a one percentage point increase relative to Step 2. Conversely, UGPA ($\beta = .142$), verbal scores ($\beta = .412$), quantitative scores ($\beta = .198$), and age ($\beta = .204$) all remained significant at Step 3. H_3 was therefore not supported with either MBA grades or assessment exam scores as the criterion.

In the Kuncel et al. (2007) meta-analysis, GMAT total scores ($\beta = .356$) predicted grades better than did either the verbal ($\beta = .288$) or quantitative scores ($\beta = .278$) alone. I tested whether this effect existed here, given multicollinearity between the verbal and quantitative subtest scores. In my sample, verbal and quantitative scores correlated .35, once correcting for student status (the regressions in Table 2 also controlled for student status). I therefore reran the analyses in Table 2, replacing the verbal and quantitative subtest scores with GMAT total scores in all steps. Although the conclusions remained the same, the effects for the GMAT were now stronger. With MBA grades, GMAT total scores produced a Step 3 β of .408, which was larger than that seen for either the verbal or quantitative subtest scores alone (.222

and .304, respectively, from Table 2). Similarly, with the program assessment exam, GMAT total scores produced a β of .485, again larger than that seen with either the verbal or quantitative scores alone (.412 and .198, respectively). In neither analysis did the GMAT writing scores add to the prediction.

Because the sample contained only 35 nonnative U.S. students, I did not have the statistical power to test for interactions between the GMAT subtest scores and student status as predictors of either MBA grades or program assessment scores. However, given the relatively strong differences in subtest scores by student status in Table 1, I opted to run two additional analyses looking only at the 147 U.S. students in my sample. The first analysis predicted MBA grades, and the second predicted program assessment scores. For MBA grades, the percentage of variance explained was 44%, with age ($\beta = .178$), UGPA ($\beta = .496$), GMAT verbal ($\beta = .191$), and GMAT quantitative ($\beta = .389$) scores all emerging as significant. Once again, the writing scores ($\beta = -.034$) were not incrementally valid as predictors of U.S. student grades.

The same pattern occurred with the assessment exam as the criterion. Here, 34% of the variance was explained, with significant predictors including age ($\beta = .201$), UGPA ($\beta = .156$), GMAT verbal ($\beta = .355$), and GMAT quantitative ($\beta = .242$) scores. Writing scores, themselves, were not reliable predictors of assessment exam performance ($\beta = .072$). In sum, across several analyses, GMAT writing scores appear to possess no incremental validity for either MBA grades or MBA program assessment scores.

DISCUSSION

Overview of Key Findings

I derived and tested three hypotheses. The first focused on whether assessment tools created to meet AACSB assurance of learning standards could predict the traditional measure of student success—final MBA grades (and thereby serve as outcome measures themselves in research aimed at validating business school admissions decisions). The second and third looked at whether the traditional variables used in validation research would also predict student success when the program assessment exam—instead of MBA grades—was the outcome measure. I framed these hypotheses as tests of incremental validity for both the GMAT subtests and UGPA predicting assessment examination scores.

H_1 was clearly supported. My college developed a content valid measure of management knowledge as a tool for program assessment to meet AACSB assurance of learning standards. Here, the assessment exam showed strong validity for predicting MBA grades (relative to validity coefficients for various selection methods predicting job performance in the human resources literature). Hence, assessment exams

can complement MBA grades as measures of student learning in an MBA program.

H_2 also received strong support. The GMAT verbal and quantitative subtests, plus UGPA, showed significant incremental validity for predicting the outcome measures. Importantly, the GMAT's validity for predicting student success beyond GPA was demonstrated. Both GMAT verbal and quantitative scores predicted performance on the program assessment exam about as well as they did MBA grades. In sum, the variables I use in admissions decisions—especially combined—possess substantial validity as predictors of both the assessment scores and MBA GPAs.

For H_3 , I addressed whether GMAT writing scores possess incremental validity, as this issue remains open in the literature (Talento-Miller & Rudner, 2008). Here, GMAT writing scores showed no evidence of incremental validity. Only the simple correlation between writing scores and the program assessment was significant, but the correlation was relatively weak (i.e., $r = .26$). More importantly, across a series of regression analyses, the writing scores failed to explain anything in the way of unique variance in either MBA grades or program assessment scores.

Implications for Management Education

The first implication concerns going beyond grades as the outcome measure for students in an MBA program. As argued previously, grades are useful but deficient measures of what students have learned. Even the link between MBA grades and success in the business world is dubious (Pfeffer & Fong, 2002). The problem is perhaps that grades are multidimensional in nature, as a host of factors influence one's GPA. Assessment tools, however, seem like unidimensional assessments that measure student learning directly. The assessment examination used here was administered to students at the end of the program, in a standardized environment, and with incentives (i.e., extra credit) for scoring well. As such, it served the dual purpose of allowing my college to objectively assess student learning and to validate admissions decisions against an outcome variable other than grades.

If business schools expend effort into program assessment for purposes of AACSB accreditation, then it seems rational to also use this information as a tool for validating existing predictors used in admissions decisions. Showing that factors considered in admissions decisions predict grades is one thing; showing that they also predict objective, well-designed measures of program assessment would further strengthen the decision-making process, and the perception that the process is fair, and job-related.

Issues such as grade inflation or protected class differences in grades could also be assessed by using alternative measures of student learning (e.g., assessment examinations). For example, if minority students have lower GPAs, it would be incumbent upon the school to see if differences also exist on an independent and objective measure of student

learning. Indeed, showing that grades predict assessment scores equally well for minority and nonminority students would support a legal inference that the grading system is fair for all students (Griggs v. Duke Power, 1971).

The present study also helps resolve the mixed literature on the validity of the GMAT writing scale (Sireci & Talento-Miller, 2006; Talento-Miller & Rudner, 2008). Given the present data, writing scores seem to add little useful information over and above that offered by verbal and quantitative scores. I therefore see little value in using these scores in admissions decisions. It can be argued that writing scores may offer more useful information for students who are nonnative English speakers (Talento-Miller, 2008). A counterargument is that well-developed standardized exams already exist that test this issue (i.e., the Test of English as a Foreign Language [TOEFL]). It seems unlikely that GMAT writing scores would add incremental validity over the TOEFL, as the former could not even add incremental validity over the verbal scores here.

Limitations and Directions for Future Research

One limitation of the present study was that scores on the GMAT subtests varied by student status. Nonnative students scored higher on the quantitative subtest, but U.S. natives scored higher on both the verbal and writing subtests. I did not have enough nonnative students to test for differential validity (Dobson et al., 1999; Koys, 2005). However, I did statistically control for student status in the regressions, and presented an analysis of U.S. natives only, which mirrored the results I found with the full sample.

Interestingly, at least with my sample, student status differences did not appear for GMAT total scores, as the U.S. native advantage on the verbal scores was washed out by the nonnative advantage on the quantitative scores. To the extent that other business schools experience subtest differences for U.S. native versus nonnative students, use of GMAT total scores is recommended, (versus verbal and quantitative scores, separately—for a counterargument, see Talento-Miller & Rudner, 2008). Low scores on either verbal or quantitative could be used to direct students to remedial classes, but the fairest admissions decisions could perhaps be reached by looking at GMAT total scores, especially for students whose native language is not English.

A second limitation concerned my focus on only one type of assessment—a content-valid test of student knowledge. Many other important dimensions (e.g., communication and leadership ability, teamwork) exist with regard to success as a manager. Future researchers should assess whether these domains also possess the validity seen here with an objective measure of student learning.

A third limitation concerned the generalizability of my results to other business schools. My college exists in a large urban environment, and my students are diverse in terms of age, ethnicity, and gender. I suspect my results would

generalize well to similarly situated schools. Moreover, the global validation of the GMAT has been established many times. Kuncel et al. (2007) recommended that researchers conduct local validation studies, and that the criterion space be expanded to include outcome measures other than grades. The present research is hopefully a step in that direction.

As a direction for future research, I invite other business schools to validate their program assessment tools against MBA grades, and then use the assessments themselves as supplemental criteria for evaluating the validity of their admissions decisions. Content-valid tests have been used in employee selection for decades (e.g., Palumbo, Miller, Shalin, & Steele-Johnson, 2005). In graduate business education, they could serve the joint purpose of program assessment for accreditation purposes, and as an additional criterion to be predicted by variables the school uses in admissions decisions.

In sum, tools used to secure and maintain accreditation presumably measure student learning directly. Here I demonstrated that these tools possess convergent validity by predicting MBA final GPAs. They also serve as useful criterion variables—by going beyond grades—for validating admissions decisions. Since many schools already possess assessments like the one featured here, there seems to be little downside to adopting these tools as criteria to be predicted when making admissions decisions.

REFERENCES

- Adams, A., & Hancock, T. (2000). Work experience as a predictor of MBA performance. *College Student Journal, 34*, 211–219.
- Anastasi, A., & Urbina, S. (1997). *Psychological testing*. Upper Saddle River, NJ: Prentice-Hall.
- Anderson, N., & Hulsheger, U. (2008). *Employee selection and performance management*. London, UK: Sage.
- Association to Advance Collegiate Schools of Business. (2009). *Assurance of learning standards*. Retrieved from <http://www.aacsb.edu/accreditation/accreditedmembers.asp>
- Dobson, P., Krapljan-Barr, P., & Vielba, C. (1999). An evaluation of the validity and fairness of the Graduate Management Admission Test used for MBA selection in a UK business school. *International Journal of Selection and Assessment, 7*, 196–202.
- Duff, A., Boyle, E., Dunleavy, K., & Ferguson, J. (2002). The relationship between personality, approach to learning and academic performance. *Personality and Individual Differences, 36*, 1907–1920.
- Fulton, E., & Turner, L. (2008). Students' academic motivation: Relations with parental warmth, autonomy granting, and supervision. *Educational Psychology, 28*, 521–534.
- Gottfredson, L. (2004). Schools and the g factor. *Wilson Quarterly, Summer*, 35–45.
- Graduate Management Admissions Council. (2009). *Profile of Graduate Management Admissions Test candidates*. Retrieved from http://www.gmac.com/NR/rdonlyres/2008_GMAC_PrintedProfile.pdf
- Griggs v. Duke Power, 401 U.S. 424 (1971).
- Gropper, D. (2007). Does the GMAT matter for executive MBA students: Some empirical evidence. *Academy of Management Learning and Education, 6*, 206–216.

- Jensen, A. (1998). *The g factor: The science of mental ability*. Westport, CT: Praeger.
- Johnson, W., te Nijenhuis, J., & Bouchard, T. (2008). Still just 1 g: Consistent results from five test batteries. *Intelligence, 36*, 81–95.
- Koys, D. (2005). The validity of the Graduate Management Admission Test for non-U.S. students. *Journal of Education for Business, 80*, 236–239.
- Kuncel, N., Crede, M., & Thomas, L. (2007). A meta-analysis of the predictive validity of the Graduate Management Admission Test (GMAT) and undergraduate GPA (UGPA) for graduate student academic performance. *Academy of Management Learning & Education, 6*, 51–68.
- Oh, I., Schmidt, F., Shaffer, J., & Le, H. (2008). The Graduate Management Admissions Test (GMAT) is even more valid than I thought: A new development in meta-analysis and its implications for the validity of the GMAT. *Academy of Management Learning & Education, 7*, 563–570.
- Palumbo, M., Miller, C., Shalin, V., & Steele-Johnson, D. (2005). The impact of job knowledge in the cognitive ability—performance relationship. *Applied H.R.M. Research, 10*, 13–20.
- Pesta, B., & Poznanski, P. (2009). The inspection time and over-claiming tasks as predictors of MBA student performance. *Personality and Individual Differences, 46*, 236–240.
- Pfeffer, J., & Fong, C. (2002). The end of business schools? Less success than meets the eye. *Academy of Management Learning & Education, 1*, 78–95.
- Plant, E., Ericson, K., Hill, L., & Asberg, K. (2005). Why study time does not predict GPA across college students: Implications of deliberate practice for academic performance. *Contemporary Educational Psychology, 30*, 96–116.
- Robbins, S., Lauver, K., Le, H., Davis, D., Longley, R., & Carlstrom, A. (2004). Do psychosocial and study skill factors predict college outcomes? A meta-analysis. *Psychological Bulletin, 130*, 261–288.
- Rothstein, M., Paunonen, S., Rush, J., & King, G. (1994). Personality and cognitive ability as predictors of performance in graduate business school. *Journal of Educational Psychology, 86*, 516–530.
- Schmidt, F., & Hunter, J. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin, 124*, 262–275.
- Schmidt, F., & Hunter, J. (2004). General mental ability in the world of work. *Journal of Personality and Social Psychology, 86*, 162–174.
- Sireci, S., & Talento-Miller, E. (2006). Evaluating the predictive validity of the Graduate Management Admission Test Scores. *Educational and Psychological Measurement, 66*, 305–317.
- Stinebrickner, R., & Stinebrickner, T. (2003). Working during school and academic performance. *Journal of Labor Economics, 21*, 473–491.
- Talento-Miller, E. (2008). Generalizability of GMAT validity to programs outside the U.S. *S, 8*, 127–142.
- Talento-Miller, E., & Rudner, L. (2008). The validity of the Graduate Management Admission Test scores: A summary of studies conducted from 1997 to 2004. *Educational and Psychological Measurement, 68*, 129–138.