



# MAT Reliability and Validity

PEARSON

**Pearson Executive Office 5601 Green Valley Drive Bloomington, MN 55437**

Copyright © 2012 NCS Pearson, Inc. All rights reserved.

Normative data copyright © 2012 NCS Pearson, Inc.

**Warning:** No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopy, recording, or any information storage and retrieval system, without permission in writing from the copyright owner.

**Pearson, Miller Analogies Test, MAT,** and the **MAT logo** are trademarks in the U.S. and/or other countries of Pearson Education, Inc., or its affiliate(s). Portions of this work were previously published.

Produced in the United States of America.

**PEARSON**

For orders and inquiries:  
MAT Customer Relations, 19500 Bulverde Road, San Antonio, TX 78259  
800.622.3231  
[www.MillerAnalogies.com](http://www.MillerAnalogies.com)

# Table of Contents

<b>Acknowledgments .....</b>	<b>iv</b>
<b>Introduction.....</b>	<b>1</b>
<b>MAT Reliability .....</b>	<b>2</b>
Internal Consistency.....	2
Standard Error of Measurement ( <i>SEM</i> ) .....	3
<b>MAT Validity .....</b>	<b>4</b>
Construct Validity—the MAT Analogy .....	4
Content Validity.....	5
Predictive Validity .....	5
<b>Glossary .....</b>	<b>9</b>
<b>References.....</b>	<b>11</b>
<b>Tables</b>	
1. Reliability Coefficients (KR-20) and Standard Errors of Measurement ( <i>SEM</i> ) for MAT Scaled Scores by Total Group and Intended Field of Study .....	2
2. Correlations ( <i>r</i> ) Between Predictor Variables and 1st Year Graduate GPA from the 1992 Psychological Corporation Study and the 2008 Pearson Study .....	7

# Acknowledgments

This document was prepared by the following members of the Pearson Assessment and Information, Clinical Content Development, Post-Secondary Education group:

Donald G. Meagher, EdD  
Senior Research Director

Tianshu Pan, PhD  
Psychometrician

Rachel Wegner, MA  
Test Development Manager

Jeffrey R. Miller, MA  
Editor

# Introduction

The *Miller Analogies Test* (MAT) is a norm-referenced standardized test that is developed to measure the analytical skills and academic content critical to success in graduate education. In an effort to ensure the continuing relevance and usefulness of the MAT for assessing the prerequisite knowledge and skills of candidates for admission to graduate programs, the test content is periodically reviewed and revised to ensure that it accurately reflects current prerequisite verbal, analytical, and academic knowledge and abilities for candidates applying to graduate school programs that do not require a quantitative score.

This document provides information and research results related to the reliability and validity of the MAT. Information in this document includes reliability and validity data based on the current normative sample—MAT candidates who took the test for the first time between January 1, 2008, and December 31, 2011 ( $n = 125,910$ )—and a summary of research findings related to the validity of the test.

Additional publications are also available on the MAT website: “MAT Basics,” with information about MAT history, contents, structure, administration, and score reporting, as well as information useful in interpreting MAT scores; “Understanding Analogies,” showing the usefulness of the analogy item format for measuring important cognitive skills; and “Predicting Performance with the *Miller Analogies Test*,” a comparison of predictive validity data obtained from independently conducted meta-analyses of MAT and GRE research.

Also available only to qualified professionals, the printed *MAT Technical Manual* contains detailed data for the current normative sample, the current percentile ranks, and a compendium table that can be used to compare the previous percentile ranks with the current percentile ranks.

To request a copy of the *MAT Technical Manual*, or to offer suggestions regarding the MAT or about this or any other related publications, please contact MAT Customer Relations at the following email address: [MATScoring.Services@Pearson.com](mailto:MATScoring.Services@Pearson.com).

# MAT Reliability

A test’s reliability is the extent to which the test yields consistent results. Considerations of the reliability of a test must show consistent data across multiple test administrations, for different methods of test delivery, or within a test form. This section will examine MAT reliability in terms of internal consistency and standard error of measurement, two statistics appropriate for use with a test like the MAT that candidates take in a single computer-based administration.

## Internal Consistency

One common measure of a test’s reliability is internal consistency, which is appropriate to use when candidates take a multiple-choice test in a single test administration. A commonly used formula for estimating the reliability of a test based on internal consistency is the Kuder-Richardson formula 20 (KR-20), which requires only a single administration of a test. KR-20 yields a reliability estimate of internal consistency (or reliability coefficient) by examining the homogeneity of the questions within a test based on the number of items in the test, the standard deviation of the total score, and the proportion of candidates correctly answering each item. KR-20 analysis assumes that the items on a test form vary in difficulty, and the resulting reliability coefficient represents an average of all possible split-half reliability estimates.

Reliability coefficients can range from 0.00 to 1.00, with a value of 1.00 indicating that the test is perfectly consistent. The reliability coefficients shown in Table 1 represent the internal consistency of all MAT test forms administered during the 2008–2011 normative sample period for the total normative sample (Total Group) and for the six intended graduate major groups that are reported on Official Transcripts. All of the reliability coefficients shown in Table 1 are at least 0.90, and all are satisfactory according to commonly accepted criteria (Anastasi & Urbina, 1997; Streiner, 2003).

**Table 1** Reliability Coefficients (KR-20) and Standard Errors of Measurement (SEM) for MAT Scaled Scores by Total Group and Intended Field of Study

Total Group and Intended Field of Study	<i>n</i>	Reliability	SEM
Total Group	125,910	0.91	7.46
Business and Management	4,689	0.90	7.02
Education	63,301	0.91	6.93
Humanities and Arts	6,212	0.92	6.98
Natural and Applied Sciences	14,078	0.90	6.88
Social and Behavioral Sciences	26,921	0.91	7.00
Other Fields	8,297	0.93	6.98

**Note.** Total Group = All normative samples from 2008 to 2011 (including those who did not report their intended major); Reliability = Average reliabilities across all forms using Fisher’s z transformation; SEM = Average SEMs of all forms reported in scaled score units.

## Standard Error of Measurement (*SEM*)

The standard errors of measurement (*SEM*) shown in Table 1 also relate to the reliability of a test. The *SEM* represents an estimate of the possible amount of error present in a test score, or the amount that scores would probably vary if a candidate were tested repeatedly with the same test. Because repeated testing always results in some variation, no single test event ever measures a candidate's actual ability with complete accuracy. For this reason, a statistical allowance must be made for a test score to represent, with a reasonable degree of certainty, an estimate of a candidate's ability. The *SEM* is used with a candidate's test score to calculate a score range that includes a reasonable estimate of his or her true score. Thus, the lower the *SEM*, the closer candidates' test scores are to their actual ability, and the greater degree of certainty that the test scores are reliable.

For example, if a MAT candidate obtains a scaled score of 400, the Total Group *SEM* of 7.46 shown in Table 1 suggests that one could be about 68% confident that the candidate would earn, upon repeated attempts, a score between 392.5 and 407.5 (i.e.,  $400 \pm 7.46$ ) and be 95% confident of a score between 385.4 and 414.6 (i.e.,  $400 \pm 7.46 (1.96)$ ).

# MAT Validity

The validity of a test is defined as the degree to which the test actually measures what it is intended to measure. There are several types of test validity evidence. For the purposes of the MAT, evidence of construct, content, and predictive validity are examined in this section.

## Construct Validity—the MAT Analogy

Construct validity is the degree to which the test measures a meaningful theoretical construct or characteristic. The construct validity of the MAT depends on the value of the analogy format in assessing abilities that are essential to success in graduate school. Evidence from research suggests that it does (Kuncel, Hezlett, & Ones, 2004).

Studies that involve analogies typically structure analogy problems first with a base or source analogue that involves a specific type of relationship between two terms. Then in a target analogue that is composed of one given term and one missing term, the missing term must be selected from among options. The correct option must be selected by inferring the relationship between the two terms in the base analogue and mapping this relationship onto the one given term and the correct option used to complete the target analogue. Mapping—seeing the relationship between pairs of related terms—is the essence of an analogy (see also the “Understanding Analogies” document).

In studies of human intelligence and reasoning, researchers have found that verbal, quantitative, and figural analogies are among the best measures of verbal comprehension and analytical intelligence (Gentner, Holyoak, & Kokinov, 2001; Gentner & Markman, 1997; Holyoak & Thagard, 1996; Sternberg, 1977, 1985, 1988). Many cognitive psychologists also suggest that the cognitive process involved in creating analogies and recognizing conceptual similarities despite surface differences has practical benefits.

Many cognitive psychologists suggest that the ability to think analogically has practical benefits in activities such as problem solving, constructing explanations, and building arguments. According to Sternberg (1985), solving analogy problems involves all seven of the information processing components characteristic of inductive reasoning:

1. Encoding—comprehending relevant information to enable interpretation
2. Inference—relating a given concept to another concept
3. Mapping—recognizing a common rule shared by two concepts
4. Application—applying a rule inferred from one set of concepts to another set of concepts
5. Comparison—choosing an option that best conforms to an ideal
6. Justification—determining the reasonableness of a choice relative to an ideal
7. Response—expressing a choice determined through inductive reasoning

In a meta-analysis of 127 studies involving more than 20,000 participants, Kuncel, Hezlett, and Ones (2004) found that the MAT measures abilities that other cognitive ability instruments measure, and suggested that analogical reasoning involves all the principles of cognition. They

also suggest that the MAT is a useful measure of general cognitive ability because it is composed of analogy items that require reasoning with vocabulary knowledge as well as knowledge of various domains, including the sciences, literature, arts, and history. The authors conclude that a strong relationship exists between general cognitive ability and the acquisition of knowledge and skills (learning), that there are consistent positive relationships between general cognitive ability and academic and work performance, and that the MAT can be a valid predictor of performance in both settings.

## Content Validity

Another important type of validity is content validity—the degree of correspondence between the contents of the test and the logical and curricular domains intended to be measured. The MAT analogy items have been designed and constructed to measure knowledge, skills, and abilities considered necessary for success in graduate school. Such traits include the ability to analyze and interpret information, to think logically and critically, and to understand relationships across a broad spectrum of subjects (see the “MAT Basics” document).

To ensure that MAT test items involve the knowledge, skills, and abilities considered necessary for success in graduate school, all items have been written and reviewed by contracted individuals with content expertise according to detailed guidelines provided to them by Pearson. Prior to field-testing new test items, all items are reviewed by Pearson subject-matter experts and editorial staff for content appropriateness, style and format consistency, and gender and ethnic bias. Only items that are judged satisfactory through this process are considered for inclusion as items on MAT test forms.

## Predictive Validity

An important way of evaluating the validity of a high-stakes standardized admission test is to examine its predictive validity, a type of criterion validity that addresses how well a test, like the MAT, predicts later criteria, such as grade point averages or program completion. Many studies conducted over the years have found the MAT to be a moderate to strong predictor of subsequent performance, with predictive validity statistics comparable to other standardized admission tests commonly used by graduate and professional schools (Kuncel, Hezlett, & Ones, 2004; Kuncel & Hezlett, 2007a; Kuncel & Hezlett, 2007b).

Studies have compared the predictive validity of the MAT to other graduate school admission tests. Robertson and Hall (1964) found that a combination of MAT scores, *Graduate Records Examinations*<sup>®</sup> (GRE<sup>®</sup>) scores, and undergraduate grade point average (GPA) was the most promising predictor of faculty ratings, peer ratings, and comprehensive examination scores. In a nine-year study, Furst and Roelfs (1979) compared the predictive validity of the MAT to the GRE used in a doctoral program in education. Both the MAT and GRE scores showed moderate to low correlations with criterion measures. In another study, DeCato (1982) suggested that neither MAT scores, GRE scores, nor previous GPAs predicted performance in a specific course, but that these variables could be useful in assessing a general factor of scholastic ability.

One study that examined the relationship between the MAT and the *Graduate Management Admission Test*<sup>®</sup> (GMAT<sup>®</sup>) found significant relationships between graduate GPA and both the MAT and GMAT, with correlations of 0.282 and 0.410, respectively (Graham, 1991).

In a comparison of meta-analyses of several standardized tests commonly used for graduate school admissions, Kuncel and Hezlett (2007a, 2007b) found nearly identical correlations (observed and corrected for measurement unreliability and restriction of range) between first-year graduate GPA and MAT scores (0.29 observed, 0.41 corrected), GRE General Test Total scores (0.27 observed, 0.41 corrected), and GMAT scores (0.27 observed, 0.41 corrected). They also found overall graduate GPA to correlate slightly higher with MAT scores (0.27 observed, 0.39 corrected) than with either GRE Total scores (0.25 observed, 0.37 corrected) or GMAT scores (0.25 observed, 0.35 corrected).

Some studies have compared the MAT to other predictors in various specific majors. In an early study of graduate psychology majors, Platz, McClintock, and Katz (1959) considered previous GPA, doctoral examinations, faculty ratings, and MAT scores as predictors of graduate course grades and scientific contributions. These authors found that MAT scores were the single best predictor of these outcomes. In a study by Littlepage, Bragg, and Rust (1978), the researchers used several variables to predict graduate school and professional performance in psychology, including GPA, MAT scores, the English Cooperative Test, GRE-Q, faculty and employer ratings of students, a survey administered to the graduate students concerning job duties, salary, job title, and undergraduate major. The researchers found that both MAT score and having an undergraduate major in psychology were significant predictors of subsequent graduate school performance. They also found that MAT scores, the undergraduate major, and faculty ratings were significant predictors of professional performance. Tyron and Tyron (1986) indicated that both MAT and GRE scores could be used to predict a psychologist-trainee's ability to readily engage clients and that high engagers might be identified at admission to graduate school by using these variables. In another study that focused on a psychology program, Huber (1999) used MAT scores, GRE scores, and previous GPA to predict applicant's GPA in a doctoral program in clinical psychology. The results indicated that MAT scores were the most useful predictor of students' subsequent GPA in this program.

Research focusing on the MAT for use with programs in education suggests that MAT scores are useful in predicting graduate course grades. Murray (1979) examined the variability of MAT scores within six departments of education and found significant differences within the departments. Students entering programs in educational psychology, secondary education, and instructional technology had significantly higher mean MAT scores than those in special education, elementary education, and educational administration programs. The researcher concluded that schools should carefully consider candidates in terms of the different areas of education to which they apply. Research by House and Keeley (1993) found that MAT scores were significantly correlated with graduate student performance. Other research findings have suggested that students in master's of education programs with higher MAT scores were more imaginative, intuitive, and abstract in their thinking (Hughes, Costner, & Douzenis, 1988). In two longitudinal studies involving candidates applying to a doctoral program in educational

leadership that bases acceptance decisions on a compensatory model, the researchers found the MAT scores to be a stronger predictor of acceptance (Young, 2007; Young & Young, 2010) and graduation (Young, 2007) than undergraduate GPAs.

The publisher of the MAT has also conducted studies involving MAT scores and subsequent GPA. In one study (The Psychological Corporation, 1992) more than 50 graduate school departments provided information related to the predictive validity of the MAT. Predictive validity in this case was described as the statistical correlation between MAT test score and undergraduate GPA (the predictors) and the first-year graduate GPA. In a more recent study by Pearson (Meagher, 2008), nine graduate schools submitted data for students entering graduate programs in the fall of 2005, including undergraduate GPAs, previous graduate GPAs, MAT scaled scores, GRE scaled scores, demographic information, and GPAs for the first academic year in a graduate program. These data were also analyzed to determine the validity of the MAT and other entering variables for predicting subsequent GPAs using correlation analyses.

Table 2 shows the correlation findings for both the 1992 Psychological Corporation and 2008 Pearson studies. The 1992 study found correlation coefficients with first-year graduate GPA of 0.23 for MAT scores and 0.29 for previously-earned undergraduate GPA. The study completed in 2008 found that the variable with the highest correlation with first-year graduate GPA was previously-earned graduate GPA (0.30), followed by MAT scores (0.27) and GRE-Quantitative scores (0.27). The similar correlation findings from these two studies provide evidence of the consistency of both MAT scores and previously-earned grade point averages in predicting subsequent performance measured by first-year grade point averages. Whether based on the old raw scores obtained exclusively from paper-and-pencil administrations or on more recently-reported scaled scores obtained primarily from computer-based administrations, correlations between MAT scores and first-year graduate school performance have remained moderately positive.

**Table 2** Correlations (*r*) Between Predictor Variables and 1st Year Graduate GPA from the 1992 Psychological Corporation Study and the 2008 Pearson Study

Predictor Variables	<i>n</i>	1st year Graduate GPA
		<i>r</i>
1992: MAT Raw Score	1,035	0.23*
1992: Undergraduate GPA	938	0.29*
2008: MAT Scaled Score	513	0.27*
2008: GRE Verbal Scaled Score	437	0.21*
2008: GRE Quantitative Scaled Score	432	0.27*
2008: GRE Analytical Writing Score	326	0.11
2008: Undergraduate GPA	639	0.24*
2008: Previous Graduate GPA	118	0.30*

Note. \* = Significant at 0.01 level.

The correlations found in both of these studies are also similar to findings in the research literature, particularly in comparison to correlations found in studies of the GRE and in meta-analyses of MAT and GRE research. Correlations between GRE General Test scores and graduate GPAs are found to be around 0.30 (Bridgeman, Burton, & Cline, 2009). In their meta-analysis of MAT studies, Kuncel, Hezlett, and Ones (2004) found a sample-size weighted average correlation of 0.29 between MAT scores and first-year graduate GPA, and an estimated true score validity (corrected for measurement unreliability and restriction of range) of 0.41 between MAT scores and first-year graduate GPA. In their meta-analysis of GRE research, the same researchers (Kuncel, Hezlett, & Ones, 2001) found correlations between GRE scores and first-year graduate GPAs that were similar to those found in the MAT study (sample-size weighted average correlations of 0.24; estimated subtest operational validities ranging from 0.34 to 0.38), which suggests that MAT scores are at least comparable in predictive strength to GRE scores (see the “Predicting Performance with the *Miller Analogies Test*” document).

Kuncel, Hezlett, and Ones’s (2004) meta-analysis represents the most extensive analysis of the predictive validity of the MAT conducted to date. In this study, the researchers found the MAT to be a valid predictor of several aspects of graduate student performance, job performance, potential, and creativity. They found that the MAT correlated highly with the GRE-Verbal section, with other cognitive ability tests from educational and work settings, and with the acquisition and demonstration of knowledge and skills. The authors found that correlations between MAT scores and first-year GPA were slightly higher than for overall graduate GPA, that correlations with faculty ratings of graduate student performance were positive, and that correlations between MAT scores and comprehensive examination scores were the highest. Correlations with degree attainment and time to degree attainment were somewhat positive, possibly moderated by motivational variables other than cognitive ability.

These researchers also found the MAT to have a small, positive correlation with internship and practicum ratings, a moderate correlation with counseling work sample performance, and a near-zero correlation with student-teaching performance, again possibly due to other motivational variables unrelated to cognitive ability. In addition, the authors examined the relationship of the MAT to faculty and work supervisors’ ratings of creativity. They found that MAT scores moderately correlate with ratings of overall potential, with the MAT most highly correlated with counseling potential ratings and less strongly correlated with educational administration performance. The authors concluded that the MAT is a valid measure of both general cognitive ability and verbal ability and that it is useful for predicting performance in both academic and work settings.

The correlations found in these and other studies are comparable to correlations typically found in similar types of analyses, where predictor variables are standardized test scaled scores with a large range of score points (e.g., MAT scaled score range of 200–600) and the outcome variable is a GPA (typically on a more restricted scale of 1–4). Regardless of the limitations inherent in such correlation analyses, consistently similar results found over many years suggest that the MAT continues to have value in predicting graduate program GPAs, particularly for the first year of graduate study, as well as other important academic and non-academic performances.

# Glossary

**Construct Validity**—The degree to which a meaningful theoretical construct or characteristic can be inferred from a test score.

**Content Validity**—Evidence for validity that relates to how adequately the content of a test represents a specified body of knowledge, and to how adequately subjects' responses represent knowledge of the content.

**Correlation**—A measure of the strength and direction of the relationship between two variables. (See “Correlation Coefficient”)

**Correlation Coefficient ( $r$ )**—A statistic ranging between  $-1$  and  $1$  that indicates the degree and direction of relationship between two variables. The strength of the relationship is indicated by the values of the coefficients (with greater values indicating stronger relationships). The direction of the relationship is indicated by either a positive sign (+) representing a positive relationship in which variables tend to increase or decrease together, or a negative sign ( $-$ ) representing an inverse relationship between variables.

**Internal Consistency**—A type of test score reliability indicating the degree of correlation among item scores from different parts of a test.

**Kuder-Richardson Formula 20 (KR-20)**—A statistic (ranging from 0 to 1) used to determine internal consistency reliability as an estimate of the average of all possible split-half coefficients, with a high value indicating a strong relationship among test items and a low value indicating a weak relationship. (See “Correlation Coefficient”)

**Mean ( $M$ )**—The average of a set of scores computed by adding all of the scores together and then dividing by the total number of scores.

**Meta-Analysis**—A method of research that analyzes the results of several independent studies by combining them to determine an overall effect or the degree of relationship between variables.

**N-count ( $n$ )**—The total number of individuals who make up a sample (e.g., the number of candidates who took a test).

**Normative Sample/Norm Group**—The group of individuals (sample) earning scores on a test whose score data are used to determine scaled scores and/or percentile ranks.

**Norm-Referenced Standardized Test**—A measurement in which an individual's scores are interpreted by comparing them to scores obtained from a defined group of individuals (a norm group or normative sample) whose scores have been used to determine scaled scores and/or percentile ranks.

**Percentile Rank (PR)**—A whole number between 1 and 99 that represents the proportion of individuals from the normative sample who earned lower than a given score on a test.

**Predictive Validity**—Evidence for validity based on how accurately test data (e.g., admission test scores) are able to predict criterion measures obtained at some later time (e.g., a grade point average earned after admission).

**Predictor Variable**—A variable that is used to predict some subsequent outcome (e.g., a grade point average earned after admission).

**Raw Score**—The number of items answered correctly by a candidate on a test.

**Reliability**—An estimate of the dependability of test scores in terms of the degree of consistency between administrations.

**Reliability Coefficient**—A statistic (usually ranging from 0 to 1) that measures the degree to which test scores are free of measurement error. (See “Standard Error of Measurement”)

**Scaled Score**—A standardized test score on a specified common scale (e.g., 200–600) with a designated mean and standard deviation that are derived from a raw score (or an ability estimate). Scaled scores are especially useful for comparing performance of individuals or groups over time.

**Standard Deviation (*SD*)**—A measure of the variability of test scores in terms of how spread out scores are from the mean.

**Standard Error of Measurement (*SEM*)**—An estimate (based on group data) of the variation in scores earned on repeated administrations of the same test by the same individual, in the absence of practice effects or change in the underlying ability.

**Validity**—Validity refers to the extent to which evidence supports appropriate interpretations and inferences from test scores regarding characteristics of a person measured (e.g., knowledge or ability) or performances other than those measured (e.g., subsequent performance or achievement).

# References

- Anastasi, A., & Urbina, S. (1997). *Psychological testing* (7th ed.). Upper Saddle River, NJ: Prentice Hall.
- Bridgeman, B., Burton, N., & Cline, F. (2009). A note on presenting what predictive validity numbers mean. *Applied Measurement in Education, 22*, 109–119.
- DeCato, C. M. (1982). Admissions criteria and later performance: Graduate record exam, Miller's analogies, and GPA as predictors of professional competency. *Psychological Reports, 51*(3, Pt. 2), 1149–1150.
- Furst, E. J., & Roelfs, P. J. (1979). Validation of the graduate record examinations and the Miller analogies test in a doctoral program in education. *Educational and Psychological Measurement, 39*(1), 147–151.
- Gentner, D., Holyoak, K. J., & Kokinov, B. N. (Eds.). (2001). *The analogical mind: Perspectives from cognitive science*. Cambridge, MA: MIT Press.
- Gentner, D., & Markman, A. B. (1997). Structure mapping in analogy and similarity. *American Psychologist, 52*, 45–56.
- Graham, L. D. (1991). Predicting academic success of students in a master of business administration program. *Educational and Psychological Measurement, 51*(3), 721–727.
- Holyoak, K. J., & Thagard, P. (1996). *Mental leaps: Analogy in creative thought*. Cambridge, MA: MIT Press.
- House, J. D., & Keeley, E. J. (1993, November). Differential prediction of graduate student achievement from Miller analogies test scores. Paper presented at the annual meeting of the Illinois Association for Institutional Research, Oakbrook Terrace, Illinois.
- Huber, L. L. (1999). *A regression equation used as a criterion for admission to a graduate program in clinical psychology*. Dissertation Abstracts International Section A: Humanities & Social Sciences, 60(2-A).
- Hughes, T. M., Costner, M. J., & Douzenis, C. (1988). Personality and motivation variables in tomorrow's teachers: A center of excellence. *Education, 180*(3), 393–403.
- Kuncel, N. R., & Hezlett, S. A. (2007a). Standardized tests predict graduate students' success. *Science, 315*(5815), 1080–1081.
- Kuncel, N. R., & Hezlett, S. A. (2007b). Supporting online material for standardized tests predict graduate students' success. [Electronic version]. *Science, 315*(5815).
- Kuncel, N. R., Hezlett, S. A., & Ones, D. S. (2001). A comprehensive meta-analysis of the predictive validity of the graduate record examinations: Implications for graduate student selection and performance. *Psychological Bulletin, 127*(1), 162–181.

- Kuncel, N. R., Hezlett, S.A., & Ones, D. S. (2004). Academic performance, career potential, creativity, and job performance: Can one construct predict them all? *Journal of Personality and Social Psychology*, *86*(1), 148–161.
- Littlepage, G. E., Bragg, D. M., & Rust, J. O. (1978). Relations between admission criteria, academic performance, and professional performance. *Teaching of Psychology*, *5*(1), 16–20.
- Meagher, D. (2008). *Miller analogies test: Predictive validity study general report* (Pearson White Paper). San Antonio, TX: NCS Pearson.
- Murray, R. E. (1979). Variability in MAT results within the field of education. *Psychological Reports*, *45*(2), 665–666.
- Platz, A., McClintock, C., & Katz, D. (1959). Undergraduate grades and the Miller analogies test as predictors of graduate success. *American Psychologist*, *14*, 285–289.
- The Psychological Corporation. (1992). *Miller analogies test validity study*. Unpublished research study.
- Robertson, M., & Hall, E. (1964). Predicting success in graduate study. *Journal of General Psychology*, *71*(2), 359–365.
- Streiner, D. L. (2003). Starting at the beginning: An introduction to coefficient alpha and internal consistency. *Journal of Personality Assessment*, *80*(1), 99–103.
- Sternberg, R. J. (1977). *Intelligence, information processing, and analogical reasoning: The componential analysis of human abilities*. Hillsdale, NJ: Lawrence Erlbaum.
- Sternberg, R. J. (1985). *Beyond IQ: A triarchic theory of human intelligence*. Cambridge, MA: Cambridge University Press.
- Sternberg, R. J. (1988). *The triarchic mind: A new theory of human intelligence*. New York, NY: Viking.
- Tyron, G. S., & Tyron, W. W. (1986). Factors associated with clinical practicum trainees' engagements of clients in counseling. *Professional Psychology: Research and Practice*, *17*(6), 586–589.
- Young, I. P. (2007). Predictive validity of grade point averages and of the *Miller Analogies Test* for admission to a doctoral program in educational leadership. *Educational Research Quarterly*, *31*(2), 44–54.
- Young, I. P., & Young, K. H. (2010). An assessment of academic predictors for admission decisions: Do applicants varying in national origin and in sex have a level playing field for a doctoral program in educational leadership? *Educational Research Quarterly*, *33*(3), 39–51.