PCAT®

PHARMACY COLLEGE ADMISSION TEST

PCAT Reliability and Validity

Effective: July 2020

Pearson

For inquiries or reordering:
800.622.3231
www.pearsonassessments.com/pcat

NCS Pearson, Inc. 5601 Green Valley Drive Bloomington, MN 55437

Produced in the United States of America.

v.A

Product Number 66108

Table of Contents

Acknowledgments

This document was prepared by the following members of the Pearson Clinical Assessment group:

Donald G. Meagher, EdD
Senior Research Director

Tianshu Pan, PhD
Senior Psychometrician

Natividad Ybarra, PhD
Test Development Manager

Rachel P. Wegner, MA
Research Director

Jeffrey R. Miller, MA
Digital Editor

Introduction

The *Pharmacy College Admission Test* (PCAT) is a norm-referenced standardized test that measures the abilities, aptitudes, and skills that pharmacy schools deem essential for success in their programs. To ensure the continuing relevance and usefulness of the PCAT for assessing the prerequisite knowledge and skills of candidates for admission to professional pharmacy programs, the test content is periodically reviewed and revised.

This publicly available document provides information and research results related to the reliability and validity of the PCAT. Information in this document includes reliability and validity data based on the current normative sample—PCAT candidates who took the test for the first time between July 2015 and April 2019 (N = 52,882)—and a summary of research findings related to the validity of the test.

During the 2011–2015 normative sample period, the PCAT consisted of a Writing subtest and five multiple-choice subtests: Verbal Ability, Biology, Chemistry, Reading Comprehension, and Quantitative Ability. However, beginning with the July 2016 PCAT administration, the test no longer contained a Verbal Ability subtest, and the four remaining multiple-choice subtests were renamed as Biological Processes, Chemical Processes, Critical Reading, and Quantitative Reasoning. So that the information in this document remains current and relevant going forward, references to the multiple-choice subtests will include only the four that are currently part of the PCAT. Also beginning in July 2016, a new test blueprint was implemented in phases, with some passage-dependent items added to the Biological Processes and Chemical Processes subtests, some Critical Reading passages added that contained humanities and social science content, and a greater proportion of scenario type items added to the Quantitative Reasoning subtest. These blueprint changes were fully implemented during the 2018–19 PCAT testing cycle.

The following additional publications are also available on the PCAT website: PCAT Basics, with information about PCAT history, content, structure, administration, and score reporting; and Interpreting PCAT Scores, with information useful in interpreting all PCAT scaled scores, percentile ranks, and Writing scores. Also available only to qualified professionals, the printed *PCAT Technical Manual* contains detailed data for the current normative sample, the current 2019 percentile ranks, and compendium tables that can be used to compare the previous 2015 percentile ranks with the current percentile ranks.

To request a copy of the *PCAT Technical Manual*, or to offer suggestions regarding the PCAT or about this or any other related publications, contact PCAT Customer Relations at Scoring.Services@Pearson.com.

PCAT Reliability

A test's reliability is the extent to which the test yields consistent results across alternate forms, occasions, or methods of administration or scoring. Considerations of the reliability of a test must account for both multiple-choice and writing components, and must show consistent data for different conditions, such as human and electronic essay scoring.

Internal Consistency

One common measure of a test's reliability is internal consistency, which is appropriate to use when candidates take a multiple-choice subtest on a single test administration. Internal consistency reflects the homogeneity of the test content, that is, the degree to which the questions measure the same dimension of ability. Internal-consistency reliability coefficients (coefficient alpha) were determined for the four multiple-choice subtest scores, 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. A reliability coefficient for the Composite score was calculated using a separate composite reliability formula (Feldt & Brennan, 1989).

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 average internal consistency of the four multiple-choice PCAT subtests and of the test as a whole (Composite) across all forms administered during the normative sample period. The reliability coefficients for all subtests and the Composite are satisfactory (≥0.80) according to commonly accepted criteria (Anastasi & Urbina, 1997; Streiner, 2003).

Table 1 Reliability Coefficients and Standard Errors of Measurement (*SEM*s) for the 2015–19 Normative Sample Multiple-Choice Scaled Scores

Subtest/Composite	Reliability	*SEM*
Biological Processes	.85	8.24
Chemical Processes	.81	11.32
Critical Reading	.83	9.37
Quantitative Reasoning	.86	7.47
Composite	.94	4.54

Standard Error of Measurement (*SEM*)

The standard error of measurement (*SEM*) for a test can be thought of as the average discrepancy between a candidate's true score and the score the candidate actually obtains. Because every test is imperfect, no single test event ever measures a candidate's true ability with perfect 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 the size of that allowance. The smaller the *SEM*, the closer candidates' test scores are to their true ability, and the greater the degree of certainty that the test scores are accurate.

The *SEM* shown in Table 1 may be used with a candidate's subtest or Composite score to calculate a score range that represents a reasonable estimate of his or her true ability. For example, if a PCAT candidate obtains a Composite scaled score of 400, the *SEM* of 4.54 shown in Table 1 suggests that one could be about 68% confident that the candidate's true score would be within one *SEM* of 400 (i.e. between 395.5 and 404.5) and be 95% confident that the true score lies within about two *SEM*s (i.e., between 391.1 and 408.9.

Subtest Blueprint Comparability

During the current normative sample period, the old PCAT subtest blueprints were in effect during the 2015–16 testing cycle, new blueprints for each subtest were partially implemented during the 2016–2017 and 2017–2018 cycles, and the subtest blueprints were fully implemented beginning with the 2018–2019 cycle. Comparisons of the means, standard deviations, minimum (Min.) and maximum (Max.) scores, and effect size data for candidates testing during these three periods are shown in Table 2. The effect size data are intended to show the magnitude of scaled score differences between the three groups, with Cohen's *d* used to determine mean differences in terms of standard deviation units. The very small effect sizes noted for each subtest between the mean scores obtained during the three testing periods (less than 0.20) suggest that the differences were negligible (Cohen, 1988).

Table 2 Means, Standard Deviations (*SD*s), and Effect Size Data for PCAT Subtest Blueprints in Effect During Three Testing Periods

						Effect size	
Subtest/Composite	**Blueprint**	**Mean**	***SD***	**Min.**	**Max.**	**2016–18**	**2018–19**
	2015–16	409.4	21.7	339	502	0.00	0.02
	2016–18	409.5	20.6	339	503		0.01
Biological Processes	2018–19	409.7	22.4	319	503		
	2015–16	411.5	26.6	317	526	0.05	0.03
	2016–18	412.8	26.0	304	529		0.07
Chemical Processes	2018–19	410.9	25.1	305	526		
	2015–16	394.7	22.9	289	511	0.08	0.02
	2016–18	393.0	21.9	285	517		0.05
Critical Reading	2018–19	394.2	23.0	309	513		
	2015–16	403.2	20.9	316	518	0.04	0.00
	2016–18	402.4	19.1	316	517		0.04
Quantitative Reasoning	2018–19	403.1	20.3	327	517		
	2015–16	404.8	18.4	342	500	0.01	0.01
	2016–18	404.6	17.7	340	478		0.00
Composite	2018–19	404.6	18.7	344	483		

Writing Score Consistency Across Human Scorers

Data in Table 3 present evidence of the reliability of valid Writing subtest scores (i.e., no 0 scores) earned by candidates in the current norm group as indicated through interscorer consistency—the degree of agreement between scores assigned to the same essay by different human scorers. Table 3 indicates the percentage of reported Writing scores for all candidates in the current norm group that had a 0, 1, or greater than 1 point discrepancy in the scores assigned by two trained scorers. The total percentage of 0–1 score point discrepancies of 99.8% suggests a very high degree of consistency between the two independent scorers in the scoring of the essays, particularly because scores differing by 1 point are averaged to obtain the candidate's reported score.

Table 3 Writing Score Interscorer Agreement for the 2015–19 Normative Sample

	Score point discrepancy			
Score	**0**	**1**	**≥2**	**0–1**
n	40,105	12,645	130	52,750
% of normative sample	75.8	23.9	0.2	99.8

Comparability of Human and Electronic Writing Scores

During a two-year trial period (September 2014 through January 2016), Pearson's Intelligent Essay Assessor™ (IEA) provided one of the two scores obtained for each PCAT Writing subtest essay.

Developed by Pearson Knowledge Technologies, IEA is an Internet-based tool designed to automatically score electronically submitted essays. IEA uses a scoring engine that analyzes text through a combination of latent semantic analysis (LSA) and methods widely used in automatic speech recognition, computational linguistics, and other forms of statistical artificial intelligence. LSA analyzes text by determining semantic similarities for words and extended passages, which enables the IEA tool to evaluate essays for both substantive content and mechanical aspects of writing.

Pearson provided Knowledge Technologies with over 23,000 PCAT essays with over 46,000 assigned scores that were used to program the IEA scoring engine for use with the seven Writing prompts administered from September 2014 through January 2016.

Table 4 shows the interscorer agreement results for essays scored by one human scorer and the IEA electronic scoring tool (the Sept. 2014 through Jan. 2016 sample) and for essays scored by two human scorers (the July 2016 through April 2019 sample). The results shown in Table 4 indicate a somewhat greater proportion of 1-point discrepancies for essays scored by human and IEA scorers compared to two human scorers (29.9% and 21.7%, respectively). However, when the 0-point and 1-point discrepancies are considered together, both tables show very similar results: 99.5% and 99.9% for human and IEA scorers and two human scorers, respectively. Even though these data show that humans agree more with each other than with IEA, the overall human-IEA agreement is still very high and is sufficient to support its operational use.

Table 4 Writing Score Interscorer Agreement: September 2014 Through January 2016 (Human / IEA Scores) Compared to July 2016 Through April 2019 (Human / Human scores)

Sample		Score point discrepancy			
		0	1	≥2	0–1
Human/IEA Scores	*n*	10,135	4,348	75	14,483
	%	69.6	29.9	0.5	99.5
Human/Human Scores	*n*	29,970	8,297	55	38,267
	%	78.2	21.7	0.1	99.9

PCAT Writing subtest scores earned beginning with the July 2020 administration include some assigned by human and IEA scorers and some assigned by two human scorers, with all cases of more than 1-point discrepancy resolved by a human resolution scorer.

PCAT Validity

The validity of test scores is the accuracy of inferences based on those scores. There can be various kinds of inference. For example, one might infer from a high score on the Chemical Processes subtest that the candidate knows a lot about chemistry, or one might infer from a high Composite score that the candidate is likely to do well in pharmacy school. Different kinds of evidence can be used to support different kinds of inference. This section presents evidence related to the content of the PCAT, its internal structure, and the relationship of PCAT scores to other variables.

Content Validity

One important type of validity is content validity—the degree of correspondence between the content of the test and the curricular domains measured. Pearson has worked closely with the American Association of Colleges of Pharmacy (AACP) PCAT Advisory Committee over the years to ensure that the test includes content that is relevant to pharmacy school prerequisite requirements and to the pharmacy school curricula.

In addition to PCAT content being monitored by the PCAT Advisory Committee, all PCAT test items are designed and constructed to measure knowledge, skills, and abilities considered necessary for success in pharmacy school (see the PCAT Basics document). Contracted individuals with content expertise have written and reviewed all PCAT test items according to detailed guidelines provided to them by Pearson. Prior to field-testing new test items, Pearson subject-matter experts and editorial staff review all items 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 PCAT test forms.

Internal Structure Validity

Internal structure validity involves the degree to which psychometric relationships among components within a test are consistent with the intended meaning of scores for those components. Each PCAT subtest is designed to measure different skills and content, and examining the correlations between subtests indicates whether each subtest is performing as expected.

Table 5 shows the relationships among the multiple-choice subtest scaled scores and Writing scores earned by candidates in the current norm group. These relationships are displayed in terms of a correlation coefficient (*r*), a measure of the strength of the relationship that can range from –1.0 to 1.0. A correlation coefficient of 1.0 would represent a perfect correspondence, suggesting that the subtests are measuring the same thing, and a coefficient of 0.0 would indicate no relationship. Higher correlation coefficients suggest that similar abilities or knowledge are being assessed, and correlation coefficients closer to zero suggest that the subtests are measuring more distinct abilities or knowledge.

Table 5 Intercorrelations Among PCAT Subtest and Composite Scores for the 2015–19 Normative Sample

Subtest/Composite	Intercorrelations				
	Biological Processes	Chemical Processes	Critical Reading	Quantitative Reasoning	Composite
Chemical Processes	.68				
Critical Reading	.50	.42			
Quantitative Reasoning	.52	.63	.45		
Composite	.83	.86	.73	.79	
Writing	.27	.26	.37	.29	.37

As shown in Table 5, the highest correlation for each of the multiple-choice subtests is with the Composite score (.73–.86). In part, this is because the Composite includes each of these subtests. Among the subtests, the highest correlations are between Biological Processes and Chemical Processes (.68) and between Chemical Processes and Quantitative Reasoning (.63). The lowest correlations are between the Writing scores and all the other subtests (.26–.37), followed by Critical Reading with the other three multiple-choice subtests (.42–.50). These lower correlations highlight the difference between the Writing task and the content knowledge required by the other subtests, and reflect the very different content and skills assessed on the Critical Reading subtest and the other multiple-choice subtests.

Relationship of PCAT Scores to Other Ability Measures

Some research over the years has addressed the validity of the PCAT through correlation analyses of PCAT scores and other assessment scores obtained prior to admission to pharmacy school. Related to the PCAT as an assessment of critical thinking skills, studies have found PCAT scores to have positive correlations with scores on the *California Critical Thinking Skills Test* and the *Health Science Reasoning Test* administered to candidates for admission during the interview process (Allen & Bond, 2001; Kelsch & Friesner, 2014; McCall, MacLaughlin, Fike, & Ruiz, 2007). The authors of these studies suggest that the PCAT represents an adequate measure of critical thinking skills.

Relationship of PCAT Scores to Success in Pharmacy School

Predictive validity is based on evidence of how well a test predicts later criteria, such as grade point averages (GPAs) or licensing-exam scores. Since the introduction of the PCAT in 1974, studies conducted over the years have found the test 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 et al., 2005; Kuncel & Hezlett, 2007).

Such studies have typically used correlation and regression analyses to examine the predictive value of PCAT scores, and some have used other variables to make predictions about subsequent GPAs earned by students in pharmacy school. Studies that have considered both academic and demographic or other noncognitive characteristics have found PCAT scores and entering GPAs to be better indicators of pharmacy school performance than noncognitive variables (Allen & Bond, 2001; Carroll & Garavalia, 2002; Hardigan, Lai, Arneson, & Robeson, 2001; Kidd & Latif, 2003; Lobb, Wilkin, McCaffrey, Wilson, & Bentley, 2006). Other studies have found the PCAT to be useful for predicting performance in pharmacy programs and on prelicensing and licensing exams (Chisholm-Burns, Spivey, McDonough, Phelps, & Byrd, 2014; Cunny & Perri, 1990; Fuller, Horlen, Cisneros, & Merz, 2007; Kuncel et al., 2005; McCall et al., 2007; Schauner, Hardinger, Graham, & Garavalia, 2013; Shaya, Lebovitz, Gaitonde, Schlesselman, & Buring, 2016; Thomas & Draugalis, 2002).

In the most comprehensive PCAT study conducted to date, Kuncel et al. (2005) found in a meta-analysis of previous research that both PCAT scores and prepharmacy GPAs were valid predictors of student success during the first 3 years of pharmacy school (especially during the first-year), and that the PCAT was a strong predictor of performance on professional licensing examinations. It has also been shown by others that undergraduate GPAs, prerequisite GPAs, and PCAT scores are predictive of success in didactic-rich pharmacy curriculums (Eiland, Gaillard, Fan, & Jungnickle, 2018; Ferrante, Lambert, Leggas, & Black, 2017; Muratov, Lewis, Fourches, Tropsha, & Cox, 2017; Tejada, Parmar, Purnell, & Lang, 2016), as well as performance on the *Pharmacy Curriculum Outcomes Assessment* (PCOA), a test that measures student performance on pharmacy curricula (McDonough, Spivey, Chisholm-Burns, Williams & Phelps, 2019). These studies suggest that, when considered together with prepharmacy GPAs, PCAT scores can be used in admissions decisions to substantially increase the likelihood of identifying students who will perform successfully in pharmacy programs.

In cooperation with the AACP PCAT Advisory Committee, Pearson conducted correlation, regression, discriminant, and diagnostic accuracy analyses to determine the validity of PCAT scores for predicting GPAs of students in years one through four of pharmacy programs (Meagher, Lin, & Perez, 2007; Meagher, Lin, & Stellato, 2006). These studies demonstrated the increase in predictability when PCAT scores were added to prepharmacy GPA alone, and showed that all the multiple-choice subtests were predictive of first-year pharmacy GPA. The results of these studies substantiated previous findings that both PCAT scaled scores and entering cumulative GPAs have moderate to strong predictive validity for success in pharmacy school.

A more recent study conducted by Pearson examined data collected on more than 2,200 students from 22 pharmacy school PharmD programs around the country in order to determine the value of recent PCAT scaled scores, PCAT Writing scores, and entering GPAs in predicting subsequent GPAs during the first year of professional pharmacy study (Meagher, Pan, & Perez, 2011). Table 6 shows correlations—both uncorrected (r) and corrected for range restriction (r^*)—found in this study between PCAT scaled scores, PCAT Writing scores, and pre-pharmacy GPAs and first-year pharmacy school GPA. All correlations shown in Table 6 are weighted means of the correlation data using Fisher's z transformation, and corrected correlations were computed using a formula to adjust for range restriction (Cohen, et al., 2003, p. 58).

Table 6 Correlations Between Selected Predictors and First-Year Pharmacy School GPA

Variable	*r*	*r**
PCAT Biology	.26	.35
PCAT Chemistry	.29	.36
PCAT Reading	.17	.24
PCAT Quantitative	.21	.28
PCAT Composite	.32	.44
PCAT Writing	.06	—
Entering cumulative GPA	.44	—

The results of this study were comparable in many ways to the findings of previous studies of PCAT and other graduate and professional school admission tests. Correlations between PCAT scores and first-year pharmacy GPAs were similar to those found in the meta-analysis of PCAT predictive validity studies by Kuncel et al. (2005), as well as those found in studies examining correlations between the *Graduate Record Examinations* (GRE) General Test and first-year GPAs in graduate school (Bridgeman, Burton, & Cline, 2009) and law school (Klieger, Bridgeman, Tannenbaum, Cline, & Olivera-Aguilar, 2018). In addition, the findings from regression analyses in the PCAT study compared favorably to the findings of a study that examined the relationship between the *Medical College Admission Test* (MCAT) and GPAs for the first 2 years of medical school (Callahan, Hojat, Veloski, Erdmann, & Gonnella, 2010; Dixon, 2012), as well as to regression results typically found for the GRE (Bridgeman, Burton, & Cline, 2009).

PCAT scores have also been found to predict whether students graduating from pharmacy school pass the *North American Pharmacist Licensure Examination* (NAPLEX). Recently, Shaya et al. (2016) reported that final cumulative pharmacy school GPA, age less than 25 years, and PCAT Chemistry and Reading Comprehension scores were predictors of higher NAPLEX scores. Likewise, Eiland et al. (2018) communicated that prepharmacy science GPAs and PCAT Biology scores predicted students' passing scores on the NAPLEX in 2014 and 2015, with higher PCAT Biology and team activity scores increasing the likelihood of earning passing NAPLEX scores in 2016.

While studies conducted over the years have continued to suggest that both PCAT scores and prepharmacy GPAs show consistently positive correlations with subsequent performance in pharmacy school (especially with first-year GPA) and on the NAPLEX (Eiland et al., 2018; Kuncel et al., 2005; McCall et al., 2007; Shaya et al., 2016), many older studies were conducted before the PharmD became the required degree for professional pharmacy practice (Lobb & Wilkin, 2003), and relatively few studies had considered demographic characteristics as predictor variables.

In addition to examining predictors of performance in pharmacy school such as PCAT scores and previously-earned GPAs, the 2009–2010 Pearson study also attempted to shed light on differing PCAT mean scores for demographic groups by including demographic characteristics as predictor variables. With regard to the influence of demographic variables, the Pearson study found PCAT scores to be equally strong predictors of first-year pharmacy school GPA, regardless of students' sex, race, or ethnicity. PCAT scores tended to overpredict the actual GPAs of Hispanic students (i.e., scores favored Hispanics) and were stronger predictors of GPA for students with lower levels of parent education. These findings provide evidence that observed score differences between demographic groups do not result in differential predictions of performance in pharmacy school in ways that would disadvantage individuals from underrepresented minority groups.

Some other studies that have examined questions related to demographic factors have found sex, ethnicity, and native language to be associated with key pharmacy school performance variables (Bandalos & Sedlacek, 1989; Chisholm, Cobb, & Kotzen, 1995; Chisholm-Burns, Spivey, Byrd, McDonough, & Phelps, 2017; Chisholm-Burns et al., 2014; Kelly, Secnik, & Boye, 2001; Tejada et al., 2016; Wu-Pong, Windridge, & Osborne, 1997). However, other studies found comparable performances among students in terms of course grades regardless of sex or ethnicity (Carroll & Garavalia, 2002), and in terms of course exam results, regardless of previous educational background, age, sex, or ethnicity (Ried & Byers, 2009).

Regarding correlations between the educational background and subsequent performance in pharmacy school, studies have shown differing results. In contrast to what Ried and Byers reported, other researchers have observed that earning a degree prior to pharmacy school is a strong predictor of subsequent performance (Myers, DeHart, Vuk, & Bursac, 2013) and that the competitiveness and level of colleges attended prior to matriculation into a PharmD program adds predictive value to PCAT scores and entering GPAs only for students' performance in the fourth year of pharmacy study (Hall & White, 2007; White & Hall, 2006). Several other studies found that pharmacy students with bachelor's degrees outperformed students with less prior education in terms of first professional-year GPAs (Chisholm, 2001; Chisholm, Cobb, Kotzen, & Lauthenschlager, 1997; Chisholm, Cobb, DiPiro, & Lauthenschlager, 1999; Houglum, Aparasu, & Delfinis, 2005; McCall, Allen, & Fike, 2006; Renzi, Krzeminski, Sauberan, Brazeau, & Brazeau, 2007).

The findings of the 2009–2010 Pearson study seem generally consistent with the findings of these other studies, especially regarding students who earned a bachelor's degree. However, unlike earlier studies, the Pearson study found that a student's age had a negative influence on first-year pharmacy GPAs. With regard to candidates' and parents' educational background, the Pearson study found that neither candidates' last previous attendance at a four-year college or a private school prior to matriculation to a pharmacy school nor parents' attainment of a bachelor's degree were significant factors in predicting first-year pharmacy school GPAs.

The results of this Pearson study are comparable in many ways to other research findings that have shown the moderate validity of PCAT scores in predicting GPAs during the first year of pharmacy school. The Pearson study findings also suggest that previously earned GPAs and PCAT scores do not unfairly represent the potential of individuals from any one demographic group to succeed in pharmacy school. Finally, the small proportions of students found to discontinue or continue on academic probation after the first year of study suggests that if success in pharmacy school is defined as maintaining good academic standing following the first year, then the criteria being used in admission decisions are appropriate and effective, including the use of PCAT scores in combination with previously earned GPAs.

Glossary

Content Validity—A type of 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 sets of 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.

Criterion Validity—A type of validity involving correlations between test scores and scores on some other measure representing a relevant criterion.

Diagnostic Accuracy Analysis—A statistical method used to determine the ability of test scores to accurately identify individuals who are likely to have a specific characteristic (e.g., first-year students likely to earn a specific grade point average).

Discriminant Analysis—A statistical method used to predict group membership (e.g., first-year students earning in the top 10% of grade point averages) from one or more predictor variables (e.g., PCAT scores).

Effect Size—The magnitude of the effect of an experimental treatment, reflected in the difference in average scores between two groups. This difference is often expressed in standard-deviation units.

Internal Consistency—A type of test score reliability reflecting the degree of correlation among scores on items within the test.

Internal Structure Validity—A type of validity involving the degree to which relationships among test items and test components conform to what the test is intended to measure.

Interscorer Reliability—A type of reliability whereby ratings assigned by two or more independent judges of the same test performance (e.g., a written essay) are compared to determine their consistency.

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* or *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 test that presents consistent content, using the same administration conditions and scoring procedures, to all examinees, and that is interpreted by comparing the individual's scores to the scores obtained by a normative sample.

Norms—Data that summarize the performance of a norm group (or normative sample) by showing how earned scores compare to one another, such as by listing scaled scores and corresponding 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—A type of criterion validity based on how accurately test data (e.g., admission test scores) predict criterion measures obtained at some later time (e.g., a grade point average earned after admission).

Predictor Variable—A variable that occurs prior to any intervention (e.g., scores on an admission test) that is used to predict some subsequent outcome (e.g., a grade point average earned after admission). (*See* Variable.)

Regression—An indication of the relationship between variables for purposes of predicting a measure of one variable (e.g., first-year grade point average) from a measure of another variable (e.g., PCAT Composite score).

Reliability—An estimate of the dependability of test scores in terms of the degree of consistency between measures of the test (e.g., comparisons of administrations of a test over time, or comparisons of items within a test).

Reliability Coefficient—A correlation 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 (SS)—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 in a content area (e.g., biology).

Standard Deviation (*SD*)—A measure of the variability of test scores.

Standard Error of Measurement (*SEM*)—An estimate of the variation in scores earned on repeated administrations of the same test by the same individual. The *SEM* is the standard deviation (*SD*) of the measurement error distribution. The *SEM* is inversely related to reliability: As the *SEM* decreases, reliability increases, and confidence in the observed test score precision increases. The *SEM* is used to calculate confidence intervals, or bands of scores around observed scores, in which true scores are likely to fall. Confidence intervals express test score precision and serve as reminders that measurement error is inherent in all test scores and that observed test scores are only estimates of true ability.

Validity—The accuracy of inferences based on test scores. Validity refers to the extent to which test scores (or other measures) support appropriate interpretations and inferences regarding characteristics of a person measured (e.g., knowledge or ability) or performances other than those measured (e.g., subsequent performance or achievement).

References

Allen, D. D., & Bond, C. A. (2001). Prepharmacy predictors of success in pharmacy school: Grade point averages, pharmacy college admissions test, communication abilities, and critical thinking skills. *Pharmacotherapy*, *21*, 842–849.

Anastasi, A., & Urbina, S. (1997). *Psychological testing* (7th ed.). Upper Saddle River, NJ: Prentice Hall.

Bandalos, D. L., & Sedlacek, W. E. (1989). Predicting success of pharmacy students using traditional and nontraditional measures by race. *American Journal of Pharmaceutical Education*, *53*, 145–148.

Bridgeman, B., Burton, N., & Cline, F. (2009). A note on presenting what predictive validity numbers mean. *Applied Measurement in Education*, *22*, 109–119.

Callahan, C. A., Hojat, M., Veloski, J., Erdmann, J. B., & Gonnella, J. S. (2010). The predictive validity of three versions of the MCAT® in relation to performance in medical school, residency, and licensing examinations: A longitudinal study of 36 classes of Jefferson Medical College. *Academic Medicine*, *85*(6), 980–987.

Carroll, C. A., & Garavalia, L. S. (2002). Gender and racial differences in select determinants of student success. *American Journal of Pharmaceutical Education*, *66*, 382–387.

Chisholm, M. (2001). Students' performance throughout the professional curriculum and the influence of achieving a prior degree. *American Journal of Pharmaceutical Education*, *65*, 350–354.

Chisholm, M., Cobb, H., DiPiro, J. T., & Lauthenschlager, G. J. (1999). Development and validation of a model that predicts the academic ranking of first-year pharmacy students. *American Journal of Pharmaceutical Education*, *63*, 388–394.

Chisholm, M., Cobb, H., & Kotzen, J. A. (1995). Significant factors for predicting academic success of first-year pharmacy students. *American Journal of Pharmaceutical Education*, *59*, 364–370.

Chisholm, M., Cobb, H., Kotzen, J. A., & Lauthenschlager, G. J. (1997). Prior four year college degree and academic performance of first year pharmacy students: A three year study. *American Journal of Pharmaceutical Education*, *61*, 278–281.

Chisholm-Burns, M. A., Spivey, C. A., Byrd, D. C., McDonough, S. L. K., Phelps, S. J. (2017). Examining the association between the NAPLEX, pre-NAPLEX, and pre- and post-admission factors. *American Journal of Pharmaceutical Education*, *81*(5), Article 86.

Chisholm-Burns, M. A., Spivey, C. A., McDonough, S., Phelps, S., & Byrd, D. (2014). Evaluation of student factors associated with Pre-NAPLEX scores. *American Journal of Pharmaceutical Education*, *78*(10), Article 181.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences.* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.

Cohen, J., Cohen, P., West, S.G., & Aiken, L.S. (2003). *Applied multiple regression/correlation analysis for social science* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

Cunny, K. A., & Perri, M. (1990). Historical perspective on undergraduate pharmacy student admissions: The PCAT. *American Journal of Pharmaceutical Education*, *54*, 1–6.

Dixon, D. (2012). Prediction of osteopathic medical school performance on the basis of MCAT score, GPA, sex, undergraduate major, and undergraduate institution. *The Journal of the American Osteopathic Association*, *112*(4), 175-181.

Eiland, L. S., Gaillard, P. R., Fan, S., & Jungnickle, P. W. (2018). Differences in predictors of academic success using multi- and individual year student admissions data. *Pharmacy Education*, *18*(1), 255–258.

Feldt, L. S., & Brennan, R. L . (1989). Reliability. In R. L. Linn (Ed.), *Educational measurement* (3rd. ed., pp.105–146). New York: Macmillan.

Ferrante, A. B., Lambert, J., Leggas, M., & Black, E. P. (2017). Predicting student success using in-program monitoring. *American Journal of Pharmaceutical Education*, *81*(6), Article 111.

Fuller, S., Horlen, C., Cisneros, R., & Merz, T. (2007). Pharmacy students' reading ability and the readability of required reading materials. *American Journal of Pharmaceutical Education*, *71*(6), Article 111.

Hall, P., & White, R. L. (2007). Completion of pre-pharmacy requirements at community colleges: What is the association with PharmD GPA? [Abstract]. *ACCP [American College of Clinical Pharmacy] 2007 Annual Meeting Guide and Abstracts*, 207, Abstract #63E.

Hardigan, P. C., Lai, L. L., Arneson, D., & Robeson, A. (2001). Significance of academic merit, test scores, interviews and the admissions process: A case study. American Journal of Pharmaceutical Education, 65, 40–44.

Houglum, J. E., Aparasu, R. R., & Delfinis, T. M. (2005). Predictors of academic success and failure in a pharmacy professional program. *American Journal of Pharmaceutical Education*, *69*(3), Article 43.

Kelly, K. A., Secnik, K., & Boye, M. E. (2001). An evaluation of the pharmacy college admissions test as a tool for pharmacy college admissions committees. *American Journal of Pharmaceutical Education*, *65*, 225–230.

Kelsch, M. P., & Friesner, D. L. (2014). The Health Sciences Reasoning Test in the pharmacy admissions process. *American Journal of Pharmaceutical Education*, *78*(1), Article 9.

Kidd, R. S., & Latif, D. A. (2003). Traditional and novel predictors of classroom and clerkship success of pharmacy students. *American Journal of Pharmaceutical Education*, *67*(4), Article 109.

Klieger, D.M., Bridgeman, B., Tannenbaum, R.J., Cline, F.A., Olivera-Aguilar, M. (2018). The validity of GRE general test scores for predicting academic performance at U. S. law schools (ETS Research Report No. RR-18-26). Retrieved from https://onlinelibrary.wiley.com/doi/epdf/10.1002/ets2.12213.

Kuncel, N. R., Crede, M., Thomas, L. L., Kleiger, D. M., Seiler, S. N., & Woo, S. E. (2005). A meta-analysis of the validity of the Pharmacy College Admission Test (PCAT) and grade predictors of pharmacy student performance. *American Journal of Pharmaceutical Education*, *69*(3), Article 51.

Kuncel, N. R., & Hezlett, S. A. (2007). Standardized tests predict graduate students' success. *Science*, *315*, 1080–1081.

Lobb, W. B., & Wilkin N. E. (2003). Admission and progression standards at U.S. schools and colleges of pharmacy: An exploration of criteria changes. *American Journal of Pharmaceutical Education*, *67*(3), Article 93.

Lobb, W. B., Wilkin, N. E., McCaffrey, D. J., Wilson, M. C., & Bentley, J. P. (2006). The predictive utility of nontraditional test scores for first-year pharmacy student academic performance. *American Journal of Pharmaceutical Education*, *70*(6), Article 128.

McCall, K. L., Allen, D. D., & Fike, D. S. (2006). Predictors of academic success in a doctor of pharmacy program. *American Journal of Pharmaceutical Education*, *70*(5), Article 106.

McCall, K. L., MacLaughlin, E. J., Fike, D. S., & Ruiz, B. (2007). Preadmission predictors of PharmD graduates' performance on the NAPLEX. *American Journal of Pharmaceutical Education*, *71*(1), Article 05.

McDonough, S. L. K., Spivey, C. A., Chisholm-Burns. M. A., Williams, J. S., & Phelps, S. J. (2019). Examination of factors relating to student performance on the pharmacy curriculum outcomes assessment. *American Journal of Pharmaceutical Education*, *83*(2), Article 6516.

Meagher, D. G., Lin, A., & Perez, C. P. (2007). Predicting pharmacy students' performance from PCAT scores: A further analysis of predictive validity study data. *American Journal of Pharmaceutical Education*, *71*(5), Article 101.

Meagher, D. G., Lin, A., & Stellato, C. P. (2006). A predictive validity study of the Pharmacy College Admission Test. *American Journal of Pharmaceutical Education*, *70*(3), Article 53.

Meagher, D. G., Pan, T., & Perez, C. P. (2011). Predicting performance in the first-year of pharmacy school. *American Journal of Pharmaceutical Education*, *75*(5), Article 81.

Muratov, E., Lewis, M., Fourches, D., Tropsha, A., Cox, W.C. (2017). Computer-assisted decision support for student admissions based on their predicted performance. *American Journal of Pharmaceutical Education*, *81*(3), Article 46.

Myers, T. L., DeHart, R. M., Vuk, J., & Bursac, Z. (2013). Prior degree status of student pharmacists: Is there an association with first-year pharmacy school academic performance? *Currents in Pharmacy Teaching and Learning*, *5*(5), 490–493.

Renzi, S. E., Krzeminski, M. A., Sauberan, M. M., Brazeau, D. A., & Brazeau, G. A. (2007). Prepharmacy years in college and academic performance in a professional program. *American Journal of Pharmaceutical Education*, *71*(4), Article 69.

Ried, L. D., & Byers, K. (2009). Comparison of two lecture delivery platforms in a hybrid distance education program. *American Journal of Pharmaceutical Education*, *31*(5), Article 95.

Schauner, S., Hardinger, K. L., Graham, M. R., & Garavalia, L. (2013). Admission variables predictive of academic struggle in a PharmD program. *American Journal of Pharmaceutical Education*, *77*(1), Article 8.

Shaya, F. T., Lebovitz, L., Gaitonde, P., Schlesselman, L. S., & Buring, S. (2016). The well-rounded applicant (3 schools, 5 years of data): Admissions and other data as NAPLEX and MPJE performance determinants. Poster session presented at the American Association Colleges of Pharmacy Annual

Meeting. Retrieved from https://archive.hshsl.umaryland.edu/bitstream/handle/10713/7161/2016%20ACCP%20Poster%20-%20NAPLEX%20PREDICTORS.pdf?sequence=1.

Streiner, D. L. (2003). Starting at the beginning: An introduction to coefficient alpha and internal consistency. *Journal of Personality Assessment*, *80*(1), 99–103.

Tejada, F. R., Parmar, J. R., Purnell, M., & Lang, L. A. (2016) Admissions criteria as predictors of academic performance in a three-year pharmacy program at a historically black institution. *American Journal of Pharmaceutical Education*, *80*(1), Article 6.

Thomas, M. C., & Draugalis, J. R. (2002). Utility of the Pharmacy College Admission Test (PCAT): Implications for admissions committees. *American Journal of Pharmaceutical Education*, *66*(1), 47–50.

White, R. L., & Hall, P. (2006). Relationships between undergraduate institution ranking and academic performance in a doctor of pharmacy program [Abstract]. *ACCP [American College of Clinical Pharmacy] 2006 Annual Meeting*, E63, Abstract #80.

Wu-Pong, S., Windridge, G., & Osborne, D. (1997). Evaluation of pharmacy school applicants whose first language is not English. *American Journal of Pharmaceutical Education*, *61*, 61–66.