REVIEW

Understanding and Interpreting Pharmacy College Admission Test Scores

Don Meagher, EdD
NCS Pearson, Inc., San Antonio, Texas

Submitted September 24, 2015; accepted October 23, 2015; published February 25, 2017.

To fairly and accurately interpret candidates’ Pharmacy College Admission Test (PCAT) scores as listed on their official transcripts, it is important to understand how these scores reflect candidates’ performances on cognitive tasks involving the identification, interpretation, analysis, and evaluation of information assumed to have been covered in pre-pharmacy science, math, and general education coursework. This paper attempts to facilitate this understanding by explaining how candidates’ responses to PCAT test items relate to their scaled scores and percentile ranks and how their writing scores reflect their performance. This paper also suggests how differences between candidates’ PCAT subtest scores may reflect different personal experiences, educational backgrounds, and cognitive abilities.

Keywords: admission, assessment, PCAT

INTRODUCTION

The PCAT is a norm-referenced standardized test developed by NCS Pearson, Inc, to measure content knowledge and cognitive abilities that pharmacy schools consider to be essential for success in their programs. In an effort to ensure the continuing relevance and usefulness of the test for assessing candidates’ prerequisite knowledge and skills, the American Association of Colleges of Pharmacy PCAT Advisory Committee works closely with Pearson to review annual score data and to suggest periodic revisions to the contents of the test. These efforts are intended to ensure that the PCAT accurately reflects current pharmacy school prerequisite requirements in the sciences, math, and language arts, and that candidates’ scores accurately suggest their abilities in these areas.

Immediately upon completing the PCAT at a Pearson Vue test center, candidates are given a preliminary score report. After Pearson verifies that no irregularities occurred that could have affected candidates’ performances, and after candidates’ essays have been scored, each candidate receives an official score report, and each school designated by the candidate receives the candidate’s scores either on a printed official transcript directly from Pearson or electronically from the Pharmacy College Application Service (PharmCAS).

Each official score report and official transcript lists the following information for a candidate’s most recent PCAT test results: scaled scores and percentile ranks for the current four multiple-choice subtests: biological processes, chemical processes, critical reading, and quantitative reasoning; a composite scaled score, indicating an average of the four multiple-choice subtest scaled scores, and a composite percentile rank; a score for the writing subtest and a mean writing score representing the average score earned by all candidates taking the test during the 12 months prior to the candidate’s test date. The official transcript also lists the candidates’ four most recent previous PCAT scores earned during the past 5 years.

The PCAT scaled scores are standardized scores that represent equal units on a continuous scale, ranging from 200 to 600, with a designated mean and standard deviation (originally established in 2004 as 400 and 25, respectively). These scaled scores are derived from candidates’ raw scores, which are the number of items answered correctly for a given subtest. Unlike scaled scores, the percentile ranks are comparable across subtests because they are based on performances relative to the current norm group, which is the 64,652 candidates taking the PCAT for the first time from July 2011 through January 2015.

Because they represent equated scores, scaled scores earned for a given subtest during one PCAT test administration are comparable to scaled scores earned for the same subtest during other test administrations, even though different forms of the test are administered. Unlike scaled scores, the percentile ranks are comparable across subtests because they are based on performances relative to the current norm group. For this reason, the percentile

Corresponding Author: Don Meagher, 19500 Bulverde Road, San Antonio, TX 78259. Tel: 210-339-5297. E-mail: don.meagher@pearson.com
ranks are more useful than the scaled scores for comparing individual candidates to one another and for determining a candidate’s strengths and weaknesses relative to other candidates. The subtest percentile ranks are most useful for comparing abilities in the specific subject areas assessed by the PCAT, and the composite percentile rank is most useful for general comparisons.

All multiple-choice scores are reported on official transcripts as scaled scores and 2015 percentile ranks, with all percentile ranks obtained before July 2016 converted to 2015 equivalents. For all PCAT scores earned prior to July 2016, composite scaled scores have been recalculated based on the four subtests included on the PCAT, rather than the five in use when candidates originally took the examination. Additionally, composite percentile ranks for scores earned prior to July 2016 are based on these recalculated scaled scores. The PCAT percentile ranks listed on transcripts issued prior to the July 2016 PCAT administration may be compared to the current percentile ranks by using the compendium tables included in the 2016 edition of the PCAT Technical Manual.1

Each raw-score point on a test form subtest corresponds to a unique scaled score which, in turn, corresponds to a percentile rank. Because each PCAT multiple-choice subtest includes 40 operational items (items on a test form that are used to determine candidates’ scores), and there are 401 possible scaled-score points (200-600), a difference of one raw-score point necessarily results in a difference of more than one scaled-score point. This does not mean, however, that an increase of one raw-score point (one more item answered correctly) necessarily results in an equal unit increase in scaled-score points (eg, 10 scaled-score points for each raw-score point).

Each test form subtest is composed of a unique set of operational items that vary in difficulty, with no two items requiring the exact same estimate of ability for candidates to answer them correctly. For this reason, a raw-score point total (1-40) earned on one test form subtest reflects a slightly different estimate of the ability required for a candidate to answer the same number of operational items correctly on another test form. It follows, then, that the scaled-score increase associated with each increase in raw-score point will also vary.

A scaled score is assumed to indicate the ability needed to answer a given number of operational items correctly for a given subtest, and the corresponding percentile rank indicates how a given subtest scaled score relates to all candidates from the current normative sample. Table 1 shows average raw-score to scaled-score to percentile-rank correspondences for all test form subtests administered during the 2011-2015 PCAT normative sample period (except for verbal ability, which is no longer part of the PCAT). This table illustrates how 1-point increases in raw score have resulted in corresponding increases in scaled-score points and percentile ranks. Each row in Table 1 shows how a 1-point raw-score increase corresponds to a given average increase in scaled score and percentile rank during the normative sample period. For example, for biological processes, an increase from a raw score of 23 to 24 (shown as “23-24” in the table) corresponds to an average 3-point scaled-score increase and a 5-point increase in percentile rank. Note that the correspondences shown in Table 1 are averages that illustrate patterns observed across multiple test forms administered over several years rather than exact score correspondences for a specific test administration.

Table 1 illustrates what is typically seen in a subtest raw-score to scaled-score relationship: a 1-point difference in the raw score near the lower and upper ends (ie, 1-5 and 35-40) results in greater scaled-score differences than a 1-point raw-score difference near the median where corresponding differences in scaled scores are less. This occurs because there are more candidates earning scores close to the median, which results in greater scaled-score precision (with smaller standard error of measurement), and fewer candidates earning scores at the low and high ends where there is less scaled-score precision.

Table 1 also shows that the pattern in the relationship between scaled-score and percentile-rank correspondences is the opposite of the pattern seen between raw scores and scaled scores. The scaled-score to percentile-rank relationships illustrated in Table 1 reveal that around the subtest means, relatively slight differences in scaled scores result in greater differences in percentile rank. The reason for this is that scaled scores are derived from an assumption that the candidates taking the PCAT constitute an approximately normal distribution of abilities. The large size of the normative sample used to determine the percentile ranks has indeed resulted in the data forming a classic bell curve characteristic of a normal distribution, with the majority of scores clustered around the mean. This cluster of scaled scores around the mean is illustrated in Figure 1 by showing the bell curve formed when these scores are plotted for the current normative sample.

As shown in Figure 1, the percent of candidates earning a given scaled score (indicated on the y-axis) is greatest near the middle of the scaled-score distribution (ie, around 400 on the x-axis). This illustrates why relatively small differences in scaled scores near the mean correspond to greater differences in percentile rank than do...
scaled scores near the extremities, where there are fewer candidates and where each difference in scaled score results in a smaller difference in percentile rank. For example, as illustrated in Figure 1, a composite scaled score of 405 includes approximately 14% more candidates scoring below that point than does a score of 399 (a 6-point increase), whereas a scaled score of 447 only includes less than 1% more candidates scoring below that point than does a score of 441 (also a 6-point increase).

Writing Scores

The PCAT writing score reflects candidates’ abilities to apply conventions of language and problem-solving skills in the composition of an original essay. Even though
no equating method is applied to writing scores to make them precisely equivalent, as is done with the multiple-choice subtest scaled scores, they are obtained through rigorous standardized administration and scoring processes that ensure their reliability. All operational writing prompts (topics on which candidates are asked to write) are selected based on criteria requiring that a similar mean score was obtained when each prompt was originally field tested.

Each PCAT essay response is assessed with reference to scoring rubrics (detailed descriptions of each score point), and specific verification procedures are followed during the scoring process to ensure the consistency and reliability of the scores assigned. Because these scores are reported for performances on specific prompts that differ from one test administration to another, and because each test administration includes multiple operational prompts, comparisons between candidates’ writing scores must be made with some caution. That said, the writing scores may be interpreted with reference to the mean score listed on the official transcript, as well as to the score-point descriptions and the sample essays of each score that are included in a document entitled “Interpreting PCAT Scores,” which is available to download from the PCAT website.

Candidates taking the PCAT earn a single writing score reported on a scale of 1.0-6.0, with 1.0 representing the lowest earned score possible and 6.0 representing the highest earned score possible. Following a test administration, a score for each essay is determined by assigning two scores ranging from 1-6. If the two assigned scores are the same (eg, 3 and 3), or differ by no more than one score point (eg, 3 and 4), the two scores are averaged, resulting in a final score represented to one decimal place (eg, 3.0, 3.5). If the two assigned scores differ by more than one score point (eg, 3 and 5), a resolution score is assigned and then combined with the higher of the two original scores, with the average of these two scores representing the final score (eg, an original high score of 5 and a resolution score of 4 would result in a final score of 4.5).
Table 2 shows a score point distribution for all the valid writing scores earned (i.e., no 0 scores) on the current 6-point scale during the current normative sample period. Most of the scores (78.5%) are in the middle of the distribution (3.0-4.0), with a relatively smaller proportion (18.7%) at the lower end (1.0-2.5) and an even smaller proportion (2.8%) at the upper end (4.5-6.0). What this distribution suggests is that approximately 82% of the candidates earn scores of 3.0 or greater, 57% earn scores of 3.5 or greater, 39% earn scores of 4.0 or greater, and 3% earn scores of 4.5 or greater. This information may help schools determine a minimally acceptable writing performance or may suggest a candidate’s need for remediation after admission.

Interpreting PCAT Scores for Admission
Pharmacy College Admission Test scores can be useful in the admissions process and as a guide for placing students appropriately by providing consistent and reliable information about the abilities of applicants. This is particularly important considering the possible variability in standards among schools and geographic regions. Scores on the PCAT may also be useful in identifying a candidate’s academic strengths and weaknesses and assessing any need for remedial support. In the process of reviewing PCAT scores, along with all other available information about a candidate, admission committees may find some discrepancies that require clarification.

PCAT Subtest Score Discrepancies
Although many items on the biological processes and chemical processes subtests do involve the application of interpretive and critical-thinking skills, particularly those associated with passages (paragraphs of text on a specific topic, which may include graphics), other items may also include specific technical content that is more easily memorized. The critical reading, quantitative reasoning, and writing subtests are more skill-based than knowledge-based: one requires the comprehension, analysis, and evaluation of complex ideas represented in text; one involves the performance of quantitative computation skills required to solve mathematical problems; and one involves the composition of an original essay that proposes a solution to a problem. These skills are not as dependent on identifying information as they are on the ability to interpret information and perform applications. For this reason, candidates with higher scores on the biological processes and chemical processes subtests relative to their scores on critical reading, quantitative reasoning, and writing may be those whose cognitive strengths are more knowledge-based than application-based.

Looked at in another way, the biological processes, chemical processes, and quantitative reasoning subtests emphasize basic scientific and mathematic principles and knowledge, and in that sense may be much more course-dependent than the critical reading and writing subtests, which involve abilities that can be developed over time in both academic and personal settings. For this reason, the language-oriented subtest scores may provide greater insight into candidates’ general academic ability than the more course-dependent biological processes, chemical processes, and quantitative reasoning subtest scores. Also, some candidates whose native language is not English may have difficulty with critical reading items that require the interpretation and analysis of extended text, or with the writing task that requires the composition of an original essay. Even though such individuals may be able to deal effectively with the scientific and mathematical material as reflected in prerequisite course grades or in performance on the PCAT biological processes, chemical processes, and quantitative reasoning subtests, they may be less able to deal with the more language-dependent critical reading and writing subtests.

CONCLUSION
Though research has consistently shown the PCAT to be a strong predictor of subsequent academic performance, PCAT scores are intended to be only one set of indicators among several sources of candidate information considered in admissions decisions. For this reason, PCAT scores should not be seen to represent a rigid performance standard that could inadvertently exclude otherwise worthy candidates. What PCAT scores do represent is a common standard against which candidates can be compared, regardless of educational background or other personal variables. When used along with all other information available to admissions committees, PCAT scores represent valuable information that each college or school
of pharmacy can determine for itself how to best use when assessing its applicants.

ACKNOWLEDGMENTS

The author thanks the members of the AACP PCAT Advisory Committee for suggesting this primer and reviewing an early draft. The author also thanks Tianshu Pan, Rachel Wegner, and Weslea Miller, colleagues from Pearson (the publisher of the PCAT) who helped with the table and figure contents and by reviewing subsequent manuscript drafts. The ideas expressed in this manuscript are those of the author, an employee of NCS Pearson, and are not meant to represent either the AACP or the PCAT Advisory Committee.

REFERENCES

1. Pearson. PCAT Technical Manual. Bloomington, MN: NCS Pearson, Inc; 2016.
2. Pearson. Interpreting PCAT scores. PCAT, NCS Pearson, Inc. http://pcatweb.info. Accessed May 22, 2015.