

The use of computer based tests (CBT) and psychometric analysis in Doctor of Pharmacy Programs to assess exam reliability and consistency

Jonathan Hernández Agosto, EdD jonathan.hernandez12@upr.edu University of Puerto Rico, School of Pharmacy

Introduction

- Doctor of Pharmacy Programs (DOPs) has been using and analyzing assessment data to improve the quality of their programs, as it is also required by the Accreditation Council for Pharmacy Education (ACPE) to demonstrate the effective preparation of student pharmacists and the extent of the quality requirements of the pharmacy education.^{1,2}
- Computerized based tests (CBT) are now an essential tool for DOPs to fulfil those requirements, by promoting a more effective learning through an assessment-delivery model that examine skills, knowledge, and understanding.³
- CBT allows for on-time electronic data collection and psychometric data analysis, by computing useful and essential statistical measures that provide exam writers and administrators with an industry-standard set of data to validate exam reliability, consistency, and quality.⁴

Objectives

This study main objectives were:

- Assess CBT psychometric data to determine University of Puerto Rico DOP exams and items reliability and consistency.
- Describe the relationship between exam reliability and consistency with course progression.

Methods

- Design:** longitudinal study.
- Sample:** 141 CBT, 5,655 items from several exams, and 17 courses (see Table1).
- Data timeframe:** three consecutive academic years (2015-16; 2016-17; 2017-18).
- Data collection:** data was collected from the electronic assessment-management system used by faculty and students for CBT and reporting. Data cleaning procedures were performed to assure data integrity and inclusion criteria were met.
- Psychometric analysis:** Kuder-Richardson Formula 20 (KR-20), Item Difficulty Index (p-value) and Point Bi-serial Correlation Coefficient were analyzed for the sample.
- Statistical analysis:** descriptive statistics and Pearson correlation analysis were performed to examine reliability and consistency of CBT over time on course progression.

Results and Discussion

- Data showed that 86 out of 141 CBT (61%) achieved an internal consistency and reliability index (KR-20) over the 0.60, commonly used as industry-standard⁴ (see Figure 1). With small sample sizes, course exams with a KR-20 score higher than 0.60 should be considered consistent and reliable, while it is recommended to maintain scores higher than 0.70.⁴
- Regression analysis for the 141 CBT showed no significant relationship ($r=0.077$, $p=0.36$, $\alpha=0.05$) on KR-20 scores over time (see Figure 2). However, about 70% of individual courses did see an improvement in KR-20 scores over time, although not statistically significant.
- Point Bi-serial Correlation Coefficient showed that 2,942 out of 5,655 items (52%) achieved, at least, the recommended 0.20, with a median of 0.21.⁵
- Students' average performance in overall assessments was 80.7% (see Table1 and Figure 3).
- Multiple-choice items' difficulty median was 0.85, while the recommended is 0.60. Also, binary items' (e.g., true or false) difficulty median was 0.90 while the recommended is 0.75 (see Figure 4).⁶

Table 1

Course	Number of assessments	KR-20 median	Punctuation average	Average answer time (s)
[FARM7225] Integrated Pharmaceutical Science and Therapeutic Agents I	17	0.81	81.1	120
[FARM7228] Integrated Pharmaceutical Sciences of Anti- Infective Agents	10	0.76	80.5	99
[FARM7226] Integrated Pharmaceutical Science and Therapeutic Agents II: Medicinal Chemistry and Pharmacology	8	0.75	84.8	120
[FARM7207] Scientific Foundations for the Professional Practice: Pathophysiology	12	0.75	79.5	127
[FARM7206] Scientific Foundations for the Professional Practice: Biochemistry	14	0.73	79.9	120
[FARM7331] Integrated Sciences, Therapeutics and Patient Care I	17	0.72	73.9	128
[FARM7307] Psychosocial Basis, Culture and Management Theory-Practice Seminar V	4	0.66	84.8	110
[FARM7332] Integrated Sciences, Therapeutics and Patient Care II	17	0.65	77.1	120
[FARM7135] Research, Education and Scientific Method Laboratory I	2	0.51	84.0	120
[FARM7137] Compounding and Manufacturing of Dosage Forms I	8	0.48	81.3	94
[FARM7227] Pharmacy and the Health Care System	6	0.43	83.2	123
[FARM7306] Psychosocial Basis, Culture and Management Theory- Practice Seminar IV	5	0.41	85.2	128
[FARM7208] Principles of Drug Discovery and Drug Development	4	0.41	80.8	105
[FARM7166] Scientific Foundations for the Professional Practice: Mathematics, Chemistry and Physics	4	0.40	72.0	120
[FARM7237] Compounding and Manufacturing of Dosage Forms II	6	0.40	90.5	115
[FARM7116] Health Promotion and Disease Prevention	2	0.31	87.0	90
[FARM7229] Basic Biopharmaceutics and Pharmacokinetics	5	0.30	92.4	UTD
Total	141	0.66	80.7	117.7

Figure 1

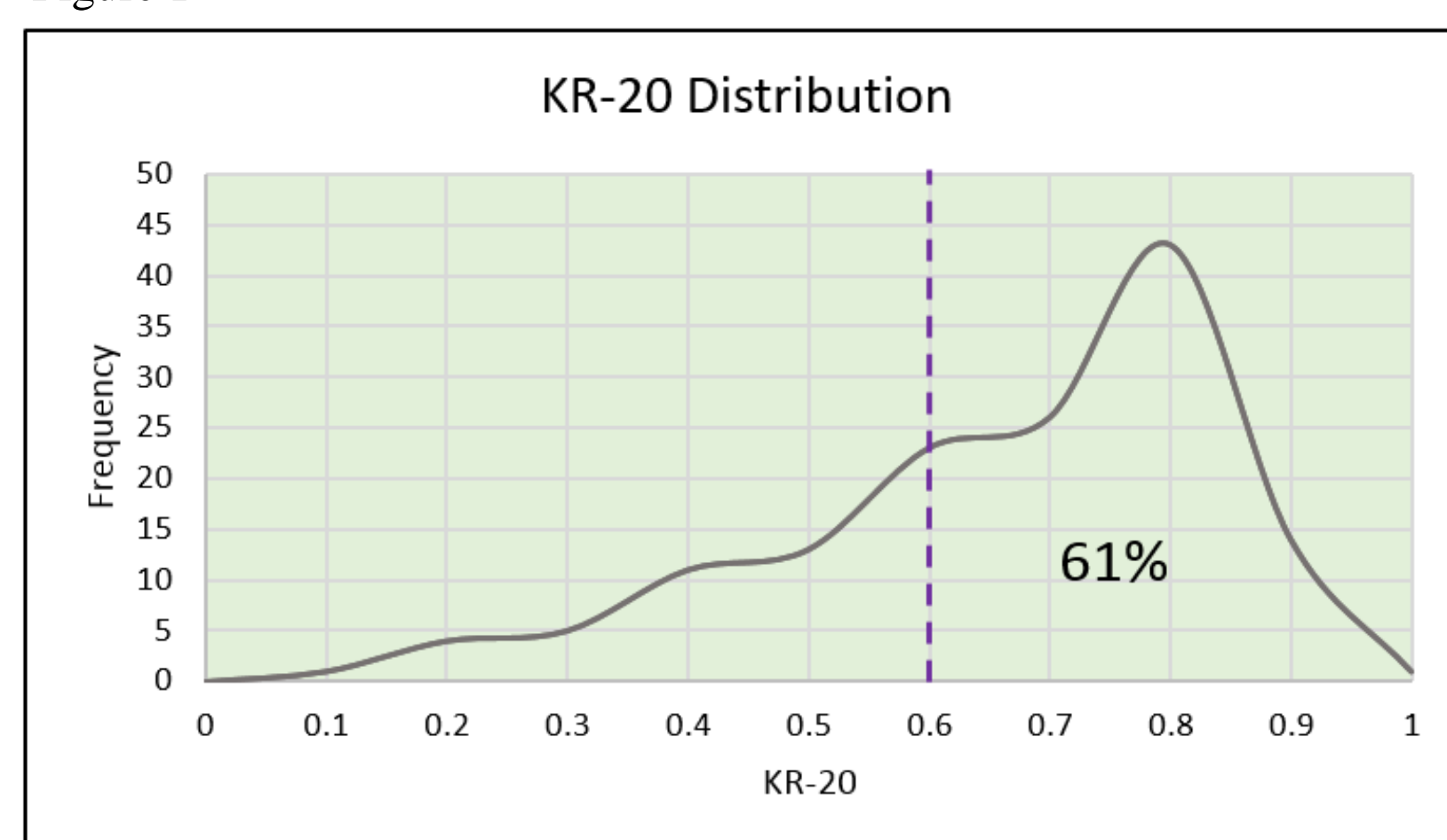


Figure 2

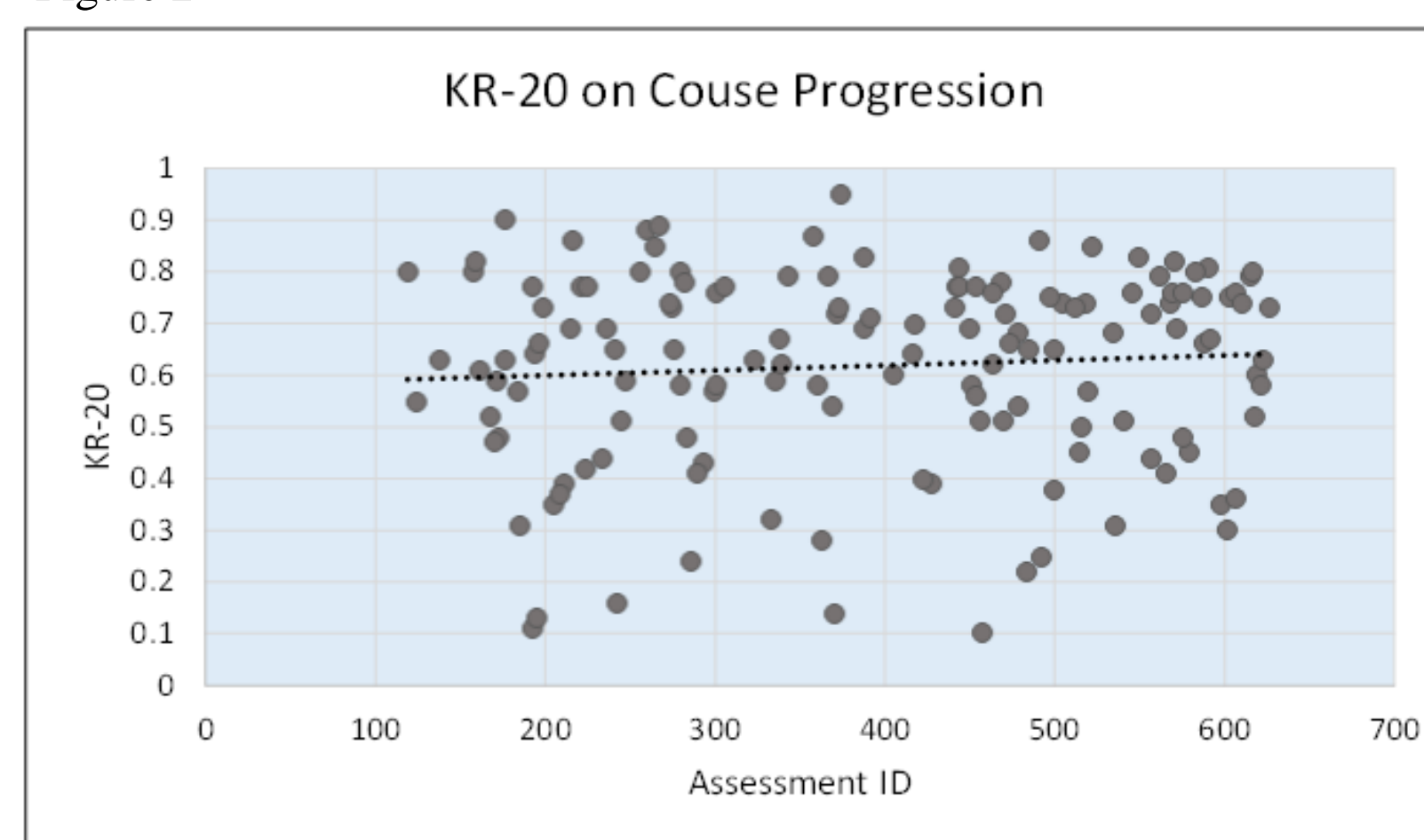


Figure 3

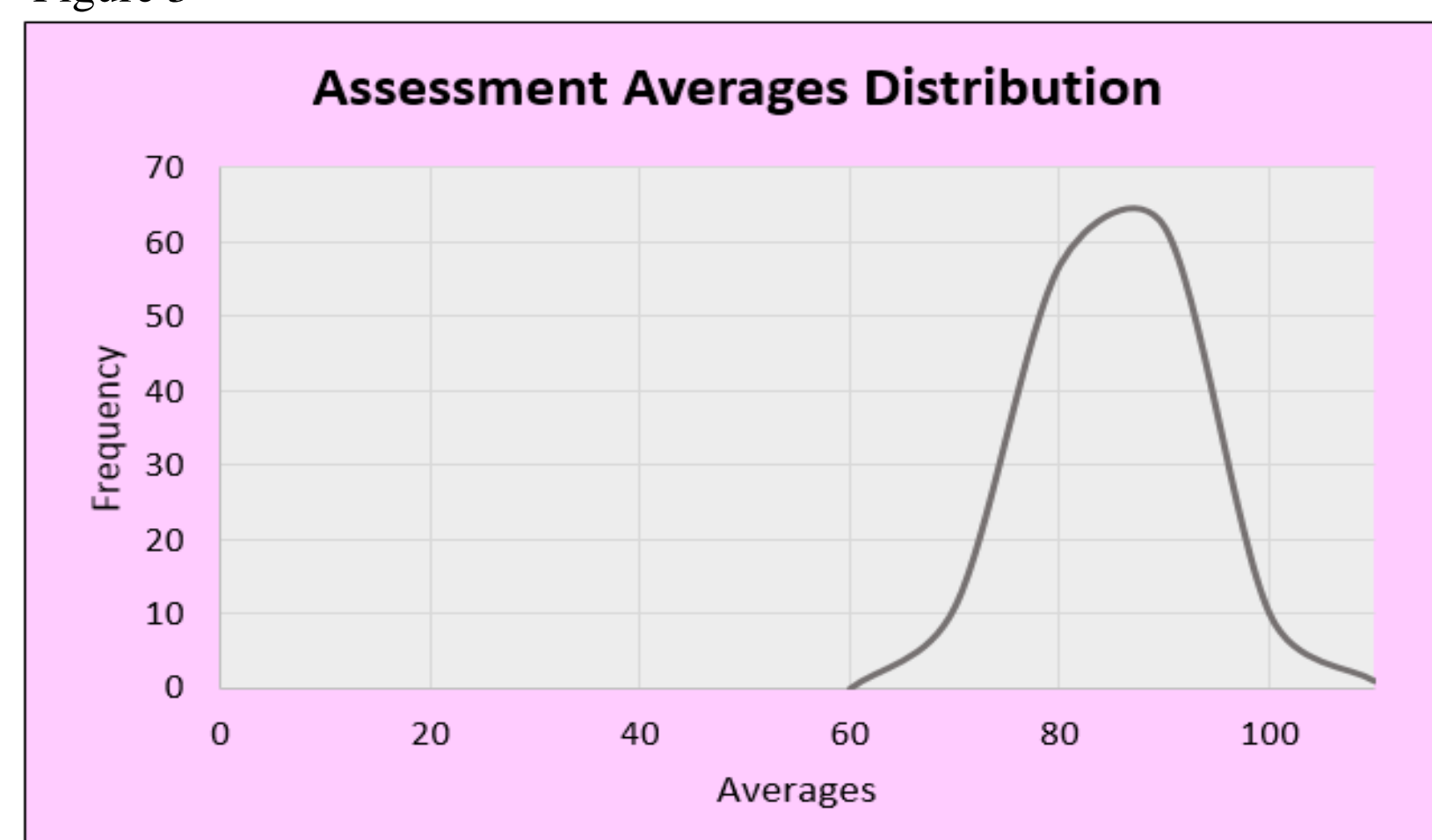
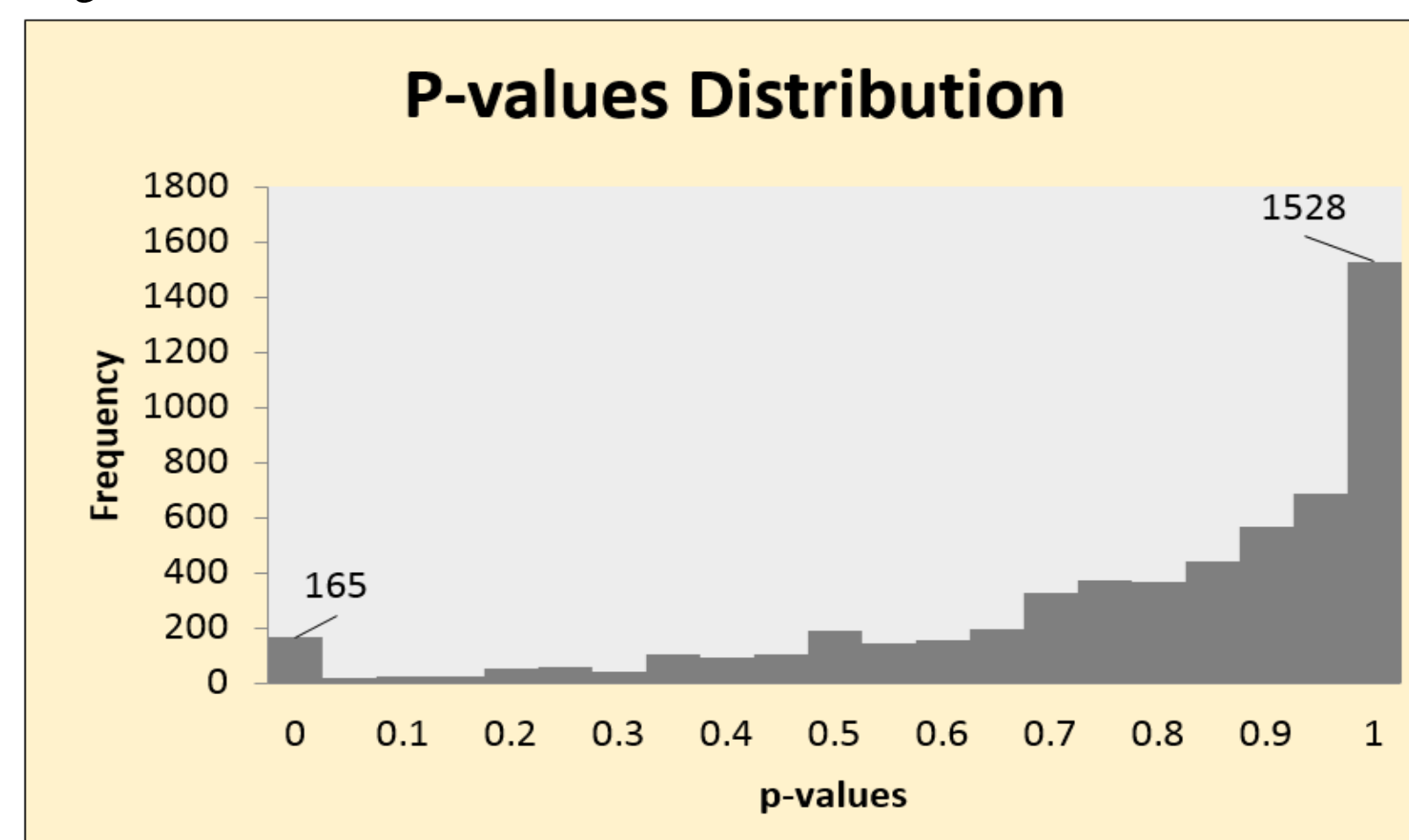


Figure 4



Conclusion

This study shows that DOP exams reliability and consistency scores frequently surpassed desired KR-20 scores of 0.60, and item correlations of 0.20. Deeper analysis is needed on items' p-values, since they were higher than recommended by literature. Future efforts will continue to build upon CBT implementation to further improve DOP assessments.

References

- Jimoh, R.G., Kehinde, A.J. & Kawu, Y.K. (2012). Students' perception of Computer Based Test (CBT) for examining undergraduate chemistry courses. *Journal of Emerging Trends in Computing and Information Sciences*, 3(2), 12-134. http://www.academia.edu/2343231/Students_Perception_Of_Computer_Based_Test_CBT_For_Examining_Undergraduate_Chemistry_Courses
- Accreditation Council for Pharmacy Education. (2015). *Accreditation standards and key elements for the professional program in pharmacy leading to the doctor of pharmacy degree*. Retrieved from <https://www.acpe-accredit.org/pdf/Standards2016FINAL.pdf>
- Jamil, M., Tariq, R.H. & Shami, P.A. (2012). Computer-based (CB) vs. paper-based (PB) examinations: perceptions of university teachers. *The Turkish Online Journal of Educational Technology*, 11(4), 371-381. <http://www.tojet.net/articles/v11i4/11437.pdf>
- Sparks, J.P. (2015). Exam Quality Through Use of Psychometric Analysis—A Primer. Retrieved from <http://learn.examssoft.com/resources/white-paper/psychometrics>
- Allen and Jarvi (2015). *When to throw away your assessment, and when to throw away the reliability statistic* [Power Point slides]. AACP & AFPC Annual Meeting 2015.
- Measurement and Evaluation Center, The University of Texas. (2003). *Test Item Analysis & Decision Making*. Retrieved from <http://www.hct.ac.ac/content/uploads/MC-item-analysis-handout.pdf>

Financial support and/or acknowledgments

University of Puerto Rico, School of Pharmacy