American Journal of Pharmaceutical Education **2015; 79 (10) Article 152.**

RESEARCH

Creating and Assessing Student Perception of an Examination Mastery Score Report for a Pharmacotherapy Course

Greg L Alston, PharmD, Donald S. Nuzum, PharmD, Angela Pegram, PharmD, John Brock Harris, PharmD

Wingate University School of Pharmacy, Wingate, North Carolina

Submitted November 14, 2014; accepted March 8, 2015; published December 25, 2015.

Objective. To compare student-perceived utility of two types of score reports.

Methods. Assessment standards were followed to create a new examination score report for pharmacotherapy coursework. Student examination scores were returned using the traditional score report, the utility of which students rated along 9 dimensions. A mastery score report was also distributed, and students rated it on the same 9 dimensions. The ratings were compared to determine which the students perceived as more useful.

Results. The students rated the mastery score report significantly better on each of the 9 dimensions and in aggregate.

Conclusion. Pharmacy students perceived the mastery score report as more useful in helping them improve their achievement of educational outcomes.

Keywords: classroom assessment, score report, outcomes assessment, CAPE 2013

INTRODUCTION

The new Center for the Advancement of Pharmacy Education (CAPE) 2013 Outcome[1] statements are deeply embedded in the Accreditation Council for Pharmacy Education (ACPE) Standards 2016.[2] The demand for assessment data that is evidenced-based, designed using sound principles, and properly validated will grow exponentially when these new standards are implemented, as schools of pharmacy scramble to document assessment outcome data for their students. A critical but frequently overlooked element in the validation process of any assessment instrument is reporting the results. If the data user, whether a student, instructor, or institution, fails to interpret the data correctly, the assessment validity will be compromised and could easily create unintended negative consequences.

Institutions of higher education typically train faculty members on creating technically sound assessments for classroom instruction. The assumption is that well-made assessments can then be used to gauge student performance in the classroom. However, little attention is given to score reporting and the ways in which the reported results of classroom assessments are organized, reported, and used.[3]

Corresponding Author: Greg L. Alston, Wingate University School of Pharmacy, 515 N. Main St., Wingate, NC 28174. Tel: 704-233-8329. E-mail: galston@wingate.edu.

The bulk of the literature on score reporting is in the K-12 education space in response to the No Child Left Behind Act of 2001. In 1980, Congress established the US Department of Education as a cabinet-level agency. The department's official mission is: "to promote student achievement and preparation for global competitiveness by fostering educational excellence and ensuring equal access." The department's jurisdiction ranges from preschool education to postdoctoral research.[4] It is possible that Department of Education standards and rules that apply to K-12 score report quality will eventually be applied in some form to college students. Ultimately, the fairness and validity of any assessment to accurately reflect student performance are critical issues. In Title I of the No Child left Behind Act, the legislation states that score reports should be "consistent with the relevant, nationally recognized professional and technical standards."[5]

Further guidance on how to construct a high quality score report comes from the Standards for Educational and Psychological Testing[6] (SEPT), the Code of Fair Testing Practices in Education (CFTPE)[7] and the Code of Professional Responsibilities in Educational Measurement (CPREM).[8] Standards from SEPT relevant to reporting student level results on assessments can be summarized as follows: only qualified individuals should interpret test scores for high stakes decisions, and administration is responsible for confirming

those qualifications; a major impact decision based on a on a single test score is inappropriate; score reports need to clearly state the degree of measurement error; and the relationship between test scores and student outcomes should be provided when making promotional decisions on this information.[6] Standards from CFTPE relevant to reporting in interpreting test results include advising test developers to report test results accurately and provide information to help test users interpret the scores correctly. The standards advise test users to, in turn, report and interpret test results accurately and clearly. In the typical classroom environment, the instructor is both the test developer and the test user.[7] These standards may be difficult to implement consistently with one party in both roles. Thus, different types of assessment instruments (other than just test scores) would be helpful for interpretation of test results.

A final foundational document is the CPREM. Published in 1995, it summarized the professional responsibilities of those who interpret, use, and communicate assessment results. This code placed the responsibility for accurate score reporting and test interpretation on those who design and administer tests with significant consequential validity for test takers.[8] These individuals are normally responsible for the final high-stakes decisions on student assessments and, as such, may carry the burden of potentially life-changing decisions for many families. Thus, their responsibilities, which can be vital to a sound assessment process, are delineated in this document. While the bulk of these guidelines were crafted to protect young students in the early stages of their academic career, the principles are appropriate for pharmacy and professional students as well because of the personal cost of failure they may experience.

Academic research on score reporting and test interpretation is modest compared to other assessment topics. The research that does exists presents a fairly consistent picture of the ineffectiveness of score reports to communicate meaningful information to stakeholders. Student examination score reports serve an instructional purpose to inform students and their advisor about the students' learning and progress towards competency within the curriculum.[9]

In their 2004 review of test score report literature, Goodman and Hambleton suggested general recommendations for score reporting:[3] (1) score reports should be clear, concise, and visually attractive; (2) score reports should include easy-to-read text that supports and improves the interpretation of charts and tables; (3) care should be taken to not try to do too much with data display; (4) devices such as boxes and graphics should be used to highlight main findings; (5) data should be grouped in meaningful ways; (6) small font footnotes and statistical jargon should be avoided; (7) key terms should be defined, preferably within a glossary; (8) reports should be piloted with members of the intended audience; (9) consideration should be given to the creation of specially designed reports that cater to the particular needs of different users. While most available research on score reporting dealt with the K-12 progress tests, Goodman and Hambleton asserted that, "many findings and principles that have emerged from this research are relevant to student-level score reports."[3]

The increasing call for sound assessment practices in pharmacy education demands that we develop evidence-based assessment practices, including the development of sound score reports. Peer-reviewed standards can provide guidance. Pharmacy educators may need to ask if they are doing everything within their power to adhere to explicit and consistent standards in the classroom assessment practices.

Armed with the background information presented, we wanted to address several concerns regarding current assessment practices in the Wingate University doctor of pharmacy (PharmD) program. First, we examined whether test review policy minimizes the formative value of failure. Multiple-choice testing represents a significant portion of the grading rubric for many courses in PharmD degree programs. Typically, the student's score report consists of a percentage-correct score.

In the Wingate University culture, individual test item performance is not discussed with students after examinations. One reason is to maintain the integrity of course question banks, which allows an instructor to reuse test items and improve examination quality over time. Anecdotal results suggest that most students pursuing this information attempt to negotiate for extra points and not necessarily identify their areas of weakness. Doing so would also disadvantage students at the distance campus who may have less opportunity to interact directly with the professor. Restricting access to test item review, however, leaves some students in need of more specific, formative guidance to understand what items they missed so they can improve their performance.

Second, we looked at the inter-rater reliability of course examinations. While each instructor received training in test item writing, there was significant variation in the weight of knowledge, application, and problem-solving questions on examinations between all classes, making it difficult for students and faculty advisors to interpret test scores from different courses.

Third, we noted that students' learning behavior centered on point gathering rather than skill mastery.

Traditional percentage score reports may have been encouraging students to seek points rather than competency.

The fourth issue was that lack of longitudinal progress measures limited the formative effectiveness of examinations. The Wingate curriculum consists of a pharmacotherapy track with a sequence of courses presented in 5-week blocks over 4 semesters. Existing assessment practices provided no ability to longitudinally assess student skill development in pharmacotherapy over the span of four semesters or between courses.

Finally, the lack of quality data for faculty advisors seemed to limit their effectiveness. Students who failed a course examination were required to meet with their faculty advisors, but the advisors were not fully aware of the content or assessment methods in every course. Therefore, advisors were not always able to give targeted guidance to a student at risk of failure.

The overall consequence associated with pharmacotherapy course examination testing remains constant: fail or pass. Traditional percentage score reports provide a single piece of information related to the global, intended consequence. A formative score report, developed using themethodology described in this article, provides a more detailed analysis of what went wrong or right, allowing students to focus on improving performance. Student formative score reports provide information during courses, which gives students the opportunity to immediately modify approaches in order to avoid certain consequences for the remainder of the course.[10, 11]

Formative assessment is more effective than summative assessment in improving academic performance.[12] Formative score reporting addresses consequential validity by providing students more information directly related to both intended and unintended consequences associated with testing. A traditional percentage report conveys a student's performance in a particular course—the intended consequence. However, a percentage report does not identify high or low performing domains, which may unintentionally imply student mastery of the entire subject.

A formative score report provides information on high and low performing subject domains or concepts as well as the intended consequence of whether the student is passing the course.[13] Therefore, improving the formative value of the score report may enhance educational performance by increasing the amount of information provided to each student about their performance.

The goal of this process, therefore, was to develop a formative score report that could address these 5 challenges and be administered in a cost-effective manner. Knowing that not all of the concerns could be addressed at one time, we looked at student perception of the new score report's usefulness as an indicator of whether the project was worth the time and effort to pursue further.

METHODS

Wingate University School of Pharmacy is a private, not-for-profit, 8-semester PharmD program housed on the main campus for Wingate University. The school enrolls 90 students on the Wingate campus and 18 students on its Hendersonville campus. Most courses are delivered via synchronous video technology between locations. Approximately 60% of entering students have at least a bachelor's degree but exceptionally performing students are accepted with a minimum of 2 years of prerequisite coursework. The first class graduated in May 2007, and the North American Pharmacist Licensure Exam (NAPLEX) board pass rate for first-time test takers exceeded the national average for the first 7 years of the program (96.8% vs 95.5%). The school is a student centered, teaching (not research) focused model. The main campus is not colocated with a medical center but the satellite campus in Hendersonville will be housed at a medical center beginning in 2016.

Most examinations at the school are administered using LXR test software (Logic eXtension Resources, Georgetown, SC). The tests can be scored using score sheets with a scanner or, more commonly, are delivered online. The scores are mechanically calculated, and a variety of score reports can be output. All instructors are educated on the the software to construct tests and evaluate item statistics reports to judge the quality of test items. Course leaders are advised to produce a test with the reliability score above 0.6 and to drop questions from the scoring if the questions are of poor quality. Most course leaders build their own question bank of multiple-choice test items to create their midterm and final examinations. One of the features of the software is customization of the score report. In 2008, the assistant dean for assessment developed a custom score report for the annual skills mastery assessment examination that was well received by students and faculty members.[14]

The primary objective of this research group was to create a student score report that would enhance the ability of students to detect weaknesses in achieving the educational outcomes of the curriculum. Potential secondary objectives for this new style of score report would be to use an aggregate mastery summary report to provide institutional effectiveness data, course improvement data, and the motivation for students to redirect their study efforts. This required the group to match the educational outcomes guidance provided by CAPE 2013 with the instructional objectives for their courses

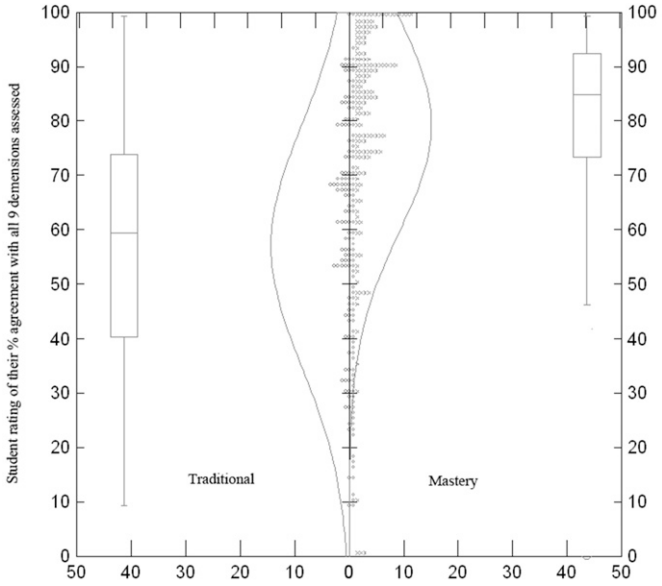

Figure 1. Comparison of student ratings of the traditional vs mastery score report for all 9 dimensions combined (N=150).

to synchronize instruction, assessment, and educational outcomes in a way that was previously not possible.

Based on Goodman and Hambleton's suggested recommendations for score reporting,[3] a customized grading table report was piloted to measure skill development in a pharmacy management course. The methodology for developing this skill-based assessment strategy was described in a previous article.[15] After a successful pilot, a group of clinical faculty was approached to consider using this style of grading report in the pharmacotherapy modules. Using the ACPE 2007 Standards Appendix B[16] as a model, the faculty group developed a pilot grading table report and tested it in an endocrinology pharmacotherapy course in spring 2014. Adjustments were made after this pilot to revise the outcomes reporting and survey procedures to reflect the new domains of the CAPE 2013 outcomes.

The process followed these steps: (1) identify the appropriate outcomes measures to be reported; (2) use the database features of the LXR software to code each test item in the question bank to reflect the outcome assessed; (3) assemble a course examination to provide at least 8 test items from each of the 4 domains being assessed; (4) administer the examination and score the tests; (5) run the new grading table reports; (6) return the individual traditional percentage correct score report to the students on their examination results; (7) survey the students on the utility of the traditional percentage score report (Appendix 1); (8) distribute the individual new mastery style color-coded report (Figure 2); (9) survey the students on the utility of the new style score report; (10) analyze the data to determine if the new score report was more useful than the traditional method.

Survey questions were written to address 9 dimensions for the 2 score reports: (1) accurate reflection of performance; (2) accurate reflection of effort; (3) identification of content areas that need improvement; (4) sufficient specificity to allow reallocation of study time; (5) ability to motivation desire to change study habits;

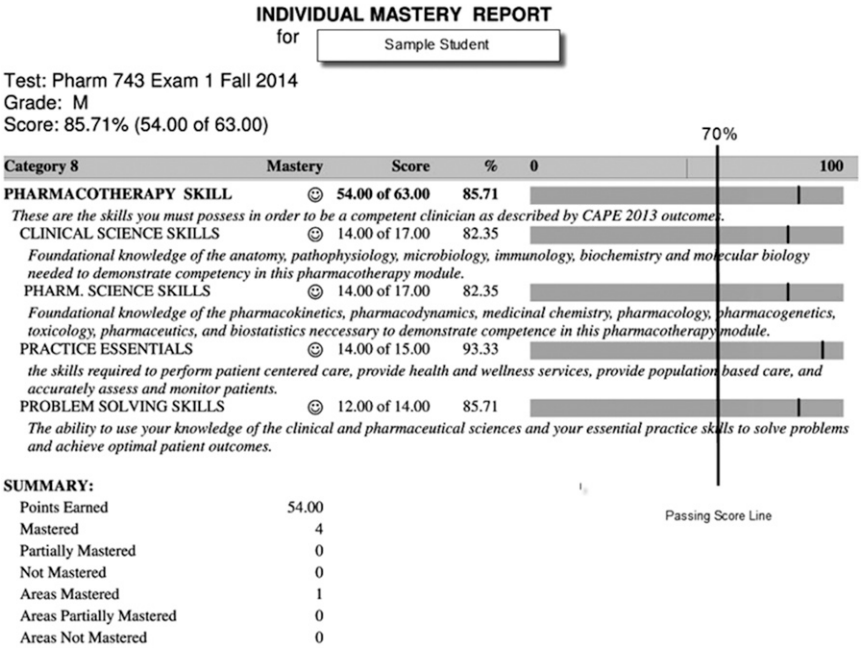

Figure 2. Individual Student Mastery Score Report.

(6) ability to identify weakness in problem-solving skills; (7) ability to gauge progress towards developing problem-solving skills; (8) ability to gauge readiness to perform on practice experiences; (9) ability to redirect learning focus from grades to mastery of competency. An affirmative statement was crafted for each dimension such as: "The individual mastery score report provides an accurate reflection of my performance on this examination." Students estimated their percentage of agreement with each statement from 0-100%. An estimate of zero was assigned a rank of 1, an estimate of 1-10% was assigned a rank of 2 and so on in 10% increments so that the minimum score for any rating was a 1 and the maximum score was 11.

The surveys were administered using SurveyMonkey (Survey Monkey, Palo Alto, CA). The second-year spring 2014 endocrine course and the second-year fall 2014 cardiology course were used to test students' reaction to the score reports. One hundred sixty students responded to the traditional score report survey, and 150 responded to the mastery score report survey. Survey results were exported to spreadsheet software and tracked by a student so that a paired *t* test could be performed to determine significance. Results were analyzed using SYSTAT, v13 (SYSTAT, San Jose, CA) and reported graphically using the paired 2-sample *t* test graph to further examine the data. A composite score of all 9 dimensions was computed for the aggregate student response for both reports and subjected to the same data analysis.

The individual student score mastery reports were printed and distributed to students 3 days after they took their examination. The mastery summary report aggregates this data to provide the instructor with a global view of their course effectiveness based on the exam scores (Figure 3). This project was reviewed and approved by the Wingate University research review board.

RESULTS

The detailed results for the 9 dimensions are reported in Table 1. The mastery score report was rated significantly higher for all 9 dimensions ($p< 0.01$). The range of improvement was from a low of 0.64 rating points for accurate reflection of performance to a high of 4.73 rating points for identification of areas of improvement. The composite scores were analyzed using the 2 sample *t* test and graph and the results appear in Figure 1. The student ratings based on their percentage of agreement with the survey statements indicate that they found the mastery score report significantly (2.66 rating points) more useful than the traditional score report ($p<0.01$).

DISCUSSION

The literature on assessment suggests score reports play an integral role in helping test users understand the meaning of their test scores.[17,18] While standardized examinations traditionally spend time and effort on perfecting their score reports, most classroom assessment has not. Students progress or fail to progress largely based on their performance on course examinations. Students and faculty members are at risk of misinterpreting results of examinations if the examinations are not designed according to a rigorous blueprint or if the scores are reported without the score placed in proper context. The role of the score report is to strengthen the validity of decisions based on examination score by making sure all users of the test understand what the results mean. Without a proper score report, the examination will likely not help the student understand what to do better next time.

To meet the CAPE 2013 outcomes goal of student development of competency, schools of pharmacy may need to develop assessment methods that can accurately inform students, educators, and institutions on the progress towards these goals. This project demonstrates that a score report centered on the CAPE domains is perceived

MASTERY SUMMARY REPORT
for Pharm 743 Exam 1 Fall 2014 from ScoresFall2014

# M	Number Mastering		% M	% Mastering			PTS	Possible Points		
# PM	Number Partial Mastering		% PM	% Partial Mastering			HIGH	Highest Points		
# NM	Number Not Mastering		% NM	% Not Mastering			LOW	Lowest Points		
							AVG	Average Points		
Category 8	**# M**	**# PM**	**# NM**	**% M**	**% PM**	**% NM**	**PTS**	**HIGH**	**LOW**	**AVG**
PHARMACOTHERAPY SKILL	67	9	19	70.53	9.47	20.00	63.00	60.00	26.00	47.24
CLINICAL SCIENCE SKILLS	59	11	25	62.11	11.58	26.32	17.00	17.00	6.00	12.88
PHARM. SCIENCE SKILLS	54	14	27	56.84	14.74	28.42	17.00	17.00	6.00	12.61
PRACTICE ESSENTIALS	64	0	31	67.37	0.00	32.63	15.00	15.00	5.00	11.22
PROBLEM SOLVING SKILLS	52	17	26	54.74	17.89	27.37	14.00	14.00	4.00	10.53
SUMMARY:										
All Category 8s	25	-	6	26.32	-	6.32	63.00	60.00	26.00	47.24
All Areas	67	-	19	70.53	-	20.00	63.00	60.00	26.00	47.24

Figure 3. Mastery Summary Report for All Test Takers.

Table 1. Comparison of Student Perceptions of the Utility of Traditional vs Mastery Style Score Reports for a Midterm Examination in a Pharmacotherapy Course

Survey Dimension	Traditional % Correct Score report[1]	Mastery Report by CAPE Competency[1]	P Value of Paired *t* test	Effect Size[2]
Accurate reflection of performance	8.44	9.08	<.01	0.64
Accurate reflection of effort	7.07	7.75	<.01	0.68
Identification of content areas that need improvement	4.95	9.68	<.01	4.73
Sufficient specificity to allow reallocation of study time	5.63	9.28	<.01	3.65
Sufficient to motivate desire to change study habits	7.02	9.10	<.01	2.08
Sufficient to identify weakness in problem solving skills	5.08	9.34	<.01	4.26
Sufficient to gauge progress towards developing problem-solving skills	5.94	9.16	<.01	3.22
Sufficient to gauge readiness to perform on practice experiences	6.35	8.84	<.01	2.49
Sufficient to redirect learning focus from grades to mastery of competency	6.83	9.00	<.01	2.17
Average for all dimensions	6.37	9.03	<.01	2.66

[1]Survey scores were based on an 11-point scale from 0-10. A rating of 10 would indicate complete agreement with the statement. A rating of 0 would indicate complete disagreement with the statement

[2]Effect size was computed by subtracting the traditional score report value from the mastery score report value

by students to be substantially better than a traditional percentage-correct score report at: identifying content areas that need improvement; providing sufficient specificity to allow reallocation of study time;, motivating students to change study habits; identifying weaknesses in problem-solving skills; gauging progress towards developing problem-solving skills; gauging readiness to perform on practice experiences; and redirecting learning focus from grades to mastery of competency.

This study was small-scale (N=150) and subject to several potential biases, including students' prior experience with a mastery-based score report for their annual assessment examination and the novelty factor of anything new being better than anything old. However, the perceived utility of the new mastery score report compared to the traditional percent-correct score report is compelling. The new report can provide critical formative assessment students may be able to use to improve their performance on future examinations and in future courses. Basing the report on CAPE outcome domains links student performance directly to institutional and curricular performance goals. In addition, the mastery summary report allows instructors to quickly identify the success of their instruction in achieving educational outcomes. This style of reporting requires faculty members to code their test item bank appropriately to tag each test item with the appropriate CAPE domain, but the score report itself adds no more than 10 minutes to the score reporting process. The key value of this approach to score reporting is that it converts a summative assessment examination in to a formative assessment tool, which strengthens performance better than summative assessment.[11]

Further work is being done to determine if the perceived ability to alter student study habits or learning styles actually occurs and whether any improvement persists across courses or over the length of the didactic coursework. A weakness in this study is that all test bank item tagging to the reported domains has not yet been peer reviewed to validate the selections of each course instructor. Finally, this study did not review the perceived benefit of faculty members using the score report to improve instructional outcomes, improve student advising, or confirm whether the quality of the course examinations improved. Those areas will be explored in future studies using the new score report format.

CONCLUSION

Pharmacy students in this program perceived a mastery style score report as more useful than a traditional percentage-based score report in helping them identify content areas that need improvement, reallocating their study time, motivating them to change study habits,

identifying competence in problem-solving skills, gauging progress towards developing problem-solving skills, gauge readiness to perform on practice experiences, and redirect their learning focus from grade achievement to mastery of competency. The score report is relatively easy to create, administer, and return to students in a timely manner. Pharmacy schools seeking to document outcomes and provide valid classroom assessment could easily adapt this mastery score report strategy.

REFERENCES

1. Medina MS, Plaza CM, Stowe CD, et al. Center for the Advancement of Pharmacy Education 2013 Educational Outcomes. *Am J Pharm Educ.* 2013;77(8):Article 162.
2. Accreditation Council for Pharmacy Education. *Standards 2016 Accreditation Standards*. https://www.acpe-accredit.org/pdf/Standards2016FINAL.pdf Accessed Dec 17, 2015
3. Goodman DP and Hambleton RK. Student test score reports and interpretive guides: Review of current practices and suggestion for future research. *App Meas Educ*. 2004; 17(2):145-220.
4. The United States Department of Education. *The Federal Role in Education, 2012*. http://www2.ed.gov/about/overview/fed/role.html. Accessed February 23, 2015.
5. No Child Left Behind Act of 2001, Title I-Improving the Academic Achievement of the Disadvantaged, US Department of Education. *Fed Regist* 2002. 67(231): 45038.
6. American Educational Research Association, American Psychological Association, National Council on Measurement in Education. *Standards for Educational Psychological Testing*. 5th ed. Washington, DC: AERA Publications; 2014.
7. Joint Committee on Testing Practices. *Code of Testing Practices in Education 2004*. http://www.apa.org/science/programs/testing/fair-testing.pdf. Accessed November 10, 2014.
8. National Council on Measurement in Education. *Code of Professional Responsibilities in Educational Measurement 1995*. http://ncme.org/resource-center/code-of-professional-responsibilities-in-educational-measurement/. Accessed November 10, 2014.
9. Ryan JM. Practice issues and trends in student score reporting. In: Downing SM and Haladyna TM, eds. *Handbook of Test Development*. Mahwah, NJ; Lawrence Erlbaum Associates, Inc.; 2006: 677-710.
10. Suskie L. *Assessing Student Learning; A Common Sense Guide*. Bolton, MA: Anker Publishing Company, Inc.; 2004.
11. Black P and Wiliam D. Developing the theory of formative assessment. *Educ Asse Eval Acc*. 2009;21:5-31.
12. Black P and Wiliam D. Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan.* 1998;80(2): 139-149.
13. Haladyna TM. Roles and importance of validity studies in test development. In: Downing SM and Haladyna TM, eds. *Handbook of Test Development*. Mahwah, NJ; Lawrence Erlbaum Associates, Inc.; 2006:739-755.
14. Alston GL and Love BL. Development of a reliable, valid annual skills mastery assessment examination. *Am J Pharm Educ.* 2010;74(5): Article 80.
15. Alston GL and Griffiths CL. The methodology for assessing skill-based educational outcomes in a pharmacy course. *Am J Pharm Educ.* 2015:Article 4940 accepted for publication.
16. Accreditation Council for Pharmacy Education. *Final 2007 ACPE Accreditation Standards*. https://www.acpe-accredit.org/pdf/FinalS2007Guidelines2.0.pdf. Accessed November 10, 2014.
17. Zenisky AL and Hambleton RK. Developing test score reports that work: the process and best practices for effective communication. *Educational Measurement: Issues and Practice.* 2012; 31(2):21-26.
18. MacIver R, Anderson N, Costa AC and Evers A. Validity of interpretation: a use validity perspective beyond the test score. *International Journal of Selection and Assessment.* 2014; 22(2): 149-164.

Appendix 1- Survey Questions Used to Compare Score Reports

1. What is your survey ID Number?
2. The individual mastery score report provides an accurate reflection of my performance on this examination.
3. The individual mastery score report provides an accurate reflection of my effort in preparing for this examination.
4. The individual mastery score report provides me with sufficient information to identify the content areas in which I need to improve.
5. The individual mastery score report provides me with sufficient information to allow me to reallocate my study time for future pharmacotherapy examinations.
6. The individual mastery score report provides me with sufficient motivation to actually reallocate my study efforts for future pharmacotherapy examinations.
7. The individual mastery score report provides me with sufficient information to identify the clinical problem-solving skills I need to improve.
8. The individual mastery score report provides me with sufficient information to accurately gauge my progress in developing clinical problem-solving skills.
9. The individual mastery score report motivates me to refocus my learning strategy from obtaining a good grade to mastering my clinical problem-solving skills.
10. I consent to the use of my de-identified answers for publication in a study. I understand that results will be reported in aggregate and no personally identifiable information will be divulged.