INSTRUCTIONAL DESIGN AND ASSESSMENT

The Development, Implementation, And Evaluation Of A Self-Assessment Instrument For Use In A Pharmacy Student Competition

Terrence R. Jackson, PharmD, MSc, and Nicholas G. Popovich, PhD
College of Pharmacy, University of Illinois-Chicago

Objectives. The purpose of this research was to develop, implement, and evaluate a 10-item assessment instrument for pharmacy students participating in a local Academy of Managed Care Pharmacy, Pharmacy and Therapeutics Committee competition. The instrument was designed to measure and document student perceived-learning processes and outcomes.

Methods. The study used a single group posttest design with a retrospective pretest component. The self-assessment instrument was comprised of 10 assessment items and 5 open-ended questions. The Rasch rating scale model was used to evaluate data provided by the instrument.

Results. The perceived outcomes from the competition in terms of student growth through participation and completion of the assessment instrument were favorable for all items. However, 3 of the 6 items administered in a retrospective pretest-posttest format became more difficult for students to endorse.

Conclusions. The Competition Student Assessment Instrument provided valuable, meaningful, and reliable information about processes and learning outcomes of the competition. Student-driven competitions represent an innovation in applying student-directed activities outside the curriculum that is consistent with expected learning outcomes of the formal curriculum. It is anticipated that this assessment instrument would be useful for assessing student competitions at other schools and colleges of pharmacy.

Keywords: Self-assessment, Rasch model, retrospective pretest posttest

INTRODUCTION

What the student learns outside of the formal classroom setting plays a central role in shaping the knowledge foundation, attitudes, and identity of the student.1 2 When the outside curriculum is linked to adult learning theory (eg, “adult learners are driven by the desire for practical knowledge”), the potential for self-directed personal growth that promotes professional competence becomes substantial.3 5 Thus, students who elect to participate in professionally related activities outside the classroom will directly benefit from their experience and progress toward the attainment of expected learning outcomes germane to the formal pharmacy curriculum.

One of the advantages that students in pharmacy have is the abundance of opportunities that facilitate professional growth and development, rooted in the practical application of knowledge. Each year, hundreds of pharmacy students self-direct their learning through active participation in competitions such as the American Pharmaceutical Association Academy of Students of Pharmacy (APhA-ASP) Patient Counseling Competition (established in 1985), American Society of Health Systems Pharmacists (ASHP) Clinical Skills Competition (established in 1995), and the Academy of Managed Care Pharmacy (AMCP) Pharmacy and Therapeutics (P&T) Committee Competition (established in 1997). The content of these competitions appears to provide students the opportunity for learning and applying pharmacy knowledge that is consistent with expected learning outcomes for the doctor of pharmacy program (Appendix 1). These competitions provide opportunities to reinforce the development of skills (eg, problem solv-
ing, clinical decision making, oral and written communication, and interpersonal skills) and the application of knowledge in a “real world” environment. It follows then that the development of critical thinking skills (ie, cycle of concrete experiences, reflection, abstract conceptualization, the application of new insights) is also facilitated by these processes. To date, however, the learning outcomes of these valuable informal curricular experiences have not been measured objectively and published.

This manuscript reports the development, implementation, and evaluation of the pilot administration of a 10-item assessment instrument completed by students from the University of Illinois at Chicago College of Pharmacy. These students participated in the 2002 AMCP Pharmacy and Therapeutics Committee Competition. The assessment instrument was developed and piloted to: (1) objectively measure and document student-perceived learning processes and outcomes, (2) effectively and efficiently provide reliable and valid information about what the students gained from the competition, (3) function as a tool for continuously improving the quality of the experience for future competitions, and (4) be adaptable for use in colleges/schools of pharmacy for similar student competitions.

METHODS

Study Design

The study used a single group posttest design with a retrospective pretest component. Items relating to perceived agreement of outcomes from the student competition were measured using a posttest design (ie, “As a result of this competition...”). Because of the challenge in evaluating what students perceive they have gained from such competition experiences when the internal construct may change as a result of the experience (ie, response-shift bias), 6 items were designed to be administered in a retrospective pretest-posttest design.

Measurement Model

The Rasch rating scale model was selected to evaluate the data provided by the instrument because it provided objective evidence that all the items measure the same construct (ie, have unidimensionality) and produced additivity of measures, allowing for subsequent statistical analyses that required interval level data. This mathematical model converts ordinal raw scores to interval level measures, if the data fit the model, on a logit measurement scale. The rating scale model describes the probability that a person (n) will respond to each Likert-type response category (x) for each item (i). This model is appropriate for estimating person self-assessment and item endorsability for responses scored in the 4 ordered categories used in this study. The equation for this probability contains 3 parameters: a self-assessment measure for each person, the item’s endorsability, and a set of step-measure parameters (Figure 1).

The model provided measures for the response categories on the attitude continuum and supported respondent’s perceptions regarding the hypothesized and required sequential ordering of category meanings for the group as a whole. From a measurement perspective, the scale locations or category thresholds (ie, weak, fair, good, very good) of the rating scale must be ordered and additive for interpretations to be meaningful. This means that the scale threshold value of one category (eg, agree) must be greater than the next adjacent category (eg, somewhat agree) and so on throughout the scale. Only when the relationship in scale function can be demonstrated is there evidence to support the validity that the rating scale is functioning in a manner required for ordered response categories. To create meaningful measures using rating scale or Likert-type questionnaire items and to evaluate model fit requirements (unidimensionality and local independence), several criteria are recommended (Appendix 2). Thus, when rating scale function, meaningfulness item hierarchy, and model fit have been established, a variable of measurement has been constructed. Rasch modeling techniques also provide information to evaluate individual item contributions not discerned by standard techniques such as Cronbach Alpha, KR20, point-biserial correlations, or exploratory or confirmatory factor analysis. Thus, when the data fit the model, the Rasch model provided strong inference that the measured behaviors were consistent with the underlying construct. Another advantage of using Rasch analysis is that as few as 30 subjects may be needed to obtain sufficient useful and reasonable estimates for detecting differences. Once a meaningful scale, unidimensionality, and model fit had been established, 2 parameters were used to evaluate student responses in the model: (1) student perceived ability or endorsement, and (2) difficulty of item endorsement as perceived by the student.

The Rasch rating scale models report values in logits similar to the dichotomous model. For a pharmacy-oriented primer on Rasch analysis see Jackson et al. However, in the Rasch rating scale model, step threshold categories are added because the response to the items is polychotomous (ie, threshold measures are
Symbol | Indicates
---|---
Beta subscript n | measure for each person
Delta subscript i | the item’s endorsability
Tau subscript j | set of step-measure parameters

Figure 1. Rasch rating scale probability equation

\[
P(X_{ni} = x) = \frac{\exp \sum_{j=0}^{\infty} \left[ \beta_n - (\delta_i + \tau_j) \right]}{\sum_{x=0}^{m} \exp \sum_{j=0}^{\infty} \left[ \beta_n - (\delta_i + \tau_j) \right]}, x = 0, 1, \ldots, m
\]

where \( P(X_{ni} = x) \) is the probability that a person \( n \) is observed in the rating scale category \( x \) on item \( i \), which has \( m + 1 \) rating scale categories, and \( \sum_{j=0}^{\infty} \left[ \beta_n - (\delta_i + \tau_j) \right] = 0 \).

the natural log odds for choosing one category versus the adjacent category). The non-Rasch practitioner need not understand the technical intricacies of the model to find it useful. This manuscript simply uses the model to evaluate the outcomes of the pilot instrument relating to outcomes of a pharmacy student competition.

Assessment Instrument Development

This portion of the project consisted of 2 components. First, the student competition evaluation instrument was developed. Then, the instrument was administered and the scored results were analyzed using Rasch modeling. A demographic information questionnaire was also designed and administered. The instrument and questionnaire format were adapted from components of instruments developed by the authors that demonstrated the collection of meaningful data from consumers receiving pharmaceutical care services and medical students evaluating the impact of an educational intervention relating to the aging adult.

Subjects

The subjects used in this study were students enrolled in the doctor of pharmacy degree program at the University of Illinois at Chicago during the academic year 2001-2002 and who competed in the AMCP Pharmacy and Therapeutics Committee Competition held in March 2002. The University of Illinois-Chicago Office for the Protection of Research Subjects granted approval and exemption status for this research project. Students were informed that participation in completing the assessment instrument and demographic questionnaire was anonymous and voluntary. The student competition instrument and demographics questionnaire were administered to 35 students participating in the Pharmacy & Therapeutics Committee competition.

Development of the Competition Evaluation Instrument

Specific items for the evaluation instrument were developed by an ad hoc committee consisting of one faculty member, a graduate student from the department of pharmacy administration, and 3 doctor of pharmacy program students all from the University of Illinois-Chicago College of Pharmacy. The 3 students were members of the local AMCP chapter, who were not participating in the competition. This committee met and deliberated 5 times on its mission to develop and refine the assessment instrument.

The Competition Student Assessment Instrument consisted of: (1) 4 learning-outcomes related items using a 4-point rating scale (ie, disagree, tend to disagree, tend to agree, and agree), (2) 6 items administered in a retrospective pre/post format using a 4-point rating scale (ie, weak, fair, good, very good), and (3) 5 open-ended items relating to competition processes (Appendix 3).

Statistical Analysis

Data (containing rating scale responses to the 10 items of the competition instrument) were entered into a data file with MS-DOS and then input into Winsteps, Version 3.37 (Mesa Press, Chicago, Ill) to calculate statistics for the Rasch rating scale model. The output generated from Winsteps included separation reliability (item separation index and item reliability, person reliability), category functioning indices (ie, average category measures and category thresholds), item INFIT and OUTFIT statistics, and the item expected score map. SPSS statistical analysis system version 11.0.1 for
Table 1. Sample Demographics (N=32)∗

<table>
<thead>
<tr>
<th>Variable</th>
<th>Number (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class standing in the doctor of pharmacy program</td>
<td></td>
</tr>
<tr>
<td>First-year students (class of 2005)</td>
<td>8</td>
</tr>
<tr>
<td>Second-year students (class of 2004)</td>
<td>9</td>
</tr>
<tr>
<td>Third-year students (class of 2003)</td>
<td>11</td>
</tr>
<tr>
<td>Fourth-year students (class of 2002)</td>
<td>4</td>
</tr>
<tr>
<td>Age of student participating</td>
<td></td>
</tr>
<tr>
<td>Range (y)</td>
<td>21-30</td>
</tr>
<tr>
<td>Mean (y)</td>
<td>24.4 ± 2.3</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>15</td>
</tr>
<tr>
<td>Male</td>
<td>17</td>
</tr>
<tr>
<td>Ethnicity</td>
<td></td>
</tr>
<tr>
<td>African American</td>
<td>1 (3.1)</td>
</tr>
<tr>
<td>American Indian or Alaska Native</td>
<td>1 (3.1)</td>
</tr>
<tr>
<td>White</td>
<td>11 (34.4)</td>
</tr>
<tr>
<td>Asian</td>
<td>19 (59.4)</td>
</tr>
<tr>
<td>Number of times students had competed in the competition</td>
<td></td>
</tr>
<tr>
<td>First time competitors</td>
<td>20 (62.5)</td>
</tr>
<tr>
<td>Second time competitors</td>
<td>7 (21.9)</td>
</tr>
<tr>
<td>Third time competitors</td>
<td>2 (6.3)</td>
</tr>
<tr>
<td>Missing data</td>
<td>3 (9.4)</td>
</tr>
<tr>
<td>Preparation time</td>
<td></td>
</tr>
<tr>
<td>Range (hrs)</td>
<td>2-55</td>
</tr>
<tr>
<td>Mean (hrs)</td>
<td>30.19 ±14.2</td>
</tr>
</tbody>
</table>

*except where noted as missing data

Windows (SPSS Inc, Chicago, Ill) was used to calculate t-tests for evaluating retrospective pretest and posttest competition evaluation data.

RESULTS

The student competition instrument and demographics questionnaire were completed by 32 of the 35 students (90%). Demographic results for the sample completing the questionnaire are displayed in Table 1. Additionally, 25 of the 32 (78%) students had participated in “other managed care experiences” (eg, AMCP rotations, conferences, course electives, and organizational meetings) and were members of one or more of 13 different professional pharmacy organizations. This suggests that participants had active interests in but not limited to AMCP. Fifty-three percent (17/32) of the students participating in the student competition reported holding leadership positions in professional organizations at the college of pharmacy (eg, president, vice president, secretary, treasurer, committee chair, program coordinator).

Rasch Rating Scale Analysis

Table 2 depicts the summary of the rating scale thresholds. The number of observations in each category
Table 2. Summary of Rating Scale Measured Steps for Evaluation of Rating Scale Function

<table>
<thead>
<tr>
<th>Category</th>
<th>INFIT MNSQ</th>
<th>OUTFIT MNSQ</th>
<th>Category Thresholds</th>
<th>Average Category Measure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weak/Disagree = 1</td>
<td>0.87</td>
<td>0.87</td>
<td>none</td>
<td>-3.50</td>
</tr>
<tr>
<td>Fair/Tend to Disagree = 2</td>
<td>1.03</td>
<td>1.07</td>
<td>-2.30</td>
<td>-1.36</td>
</tr>
<tr>
<td>Good/Tend to Agree = 3</td>
<td>0.95</td>
<td>0.96</td>
<td>-0.37</td>
<td>1.18</td>
</tr>
<tr>
<td>Very Good/Agree = 4</td>
<td>1.00</td>
<td>1.00</td>
<td>2.47</td>
<td>3.81</td>
</tr>
</tbody>
</table>

was greater than 10, average category measures increased with the rating scale categories, MNSQ statistics for the measured steps were between 0.6 and 1.4, category thresholds increased with the rating scale categories, and category thresholds were at least 1.4 logits apart and no more than 5 logits apart. Additionally, the shape each rating scale distribution was peaked (Figure 2). Thus, verifying the rating scale was functioning as intended and the criteria noted (listed under Rasch Rating Scale Diagnostic Criteria in Appendix 2) had been met.

Model Fit and Unidimensionality

Evaluation of INFIT and OUTFIT statistics for the items in the student competition evaluation instrument showed that MNSQ values were less than 1.40 and greater than 0.6 and supported the evaluation criteria for the rating scale model. This indicated that there were no unexpected responses for items near ability levels or extreme outliers by any student and that the INFIT and OUTFIT statistics exhibited good fit and supported the required unidimensionality and local independence requirements of the model.

The separation index (ie, the extent to which items are sufficiently spread out to define distinct levels of ability) for the 10-item competition evaluation instrument was 3.0, and item reliability (ie, the estimate of reproducibility of item placement within the hierarchy of difficulty across students of differing abilities) was 0.90. These two results indicated that the items created a variable that was well spread out and that the reliability of item placement along the scale was acceptable. The student separation index was 2.0 and reliability was 0.80, supporting the reliability of the student responses. These reliability values were analogous to Cronbach alpha and indicated that items separated the student measures, enabling the identification of individual differences.

Model Fit and Unidimensionality

Analysis of data demonstrated that the rating scale structure was time dependent (ie, that lower categories were rarely used in the posttest data). To measure persons and items encompassing both time measurements in the same clearly defined frame of reference (ie, shared origin and scale), student responses to the 10 items in the posttest were anchored on combined pre/post step calibrations, enabling comparisons between and pretest and posttest responses.
The item and person INFIT and OUTFIT statistics continued to meet model evaluation criterion. The results of analysis for the 10 items and the student distribution are presented with the expected score map in Figure 3. The information in this map represents what can be expected by each person and item interaction (ie, person endorsement for each item) in the data fit model. The expected score map provided for an interpretation for each item for each person measure. That is, one can observe what the expected response to each item is for a particular person measure. Responses for each item relative to the group mean and +/- 1 and 2 standard deviations may be interpreted by using the student distribution curve located at the top of the figure. For example, the distribution for item 7 (“As a result of this competition, my ability to blend knowledge with skills in the practice setting improved”) demonstrated that the mean for the student distribution (ie, dashed line) intersected the item response set just to the right of the colon between TA and A. Thus, students from the “:” located to the left of the mean (M) and to the right of the distribution would have a higher probability of responding, “agree,” to this statement than the rest of the sample. Between the “:” located to the left of the mean (M) and to the left of the distribution would have a higher probability of responding, “tend to agree.” Similarly, the results for the other 9 items would be interpreted accordingly.

The right side of the expected score map lists the 10 items in order of their difficulty to endorse. The item at the top, item 5, “My oral communication skills,” was the most difficult item for the group to endorse followed by item 1, “My knowledge about formulary management.” The easiest item for students to endorse was item 7, “As a result of this competition, my ability to blend knowledge with skills in the practice setting improved.”
Table 3. Posttest and Retrospective Pretest Item Difficulty Measures (in Logits) and Results of the t-Test

<table>
<thead>
<tr>
<th>#</th>
<th>Item Description</th>
<th>Retrospective Pretest Measure (in Logits)</th>
<th>Posttest Measure (in Logits)</th>
<th>Difference Between (in Logits)</th>
<th>t-test</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Knowledge about formulary management</td>
<td>1.42</td>
<td>0.62</td>
<td>0.80</td>
<td>-1.98*</td>
</tr>
<tr>
<td>2</td>
<td>Knowledge about clinical decision making</td>
<td>0.91</td>
<td>0.54</td>
<td>0.37</td>
<td>-0.93</td>
</tr>
<tr>
<td>3</td>
<td>My ability to problem solve</td>
<td>-0.17</td>
<td>0.20</td>
<td>0.37</td>
<td>0.92</td>
</tr>
<tr>
<td>4</td>
<td>My ability to work with others in a team</td>
<td>-2.05</td>
<td>-0.79</td>
<td>1.26</td>
<td>2.92*</td>
</tr>
<tr>
<td>5</td>
<td>My oral communication skills</td>
<td>0.86</td>
<td>0.86</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>6</td>
<td>My self-confidence</td>
<td>-0.66</td>
<td>-0.14</td>
<td>0.52</td>
<td>1.29</td>
</tr>
</tbody>
</table>

* indicates statistical significance, alpha = 0.05

**DISCUSSION**

**Interpretations for Items 7-10**

Items 7 through 10 were designed to measure outcomes in a posttest-only administration and as such each item begins with “As a result of this competition…” As previously discussed, the normative distribution showed the probability that all students agreed or tended to agree with item 7, “As a result of this competition, my ability to blend knowledge with skills in the practice setting improved.” Similarly, the results for item 9, “As a result of this competition, my ability to self-assess my learning needs improved,” may be interpreted. The positive endorsement of these 2 items (which also represent UIC core curricular expected learning outcomes) documents objectively, student perceived competition outcomes showing improvement in an area that characterizes what was believed to be happening in the competition process. That is, the competition itself provided an opportunity for the application (blending) of knowledge with skills in a practice setting and also facilitated the student’s ability to self-assess during the process.

Item 10, “As a result of this competition, my interest in pursuing a career in managed care has increased,” was also interpreted as “agree” or “tending to agree.” This item represents content specific to managed care. Thus, while most students have had multiple exposures to this material, the competition continued to stimulate interest in students, which could potentially impact career decisions after graduation. It is important to note that the core of this item for the competition evaluation instrument would be modified from “managed care” to “clinical care” for the ASHP clinical skills competition and from “managed care” to “patient care” for the APhA-ASP patient counseling competition.

Item 8, “writing skills improved,” was positively endorsed (agree or tend to agree) one standard deviation to the left of the mean (“;” depicts category break point) and represents the probability of expecting those responses for about 90% of the student sample. Thus, approximately 10% would “tend to disagree” or “disagree.”

**Posttest Interpretations for Items 1-6**

Items 1 through 6 were designed to detect and quantitatively measure student-perceived improvement for specific knowledge or skills when administered in a retrospective pretest-posttest format. The area under the normative distribution curve in Figure 3 demonstrated that for items 1-6 most students (90%) would respond “good” or “very good” regarding knowledge about clinical decision making or formulary management, ability to problem solve or work with other students in a team setting, oral communication skills, or self-confidence. This information suggested that overall the perceived outcomes from the student after completing the competition were favorable. However, because this posttest format did not differentiate what students
gained as a result of the competition process, a retrospective pretest component was needed to discern this information.

**Retrospective Pretest, Items 1-6**

The retrospective pretest allowed the participants to think back in time in evaluating their perception of how they would quantify their knowledge or ability before the competition experience, anchored in a stable construct. Those item values that changed less than 0.3 logits from pretest to posttest were considered stable (invariant) in pretest-posttest item calibration measures. Five of the 6 item measures from the retrospective pretest exhibited change from pretest to posttest. To compare the item measures for items that varied using a conservative approach, the retrospective pretest was anchored on posttest item 5 (ie, oral communications skills), because its placement on the measurement scale remained stable in both data sets. The posttest responses were used as the benchmark in this retrospective application because this was the point of time reference in which the instrument was being completed. Item difficulty measures for the posttest and the retrospective pretest with their respective t-tests for differences are listed in Table 3. The greater the logit value, the more difficult the item was to endorse. Item 1 (ie, knowledge about the formulary management system) became significantly easier on a statistical basis to endorse as a result of the students exposure to the competition (1.42 to 0.62, t = 1.98, α = 0.05). Because the competition used for the piloting of this instrument was focused on formulary management and each team spent an average of 120 combined hours in preparing their presentations, it logically follows that knowledge of formulary management would improve.

Similarly, item 2 (ie, knowledge about clinical decision making) became easier for students to endorse as evidenced by the change in item calibration value (0.91 to 0.54 logits). This made sense because the competition involved the review and evaluation of a substantial body of literature (albeit contrived literature) needed to support clinical decision making throughout the preparation period. That is, because each group had to make a final decision about whether to recommend the medication for addition to the formulary, the use of clinical information was critical to the decision-making process. This was one of the “big benefits” of participating in the competition as verified by conversations with several students following the competition.

Interestingly, items 3, 4, and 6 became more difficult for students to endorse (ie, their perceived ability to work in a team setting and solve problems, and their level of self-confidence was greater before the competition experience). Item 4, “My ability to work in a team with others,” demonstrated the largest item value difference from retrospective pretest to posttest. The item became more difficult to endorse by 1.26 logits, which was also statistically significant (t = 2.92, α = 0.05). Intuitively, the resultant decrease in student endorsement of “my ability to work with others in a team” from retrospective pretest to posttest may reflect the consequence of “new-group” dynamics. Having little or no experience in resolving issues related to effective team functioning, students worked intensely as a team for 3 weeks. While pharmacy students occasionally work in teams on various projects throughout their 4 years of pharmacy education, skill development in group dynamics is not part of the formal curriculum. The interpretation for this item is further delineated in the discussion of responses to the open-ended items.

Likewise, self-confidence (item 6) was easier for students to endorse in the retrospective pretest. Given the nature of competition and not knowing the competition outcomes at the time the survey instrument was administered, this was not surprising because the students participating may not have been confident of their placement as “the winner.” Additionally, because the competition occurred as part of the outside curriculum and there was no formal feedback given regarding their performance, competing students had to rely on self-assessment to guide their performance. This was further reinforced by the finding that the majority of students competing were first time competitors who did not have the benefit of previous competitions to aid in guiding their performance. Thus, while students were gaining experience in self-assessing their work as they proceed through the competition, their self-confidence decreased.

The relatively small finding of decreased endorsement from retrospective pretest to posttest for item 3, “my ability to problem solve” (difference from pre to post of 0.37 logits) was surprising to the researchers. Because the competition was a large problem-solving exercise, the researchers expected to observe student-perceived improvement. In discussing the finding of this item with several of the students who competed in the competition, 2 issues were identified. The first related to how students defined “problem solving.” Problem solving, as interpreted by the students, also included solving problems associated with getting the group to perform as a team beyond just collating the contributions of 4 individual members. It appeared that many teams were still working in or had just moved beyond a “storming”
phase in the group process. This interpretation was consistent with the findings regarding item 4, “My ability to work with other students in a team.” The second issue related to the nature of the case design itself. That is, the competition case used was an artificial case with fictitious information and medication names. The contrived nature of the case prohibited the use of problem-solving tools with which students were familiar and that might be used to solve problems presented in the formal curriculum, ie, use of therapeutics textbooks and additional searches of medical literature via Medline or International Pharmacy Abstracts. This interpretation was also consistent with student responses from open-ended items.

Open-ended Items 11-15

Two major issues surfaced regarding the adequacy of resources (item 11). The first issue reflected the amount and type of information. Fifteen of 32 students (47%) described the need to include more detailed (and authentic) information critical to the decision-making and problem-solving process, ie, information that is “integral to formulary management.” Only 3 of the 32 (9%) students cited that the resources were adequate. Secondly, 11 of the 32 students (34%) described the need for more time. The timeline for the competition was spread out over 4 weeks.

Students’ comments regarding questions or concerns (item 12) were few and concerns of the group were most often resolved within the groups themselves. However, when questions were asked, the competition coordinator addressed them in a timely and effective manner. Regarding item 13, “What would you recommend to improve the competition in the future?” there were several suggestions for using “real” medications with sources of information other than “fictitious manufacturer claims.”

The student responses to items 14 and 15, “Describe the contributions of the team that worked well” and “...could have worked better,” reported that overall, individual members were committed to making a successful “individual” effort, were willing to put in the necessary time, and worked hard on the tasks they had each selected. Comments about things that could have worked better focused primarily on issues relating to time management (eg, meeting on time, time prioritization during the project, needing more time to work together as a group, meeting more often as a group) and group dynamics (eg, working together better, using others’ opinions better, and avoiding going off on a tangent during discussions). Specific comments from individuals included: “A desire (for each individual student) to want to explore the components other team members were working on could be improved” and “We were all very intense and often shifted away from the big picture.”

CONCLUSIONS

When analyzed using Rasch analysis, the Competition Student Assessment Instrument provided valuable, meaningful, and reliable information about processes and learning outcomes of the competition. The items in the instrument were sufficiently distributed on a hierarchical scale and functioned unidimensionally, providing evidence to support construct validity. Additionally, evidence supporting the validity of their perceptions retrospectively was provided when new knowledge obtained during the competition became stabilized during the posttest phase, Campbell and Stanley (1963) suggest that the potential threat to internal validity with the retrospective pretest-posttest is to distort past attitudes into agreement with present ones, ie, social desirability.\(^\text{25}\) The results from the instrument support that students took seriously the evaluation of the competition process. That is, 3 of the 6 pretest-posttest items were endorsed in the opposite direction as would be expected if social desirability had been a major factor biasing the results.

Those items that were hard to endorse generally asked students to evaluate skills and knowledge. This finding is reasonable when one considers that students may not be comfortable with making these assessments for a variety of reasons. One factor, student self-confidence, has a positive influence on performance and self-assessment.\(^\text{26}\) Additionally, performance accomplishment leads to building confidence because it is based on actual mastery experiences.\(^\text{27}\) Thus, these student-centered competitions would be a good source for self-confidence information.

While self-evaluation is an essential component to developing life-long learning skills, it takes experience and practice with self-assessment techniques to develop comfort with producing an objective reflection on one’s performance.\(^\text{28}\) Additionally, congruency between the educational tasks and the self-assessment instrument is a prerequisite for making predictions from self-evaluations.\(^\text{29}\) Intuitively, more accurate self-assessment will come after the response to the stimuli. This study used a design in which students were asked to self-assess their skills immediately following their use, thus providing a close proximity between self-assessment of their skills and their performance accomplishment.

Consistent with principles of continuous quality improvement, the authors have 3 recommendations for improving the competition process and student learning outcomes based on the objective evidence obtained from
the competition assessment instrument. First, the development and implementation of an educational intervention aimed at helping students develop skills and resolve issues associated with group dynamics via a team-building seminar is suggested. Improving group dynamics represented the single largest competition area in need of improvement.

Second, actual or real medications should be used in the competition portfolio. The departure from authenticity in developing cases for the competition was problematic for the students. This finding was exclusive to the AMCP competition because the APhA-ASP Patient Counseling Competition and the ASHP Clinical Skills Competition derive their competition cases using “real” medications that more closely parallel what students would expect to see in actual practice. When the contrived nature of the cases was combined with problems associated with group dynamics, student-perceived ability to efficiently and effectively problem solve was impaired.

Lastly, the addition of specific feedback for each team from competition judges (eg, academicians, managed care executives, and practitioners) that is directed at helping students improve their application of knowledge and skill development would be helpful to the students and may result in improved self-confidence. However, as previously discussed, the nature of “winning” may cloud the students’ ability to adequately measure this outcome. Revision of this item will be discussed prior to using the assessment in further competitions.

Student-driven competitions represent an innovation in applying student-directed activities outside the curriculum that is consistent with expected learning outcomes of the formal curriculum. Because of this, recommendations that address the needs of the competition, also complement needs for improving the process and learning outcomes in projects assigned in the formal classroom setting.

Typically, formative and summative student evaluations of courses include an assessment of the course itself and the instructor(s). However, students do not have the opportunity to evaluate reflectively how they have grown or benefited from the instruction. Slight modifications of this retrospective pre/post instrument would allow students to evaluate themselves retrospectively on how they achieved course goals and objectives. Also, nested in the evaluation could be an assessment of the student’s progress in developing performance-based abilities (eg, oral/written communication, problem solving, critical thinking) encouraged through course instructional methodology and activities.

The authors look forward to using this instrument to evaluate the student-learning outcomes of the APhA-ASP Counseling Competition and the ASHP Clinical Skills Competition.

ACKNOWLEDGEMENTS

The authors would like to acknowledge J. Warren Salmon, PhD, Professor of Pharmacy Administration, University of Illinois College of Pharmacy, and Margaret Byun, PharmD, for their untiring devotion to facilitating managed care instruction within schools and colleges of pharmacy. The AMCP Pharmacy and Therapeutics competition, which served as a basis for this research endeavor, was conceived and implemented by Dr. Salmon and now takes place at 13 schools and colleges of pharmacy nationwide where AMCP student chapters reside. Further, the authors acknowledge the contributions of the following UIC-COP doctor of pharmacy students who contributed their time and provided valuable suggestions during the development of the assessment instrument (Ryan Gehrke, Maria Onayemi, and Lydia Shenouda) and the students who served on the competition planning committee (Joy Zarlinga, Heidi Anetsberger, and Jenny Legon). Lastly, we acknowledge the encouragement and valuable input of Dr. Everett V. Smith, Assistant Professor at the University of Illinois College of Education. His expertise in psychometrics and objective measurement was especially helpful in producing this manuscript.

REFERENCES

Appendix 1: Selected University of Illinois-Chicago College of Pharmacy General Outcome Abilities

1. Conceptual competence
   The student shall understand the theoretical foundations of the pharmacy profession and its position in health care.

2. Integrative competence
   The student shall be able to meld effectively theory and skills in the practice setting.

3. Critical thinking and decision making abilities
   The student shall acquire, evaluate, synthesize and apply information, knowledge and processes relevant to the solution of an identified problem and make sound decisions.

4. Social Interaction
   The student shall read, write, speak, listen and use appropriate data, media and technology for varied purposes and audiences and communicate effectively in a variety of situations and circumstances.

5. Self-learning abilities and habits
   The student shall effectively self-assess and satisfy his or her learning needs on an ongoing basis.

Appendix 2: Rasch Rating Scale Diagnostic Criteria and Model Evaluation Indicators

I. Rasch Rating Scale Diagnostic Criteria:

A. a minimum of ten observations in each category
B. the shape of the rating scale distribution should be peaked
C. the average category measures should increase with the rating scale categories
D. the mean square (MNSQ) statistics should be between 0.6 and 1.4
E. category thresholds should increase with the rating scale categories
F. the category thresholds should be at least 1.4 logits apart and no more than five logits apart.

II. Rasch Model Evaluation Criteria

A. Separation reliability
   1. Separation index should be greater than 2.0.
   2. Item and person reliability should be at least 0.80.

B. INFIT and OUTFIT
   1. Mean square (MNSQ) statistics for items should be greater than 0.6 and less than 1.4 for rating scale application. Mean square (MNSQ) statistics for person measures should also be less than 1.4.
Using the scale provided, please rate each statement by circling the response that best describes:

as you *initially* felt before you began the competition process
as you feel *now*, after the competition presentation

<table>
<thead>
<tr>
<th>Statement</th>
<th>Initially</th>
<th>Weak</th>
<th>Fair</th>
<th>Good</th>
<th>Very good</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. My knowledge about the formulary management system</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Initially</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. My knowledge about clinical decision making</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Initially</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. My ability to problem-solve</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Initially</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. My ability to work with other students in a team setting</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Initially</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. My oral communication skills</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Initially</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. My self-confidence</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Initially</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. As a result of this competition, my ability to blend knowledge with skills in the practice setting improved.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Disagree</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. As a result of this competition, my writing skills improved.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Disagree</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. As a result of this competition, my ability to self-assess my learning needs improved.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Disagree</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. As a result of this competition, my interest in pursuing a career in managed care has increased.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Disagree</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11. Describe the adequacy of resources for this competition (eg, books, time).</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12. How were questions and/or concerns you had about the competition addressed?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13. What would you recommend to improve the competition in the future?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
14. Describe the contributions of the team that worked well.

15. Describe the contributions of the team that could have worked better.