Use of Situational Judgment Tests to Teach Empathy, Assertiveness, Communication, and Ethics

Kathryn J. Smith, PharmD, Stephen Neely, MPH, Vincent C. Dennis, PharmD, Misty M. Miller, PharmD, Melissa S. Medina, EdD

a University of Oklahoma, College of Pharmacy, Oklahoma City, Oklahoma
b Associate Editor, American Journal of Pharmaceutical Education, Arlington, Virginia

Corresponding Author: Kathryn J. Smith, University of Oklahoma, College of Pharmacy, 1110 N. Stonewall Ave., CPB 229, Oklahoma City, OK 73117. Tel: 773-368-9711. Email: kathryn-j-smith@ouhsc.edu

Submitted June 8, 2021; accepted October 8, 2021; ePublished October 2021

Objective. The objectives of this study were to (1) develop SJT scenarios for use in teaching empathy, assertiveness, and interprofessional communication in second-year pharmacy (P2) students, and ethics in third-year (P3) pharmacy students, (2) determine if SJTs developed were effective at measuring student’s communication skills and ethical judgment, (3) compare the performance of individual P3 students to the performance of teams of P3 students on the ethics SJT and (4) evaluate student feedback about SJTs as a teaching tool.

Methods. Pharmacy faculty developed five SJT scenarios related to communications and five SJT scenarios related to ethics and piloted the scenarios with P2 and P3 students, respectively. P2 students completed SJTs individually while P3 students completed theirs individually and as a team. Scenarios and responses were discussed with faculty after completion of the SJT. Students completed a questionnaire to provide feedback on the SJT.

Results. The communications SJT was completed by 59 P2 students with a mean score of 67.5% and reliability of $\alpha=0.594$. The ethics SJT was completed by 57 P3 students with a mean score of 80.1% and reliability of $\alpha=0.789$. The ethics SJT was also completed by 10 teams of P3 students resulting in a mean score of 93.2%. Students indicated the SJT content was realistic and the tests provided the opportunity to reflect on how to approach challenging situations.

Conclusion. SJTs were useful for teaching empathy, assertiveness, interprofessional communication, and ethics in pharmacy students. Future research should focus on predictive validity of SJTs for these content areas.

Keywords: situational judgment test, communication, teamwork, ethics, interprofessional

INTRODUCTION

Situational Judgment Tests (SJTs) are an emerging method of teaching and assessing non-cognitive attributes in pharmacy education.1-5 Previous studies have described SJTs for use during the pharmacy curriculum to develop or assess empathy, teamwork, ethics, self-directed learning, and problem-solving.1-3 SJTs consist of written or video scenario-based tests that assess judgments common in the test takers’ typical work environment that should not require clinical or procedural knowledge to complete.5 Test takers are presented with a scenario, then asked to rate or rank the appropriateness of various responses to the scenario. Test taker answers are compared to the answers of Subject Matter Experts (SMEs) and can be scored in a variety of ways, including giving higher scores for answers that more closely align with the SME answer.6,7 SJTs have been shown to be a valid and reliable way to assess attributes in the affective domain such as professionalism and leadership.8,9 They have also been shown to predict future behavior of medical and dental students and medical residents.10,11

The current study seeks to understand the difference between individual performance and team performance on the same SJT. This question is important because SJTs are usually completed independently, so it is not well understood how peer-teaching affects SJT performance. Evidence from Team Based Learning suggests students would perform better on an SJT completed with a group of peers rather than individually.12

OBJECTIVE

The objectives of this study were to (1) develop SJT scenarios for use in teaching empathy, assertiveness, and interprofessional (IP) communication in second-year pharmacy (P2) students, and ethics in third-year pharmacy (P3) students, (2) determine if the SJTs developed were effective at measuring communication skills (empathy, assertiveness, IP Communication) in P2 students and ethical judgment (ethics) in P3 students, (3) compare the performance of individual
P3 students to the performance of teams of P3 students on the ethics SJT and (4) evaluate student feedback about SJTs as a teaching tool.

METHODS

The development of the two SJTs in this study began by identifying desired professional attributes to assess in a second-year (P2) Clinical Communications course and a third-year (P3) Pharmacy Law and Ethics course (Figure 1). These courses were selected because the content area relates to non-cognitive attributes desired of pharmacy graduates. Professional attributes selected included Empathy, Assertiveness, and IP Teamwork for the Clinical Communication SJT and Ethics for the Pharmacy Law and Ethics SJT. An author (KJS) with previous training in SJT development and delivery interviewed faculty teaching in the Clinical Communication and Pharmacy Law and Ethics courses (MMM, VCD and MSM). These faculty were selected at SMEs because they hold advanced degrees in the area, teach the topic areas in the curriculum, and work closely with pharmacy students on a regular basis.

The SMEs were interviewed to gather examples of common scenarios from classroom or pharmacy practice settings in which they have observed pharmacy students act appropriately or inappropriately. These examples were used to develop SJT scenarios with eight to eleven possible responses. SJTs were written using a rating-type question format, wherein test takers are asked to rate the appropriateness of each possible response to the scenario presented. An example SJT can be found in Appendix 1. The SMEs reviewed the scenarios for their respective courses to ensure SJT best practices. The response key was finalized based on consensus among the SMEs. The items were uploaded to the learning management system (LMS) ( Desire2Learn, D2L Corporation, Kitchener, ON) and reviewed to ensure consistency with visual representation and content. Items were scored using a previously published matrix, based on a near-miss system that allowed for awarding partial credit (Table 1). Higher scores were given when ratings of appropriateness more closely aligned with those of SMEs.

A pilot study was conducted to assess reliability and impact of these SJTs. P2 students completed the Clinical Communications SJT in two sessions, one related to empathy and the other related to assertiveness and IP teamwork. P2 students were not pre-assigned to groups for other team-based activities, therefore, were not asked to complete the SJTs as a team. P3 students completed the Ethics SJT with five scenarios related to ethics during a single class session, first independently and then again as a pre-assigned team of four or five P3 students with whom they had been working during the course. All SJTs were administered in Fall 2020 when classes were being offered through synchronous video conferencing. Students were instructed to complete a questionnaire in the LMS after their final SJT to evaluate their perceptions of the SJT as a teaching tool. The questionnaire was similar to test-taker perception questionnaires from other SJTs in the literature and included up to seven items on a four-point Likert-scale (strongly agree, agree, disagree, strongly disagree). This study was deemed exempt by the Institutional Review Board (IRB) at the University of Oklahoma Health Sciences Center.

Internal consistency of the SJTs was measured using Cronbach’s alpha. Missing data were imputed using mean scores rather than to delete individuals listwise. To improve the internal consistency of each SJT, scenario items that reduced Cronbach’s alpha by .005 or more were removed. Only the retained items were used in the final psychometric analysis to evaluate the internal consistency, test difficulty, fairness and face validity. Item quality was determined via item discrimination index (DI), using quality delineations of good (DI>0.25), satisfactory (DI 0.24-0.17), moderate (DI 0.16-0.13), or limited (DI<0.13). Test difficulty was determined using mean scores, standard deviation (SD), and score distribution. Face validity was measured using survey results, which were summarized using frequency (percent). Data analysis was conducted using SAS software, Version 9.4 of the SAS System for Windows and JMP Pro Version 15.2.1 (SAS Institute Inc., Cary, NC, USA.).

RESULTS

Four scenarios with eight to eleven possible response items were included for the pilot communication SJT completed by 59 P2 students. Four percent of 1550 item responses were missing and imputed with mean score rather than drop individual row-wise. Out of the 37 response items, 25 (67.6%) performed well psychometrically (based on impact on Cronbach’s alpha) and were included in the final analysis. Item quality distribution determined by DI included 7 good items (28%), 6 satisfactory items (24%), 4 moderate items (16%) and 8 limited items (32%) which is consistent with other SJTs published in the pharmacy education literature. Scenarios retained between 50% and 88% of items. The overall mean score for the communication SJT was 67.4% (raw mean score 58, SD=6.4, range 42 to 70, maximum possible score 86). The communication SJT demonstrated good reliability with regards to internal consistency (α=0.594).

Five scenarios with six to ten possible responses each were included on the pilot ethics SJT completed by 57 P3 students. One percent of 1710 item responses were imputed. Out of the 38 response items, 30 (78.9%) items performed...
well psychometrically (based on impact on Cronbach’s alpha) and were included in the final analysis. Item quality distribution determined by DI included 14 good items (47%), 3 satisfactory items (10%), 4 moderate items (13%) and 9 limited items (30%) which is also consistent with other SJTs published in the pharmacy education literature. Scenarios retained between 71% and 88% of items. The overall mean score for this ethics SJT was 80.2% (raw mean score 87.4, SD=8.7, range 51 to 100, maximum possible score 109). The ethics SJT demonstrated good reliability with regards to internal consistency (α=0.789).

P3 students were divided into ten teams to complete the same ethics SJT. The overall mean score for the team ethics SJT was 93.2% (raw mean score 101.6, SD=6.9, range 86 to 109, maximum possible score 109). Individuals increased their scores on 36% of scenario items when answering as a team and decreased their scores on only 5% of items. Differences in individual vs. team performance on the ethics SJT were observed for item responses where students were asked to report unethical behavior to a preceptor or faculty member.

The response rate for the student evaluation questionnaire was 100%. Students were asked to indicate their level of agreement about the SJT regarding its relevance, difficulty, clarity, and ability to differentiate responses (Table 2). Eighty-eight percent of P3s (43 out of 49) and 60% of P2s (34 out of 57) said they would like to see more scenarios like the SJT in the curriculum. Examples of open-ended feedback from students included statements such as “Good exercise to promote conversation,” “The scenarios were helpful to get me thinking about these types of situations, but the debrief portion was when I really felt like I got the full benefit of the exercise,” and “This would be a great thing to add and ask about as we go through pharmacy school so it gets us more use to being active thinkers.”

**DISCUSSION**

The SJTs developed in this study were designed to teach empathy, assertiveness, communication, and ethical decision-making. The Cronbach’s alpha for the communication SJT (α=0.594) and ethics SJT (α=0.789) suggest scores accurately reflect the test takers’ skills in these domains. While the minimum desired level of reliability for a summative assessment is α=0.70, SJTs used for formative assessments may serve as a useful learning tool even if they do not reach this level. The difficulty of the items in this SJT as measured by DI were similar to other SJTs reported in pharmacy education. (Patterson 2019, Smith 2020) Furthermore, these SJTs were developed with fewer items than other SJTs reported in pharmacy education, while still reaching desired levels of reliability. Development and piloting of an SJT can be a significant time and/or financial commitment, but this study demonstrates the potential value of smaller, more focused SJTs to develop these skills in pharmacy students.

In the scoring method selected for this SJT, students received higher scores for ratings which more closely aligned with faculty SME rating. Therefore, it may be helpful to understand where students differ from SMEs in the rating of appropriateness in order to adjust or augment content in the curriculum related to areas where the greatest differences are seen. Two themes arose in responses where students differed from SME ratings of appropriateness: potential confrontation and timing of intervention. Student responses reflected a higher fear of conflict, especially with peers. This tendency may reflect hesitancy to act, even when inaction may not uphold professional standards or obligations. It is also possible that students are not willing to invest the time and energy into reporting unethical or dangerous behavior which does not directly affect their own learning, or seem to lead to self-harm, though corresponding behaviors in practice could have broader implications. SME ratings of appropriateness of responses to these complex situations seem to show the SME faculty were willing to intervene sooner than students. Students were more likely to rate “waiting” as appropriate or very appropriate, regardless of the situation. This tendency may represent an underdeveloped sense of responsibility to others (students, patients, the public) and/or an unawareness of risk management concerns, both of which continue developing after graduation. These results suggest that activities related to “Difficult Conversations” should be included in personal and professional development curricula throughout pharmacy school.

As expected, the performance on the ethics SJT improved when students completed the test with a team of peers (individual performance mean score 80.1% vs team performance mean score 93.2%). Team answers aligned with SMEs more frequently (range: 60 to 100%) than individuals (range: 5.6 to 98.2%). Change in agreement with SMEs ranged from 1.8% to 64.2% when completing as a team vs individually. When completing the SJT as a team, 38% of students increased their score (positive influence of team), 6% decreased their score (negative influence of team), and 56% had the same score. One scenario answer which had different ratings of appropriateness between the individual SJT and the team SJT can be found in Appendix 1. The SME rated the response, “Ignore the text from Kendra and continue studying on his own” as “Inappropriate, but not awful” because while it was appropriate that Adam did not join the group in looking at the old exam, it would have been best if he had alerted a faculty member to the potential academic misconduct. When completing the SJT as individuals, 16% of students agreed with the SMEs, but when completing the SJT as a team, 80% of teams agreed with the SMEs. We hypothesize students were avoiding conflict or confrontation when asked to respond
as an individual, but when among a group of their peers were emboldened to report the unethical behavior observed. This speaks to the significant positive or negative influence one may experience from a group of peers. Team-based learning is known to have a moderating effect on individuals within teams. Team participants who are naturally more assertive may become more cautious once they discover their teammates disagree with them, while participants who are less likely to speak up may be put on the spot and asked to defend their choice. While this activity was low-stakes, we hypothesize the team dynamics are similar since these student teams had been working together for the duration of the course.

Results of the student evaluation questionnaire indicated students found the SJT scenarios were realistic and allowed them to better understand the complexity and nuances associated with decision making in practice. While not all students felt they had all information necessary to make the most appropriate decision related to the scenarios, this mirrors practice, where practitioners are often required to make decisions without having all information desired. Students would benefit from additional modeling of decision-making in the face of uncertainty and complexity.

Limitations of this study include SJTs being only in written format. Emotions and tone would likely be better communicated through video SJTs, which could have led to different rating of appropriateness on the same scenarios. Additionally, there is inherent human error when one team member is entering an answer into a learning management system for a team of students, although we assumed that the answer put into the LMS did indeed reflect the opinion of the entire team. Because SJTs are measuring more than one characteristic, it is possible that a test-takers answers may be confounded by their preferred conflict or assertiveness styles. For example, a student could be aware that an observed behavior is unethical, but lack the assertiveness needed for reporting. Future studies could evaluate whether other personal attributes impact responses to SJT items. Further research may also include the use of SJTs to assess additional attributes (leadership, emotional intelligence) or assess professional identity development longitudinally throughout the curriculum, especially a sense of responsibility to others. Wolcott and colleagues have begun to study how prior knowledge and experience affects decision making on SJTs. As we improve our understanding of these potential influences, refinement of SJTs may improve their utility. Predictive validity of SJTs, or the correlation between SJT performance and future performance on professionalism assessments (for example, by preceptors on clinical learning experiences, or by state board of pharmacy), is an additional area of further inquiry.

CONCLUSION
SJT s are an emerging method of teaching and assessing essential professional attributes in pharmacy education. In this study, SJTs were developed to teach communication and ethical decision making to P2 and P3 pharmacy students. P3 students influenced each other to respond more like the faculty SMEs when taking the SJT with peers, demonstrating the positive impact peers can have on each other. Not only were the SJTs developed found to be valid and reliable, but students found the SJTs realistic. Discussions with faculty modeling their thought process on the SJT scenarios after completion of the SJTs by the students improved student understanding of decision making in pharmacy practice.

REFERENCES
1. Patterson F, Galbraith K, Flaxman C, Kirkpatrick CMJ. Evaluation of a Situational Judgement Test to Develop Non-Academic Skills in Pharmacy Students. Am J Pharm Educ. 2019;83(10). doi:10.5688/ajpe7074
2. Smith KJ, Flaxman C, Farland MZ, et al. Development and Validation of a Situational Judgement Test to Assess Professionalism. Am J Pharm Educ. 2020;84(7). doi:10.5688/ajpe7771
3. Wolcott MD, Lupton-Smith C, Cox WC, McLaughlin JE. A Five-Minute Situational Judgement Test to Assess Empathy in First-Year Student Pharmacists. Am J Pharm Educ. 2019;83(6):Article 6960. doi:10.5688/ajpe6960
4. Wolcott MD, Lobczowski NG, Zeeman JM, McLaughlin JE. Situational judgment test validity: an exploratory model of the participant response process using cognitive and think-aloud interviews. BMC Med Educ. 2020;20(1):506. doi:10.1186/s12909-020-02410-z
5. Wolcott MD, Lobczowski NG, Zeeman JM, McLaughlin JE. Role of Knowledge and Experience in Situational Judgment Test Responses of Pharmacists and Pharmacy Students. Am J Pharm Educ. 2021;85(1):Article 8194. doi:10.5688/ajpe8194
6. Bergman ME, Drasgow F. Scoring Situational Judgment Tests: Once You Get the Data, Your Troubles Begin. Int J Sel Assess. 2006;14(3):13. doi:10.1111/j.1468-2389.2006.00345.x
7. De Leng WE, Stegers-Jager KM, Husbands A, Dowell JS, Born MPH, Themmen APN. Scoring method of a Situational Judgment Test: influence on internal consistency reliability, adverse impact and correlation with personality? Adv Health Sci Educ. 2017;22(2):243-265. doi:10.1007/s10459-016-9720-7
8. Patterson F, Zibarras L, Ashworth V. Situational judgement tests in medical education and training: Research, theory and practice: AMEE Guide No. 100. Med Teach. 2016;38(1):3-17. doi:10.3109/0142159X.2015.1072619
9. Webster ES, Paton LW, Crampton PES, Tiffin PA. Situational judgement test validity for selection: A systematic review and meta-analysis. *Med Educ.* 2020;54(10):888-902. doi:10.1111/medu.14201

10. Patterson F, Lopes S, Harding S, Vaux E, Berkin L, Black D. The predictive validity of a situational judgement test, a clinical problem solving test and the core medical training selection methods for performance in specialty training. *Clin Med.* 2017;17(1):13-17. doi:10.7861/clinmedicine.17-1-13

11. Sahota GS, Taggar JS. The association between Situational Judgement Test (SJT) scores and professionalism concerns in undergraduate medical education. *Med Teach.* 2020;42(8):937-943. doi:10.1080/0142159X.2020.1772466

12. Parmelee D, Michaelsen LK, Cook S, Hudes PD. Team-based learning: a practical guide: AMEE guide no. 65. *Med Teach.* 2012;34(5):e275-287. doi:10.3109/0142159X.2012.651179

13. Patterson F, Driver R. Situational Judgement Tests (SJTs). In: Patterson F, Zibarras L, eds. *Selection and Recruitment in the Healthcare Professions: Research, Theory and Practice.* Springer International Publishing; 2018:79-112. doi:10.1007/978-3-319-94971-0_4

14. Cortina J. What is coefficient alpha? An examination of theory and applications. - PsycNET. *J Appl Psychol.* 1993;78(1):98. doi:10.1037%2F0021-9010.78.1.98

15. Goss BD, Ryan AT, Waring J, et al. Beyond Selection: The Use of Situational Judgement Tests in the Teaching and Assessment of Professionalism. *Acad Med.* 2017;92(6):780-784. doi:10.1097/ACM.0000000000001591

16. Epner DE, Baile WF. Difficult Conversations: Teaching Medical Oncology Trainees Communication Skills One Hour at a Time. *Acad Med.* 2014;89(4):578-584. doi:10.1097/ACM.0000000000000177

17. Eukel HN, Morrell B, Holmes SM, Kelsch MP. Managing Difficult Patient Encounters: Simulation Design, Findings, and Call to Action. *Am J Pharm Educ.* Published online January 13, 2021. doi:10.5688/ajpe8327

18. Smith KJ. Wicked Problems in Pharmacy Education. *Am J Pharm Educ.* Published online March 12, 2021. doi:10.5688/ajpe8491

5
Table 1. Scoring Matrix Detailing Points Awarded based on alignment with SME answer for this Situational Judgment Test

| Correct Answer* | Very appropriate thing to do | Appropriate, but not ideal | Inappropriate, but not awful | Very inappropriate thing to do |
|-----------------|-----------------------------|---------------------------|-------------------------------|-----------------------------|
| Very appropriate thing to do | 4* | 3 | 1 | 0 |
| Appropriate, but not ideal | 2 | 3* | 1 | 0 |
| Inappropriate, but not awful | 0 | 1 | 3* | 2 |
| Very inappropriate thing to do | 0 | 1 | 3 | 4* |

*Correct answer for the item based on subject matter expert (SME) consensus.

Table 2. Student Evaluation of the Situational Judgment Test

| Clinical Communications SJT | Strongly Agree or Agree (%) | Strongly Disagree or Disagree (%) |
|-----------------------------|-----------------------------|----------------------------------|
| The content of the empathy scenario was realistic with respect to pharmacy practice. | 86.0 | 14.0 |
| The content of the assertiveness scenarios was realistic with respect to pharmacy practice. | 96.5 | 3.5 |
| The content of the interprofessional scenario was realistic with respect to pharmacy practice. | 91.2 | 8.8 |
| I had enough information to make the most appropriate and subtle distinctions of judgment for the scenarios I examined. | 73.7 | 26.3 |
| After completing these scenarios in D2L, I better understand the complexity and nuances of making judgments in a pharmacy setting. | 84.2 | 15.8 |
| After debriefing the scenarios with the faculty, I better understand the complexity and nuances of making judgments in a pharmacy setting. | 93.0 | 7.0 |
| These scenarios gave me the opportunity to reflect on how I would approach challenging situations as a pharmacy student. | 89.5 | 10.5 |

| Ethics SJT | Strongly Agree or Agree (%) | Strongly Disagree or Disagree (%) |
|-----------------------------|-----------------------------|----------------------------------|
| The content of the ethical scenarios was realistic with respect to pharmacy practice. | 100.0 | 0.0 |
| I had enough information to make the most appropriate and subtle distinctions of judgment for the scenarios I examined. | 91.8 | 8.2 |
| After completing these scenarios in D2L, I better understand the complexity and nuances of making judgments in a pharmacy setting. | 98.0 | 2.0 |
| After debriefing the scenarios with the faculty, I better understand the complexity and nuances of making judgments in a pharmacy setting. | 100.0 | 0.0 |
| These scenarios gave me the opportunity to reflect on how I would approach challenging situations as a pharmacy student. | 98.0 | 2.0 |

SJT=Situational Judgment Test; D2L= Desire to Learn Learning Management System used for delivery of the SJT scenarios.
Figure 1. Steps Taken to Develop This Situational Judgment Test

Abbreviations: SJT, situational judgment test; SME, subject matter expert; LMS, learning management system IP, interprofessional
Adapted from reference 2, 13
Appendix 1. Example SJT Scenario with Items

Scenario

Adam is a third-year pharmacy student. While studying for an upcoming exam, Adam receives a group text message from his classmate Kendra which includes about 10 members of the third-year pharmacy class. The text message says “This material is hard! I’m so glad I got a copy of last year’s exam. Who wants to meet up later to go over it together?”

Response Instructions:
How appropriate are each of the following responses by *Adam* in this situation?

Rating Options:
1 = A very appropriate thing to do
2 = Appropriate, but not ideal
3 = Inappropriate, but not awful
4 = A very inappropriate thing to do

Possible Responses:
- Ask his classmate, Roy, who was not on the group text, if he wants to go to the meet up with him.
- Reply to the group text saying he does not think looking at last year’s exam is a good idea.
- Ignore the text from Kendra and continue studying on his own.
- Ask his classmate, Roy, who was not on the group text, what Roy thinks he should do.
- Reply to the group text saying he would like to meet up to talk about last year’s exam.
- Ask his faculty advisor what he should do if he suspects cheating might be occurring.
- Send an email to the course coordinator stating Kendra has a copy of last year’s exam.
- Send an email to the course coordinator stating he has heard there may be cheating occurring.