RESEARCH

Exploring the Role of Item Scenario Features on Situational Judgment Test Response Selections

Michael D. Wolcott, PharmD, PhD,[a,b] Nikki G. Lobczowski, PhD, MEd,[c] Jacqueline M. Zeeman, PharmD,[a] Jacqueline E. McLaughlin, PhD, MS[a]

[a] The University of North Carolina, Eshelman School of Pharmacy, Chapel Hill, North Carolina
[b] The University of North Carolina, Adams School of Dentistry, Chapel Hill, North Carolina
[c] Carnegie Mellon University, Human-Computer Interaction Institute, Pittsburgh, Pennsylvania

Objective. To explore pharmacists' and pharmacy students' perceptions regarding the significance of changing the features of test item scenario (eg, switching from a health care to a non-health care context) on their situational judgment test (SJT) responses.

Methods. Fifteen Doctor of Pharmacy students and 15 pharmacists completed a 12-item SJT intended to measure empathy. The test included six scenarios in a health care context and six scenarios in a non-health care context; participants had to rank potential response options in order of appropriateness and no two items could be of equal rank. Qualitative data were collected individually from participants using think-aloud and cognitive interview techniques. During the cognitive interview, participants were asked how they selected their final responses for each item and whether they would have changed their answer if features of the scenario were switched (eg, changed to a non-health care context if the original item was in a health care context). Interviews were transcribed and a thematic analysis was conducted to identify the features of the scenario for each item that were perceived to impact response selections.

Results. Participants stated that they would have changed their responses on average 51.3% of the time (range 20%-100%) if the features of the scenario for an item were changed. Qualitative analysis identified four pertinent scenario features that may influence response selections, which included information about the examinee, the actors in the scenario, the relationship between examinee and actors, and details about the situation. There was no discernible pattern linking scenario features to the component of empathy being measured or participant type.

Conclusion. Results from this study suggest that the features of the scenario described in an SJT item could influence response selections. These features should be considered in the SJT design process and require further research to determine the extent of their impact on SJT performance.

Keywords: cognitive interview, empathy, response process, situational judgment test, think-aloud protocol

INTRODUCTION

Situational judgment tests (SJTs) are increasingly being used in health professions assessment models to evaluate attributes such as empathy, adaptability, integrity, and collaboration.[1-6] During an SJT, examinees review hypothetical scenarios and then select potential responses based on their appropriateness. Each scenario is intended to measure a construct of interest (ie, empathy, adaptability,

integrity) that is expected for optimal job performance.[7,8] Examinees are then given a score based on how well their selections correspond with a key established by a panel of subject matter experts. Higher scores often indicate higher standing on the measured construct.[9]

Situational judgment tests can add value in pharmacy education as a tool to evaluate social and behavioral attributes of individuals during admissions and capstones and at critical timepoints in the curriculum. For example, an SJT can be used as an additional screening tool to supplement traditional admissions criteria or it can be used as an assessment to ensure pharmacy students' readiness to begin introductory or advanced pharmacy practice experiences.[3,4,10,11]

Corresponding Author: Michael D. Wolcott, The University of North Carolina at Chapel Hill, Eshelman School of Pharmacy, 321 Beard Hall, Chapel Hill, NC 27599. Tel: 919-537-3854. Email: wolcottm@email.unc.edu

Situational judgment tests offer unique advantages because they require minimal resources to administer and grade compared to other techniques, such as Multiple Mini Interviews (MMIs).[6]

As SJTs become more popular, validity research continues to increase to support score interpretation and utility as an assessment method. Of the sources for validity evidence, research about the response processes and response selections during an SJT remains under-studied.[12-17] The response process refers to the moment-to-moment cognitive steps that are required to think and make decisions. This includes how information is accessed, manipulated, and applied to answer the posed question during an assessment.[18,19] Validity research on SJT response processes evaluates whether the SJT items elicit the knowledge and experience that is intended to be measured during the assessment. If the SJT does not elicit the appropriate knowledge and experience from a participant, then it is not a reliable measure of the intended construct.[13,17]

One question about SJT response processes is the extent to which an SJT is truly situational. In other words, does the context and features presented in the SJT item scenario impact how an individual will respond to it? For this study, the scenario context refers to the general setting, environment, or location of the event described, such as at a grocery store, bank, hospital, or clinic. The scenario features refer to details about the situation, such as who is involved in the encounter, observations, and pertinent information. Situational judgment test design focuses on creating realistic scenarios that parallel job experiences. This often includes integrating details about the scenario in which these experiences would occur.[20,21] However, the use of scenario context and features information is debated (eg, does it have an impact on response selections?). Moreover, it if does not have a substantial impact, should details about the context be removed to reduce cognitive load for examinees?[22-25]

Lievens and Motowidlo argue that scenario details can be stripped from an SJT item and not dramatically influence examinee performance.[13] They posit that general domain knowledge and specific job knowledge relate equally to SJT performance; therefore, SJTs could be considered tests of general domain knowledge instead of situationally specific knowledge and therefore do not require information about the scenario setting. However, few studies have explored this theory. In one study involving an SJT intended to measure teamwork, researchers used two versions of the test, one with scenario descriptions and one without, to evaluate the difference in subjects' performance. For example, one question described a scenario that detailed the examinee being in an argument with their co-workers and asked the examinee to identify an effective way to resolve the situation from a list of options. The alternative question without a scenario asked the examinee to compare the behaviors and select the one that was most effective or reflected an ideal behavior for resolving a conflict.[14] The authors concluded that having scenario descriptions did not impact examinees' performance on 71% of items.[14] In addition, the authors evaluated whether the impact of scenario features differed based on the construct being assessed. A 30-item SJT was created to measure teamwork, integrity, and decision-making by pilots. Results showed that scenario descriptions did not impact the subject's performance on 63% of items. The authors, however, did suggest that test results might vary based on the construct measured.[14]

Understanding this relationship of SJT scenario context and features to response selections is important as it has implications for design and score interpretation of SJTs if they are to be used in pharmacy education. Thus far, research about SJT item scenario features has been conducted exclusively outside of the health professions despite the increased use of SJTs in health professions education.[20,26,27] The purpose of this study was to explore the role of item scenario features in health care and non-health care contexts with respect to a participants' response selections on an SJT intended to measure empathy. In addition, the study explored whether there were patterns based on participant type (ie, student or practicing pharmacist) and empathy component (ie, affective or cognitive).

METHODS

This report highlights one component of a comprehensive research project focused on exploring SJT response processes. Additional details about the design, implementation, and results of the broader study can be found elsewhere.[16,28]

This study used a convenience sample of 15 students completing a Doctor of Pharmacy program and 15 pharmacists with at least five years of practice experience. This sample size was selected due to the exploratory nature of the study and because previous work suggested that as few as 11 participants was sufficient to achieve coding saturation when evaluating SJT response processes.[15] Participants were assigned an alphanumeric identifier to designate whether they were a student (labeled "S" and numbered 1 to 15) or a pharmacist (labeled "P" and numbered 1 to 15). References to specific participants use these alphanumeric identifiers to ensure anonymity.

Development of the study SJT included a construct-driven approach created in collaboration with 11 faculty pharmacists at the University of North Carolina Eshelman (UNC) School of Pharmacy.[16,28,29] Details about the design and psychometric analysis of the SJT developed for this study are provided in detail elsewhere.[16,28] The SJT was

designed to measure empathy, a critical construct in health professions education and health care.[30,31] Empathy is considered a multidimensional construct with two factors: *cognitive empathy*, the ability to understand another person's perspective, and *affective empathy*, the ability to understand and internalize the feelings experienced by others.[32-36]

The SJT included 12 items that were equally distributed throughout the test based on the empathy component (ie, six items that measured either affective or cognitive empathy, labeled with an "A" or "C," respectively) and scenario context (ie, six items with scenarios that took place either in a health care or non-health care environment, labeled with an "H" or "N," respectively). Having SJT items with situations that occurred in different contexts was intentional to explore how participant perceptions about the features in these contexts influenced their response process. Each item included a brief scenario followed by five plausible response options (samples are provided in Appendix 1, along with a summary of the SJT item cases). Situational judgment test items used a knowledge-based format, ie, examinees were asked how they *should* respond to a given scenario. Examinees then ranked the response options from most appropriate to least appropriate and were not allowed to have two or more responses be ranked equally. The ranking response format requires participants to be more explicit in their decision-making process and thus offered an advantage compared to other response formats.[20,21] Examinees had approximately two minutes per question to select their response.

The study followed a data collection procedure that was approved by the UNC Institutional Review Board. Participants met with the researcher for a 90-minute one-on-one interview, which included a think-aloud interview, a cognitive interview, and a demographic survey. The interviews were audio recorded for transcription purposes (consent was obtained from the participant beforehand). During the think-aloud interview, participants completed the full 12-item SJT, verbalizing their thoughts as they worked through the items. The interviewer did not intervene unless there was silence for more than five seconds, at which point they said, "keep talking."[37,38]

Next, the cognitive interview was conducted, in which participants were asked about their understanding and thought process about eight preselected SJT items.[38,39] The selected items were randomized and stratified so that each participant reviewed four items related to non-health care settings, health care settings, affective empathy, and cognitive empathy. Participants had the opportunity to review each item during the cognitive interview; however, they could not change their responses. In summary, each item on the SJT was discussed in 30 think-aloud interviews (15 from both pharmacists and students) and 20 cognitive interviews (10 from both pharmacists and students).

During the cognitive interview, each participant was asked about how their response selections would have changed, if at all, if the item were in a different context (eg, switched to a health care environment from a non-health care environment or vice versa). The interviewer verbally described an alternative scenario to the participant (Appendix 1). For example, one non-health care scenario involved a relative who was upset about having difficulty conceiving a child. The alternative scenario in a health care setting involved a patient who was upset about having difficulty conceiving. The alternative scenarios were presented consistently by the interviewer across participants. Participants were asked to verbally share whether their answer would change and to provide a rationale for their response. The number of participants who said their responses would have changed were summarized at the item level. Using the Fisher exact test, statistical comparisons were conducted to explore if there were differences in responses based on participant type (ie, student or pharmacist). Statistical analyses were conducted using Stata, version 15 (StataCorp, LP). A relationship with a *p* value of less than .05 was considered statistically significant.[40,41]

Qualitative analysis of de-identified transcript data was conducted to identify features about the item scenario that may have influenced participants' response selections. The analysis focused on responses to cognitive interview questions about the test item scenario.[42] Transcripts were coded by two independent reviewers in a three-stage process (ie, double coding, single coding plus auditing, single coding) using an *a priori* codebook developed from existing theoretical and empirical understanding of SJT response processes.[13,17,43,44] The coding process also allowed for inductive codes to emerge when agreed upon by the reviewers in a process consistent with standards in qualitative data analysis.[45,46] Additional details about the coding process and the full codebook are available in other publications.[16,28] Coded transcripts were reviewed to identify patterns in the features related to the item scenario that may influence SJT response selections. These features were summarized and sample quotations included as supporting evidence.

RESULTS

The study included 30 participants: 15 students and 15 licensed pharmacists with at least five years of experience. The student group was predominantly female (n=11, 73.3%) with a median age of 24 years (range, 22-45 years). Most students were entering their third or fourth year of pharmacy school (n=11, 73%), and 13 students (87%) worked in a health care-related job outside of school. The pharmacist group was also predominantly female (n=13,

86.6%) with a median age of 36 years (range, 29-51 years). Participating pharmacists had a median of 8 years of experience as a licensed pharmacist (range, 6-23 years), and all were employed in a university hospital setting across multiple disciplines. Additional details about the participants and SJT performance are available in other publications.[16,28]

On average, participants stated their responses would have changed 51.3% of the time (range 20-100%) if the item scenario features were switched from a health care context to non-health care context or vice versa. If the original item was in a health care context, all participants would have changed their response 50.8% of the time (range 20-80%) and 51.6% of time (range 30-100%) if the original item was originally in a non-health care context. Pharmacists reported they would change their answers on average 46.7% of the time (range 0-100%), whereas students suggested they would change their responses on average 55.8% of the time (range 20-100%). Participants shared they would have changed their answer 43.3% of the time for items that measured affective empathy and 59.2% of the time for items that measured cognitive empathy. The number of participants who stated their response would have changed at the item level are presented in Table 1. The Fisher exact test found no statistical differences in the number of students vs pharmacists who believed they would change their selections if the item scenario context changed. Qualitative examples to describe how they would have altered their responses were also summarized to identify how participants would have liked to change their responses. Pharmacists and students often reported making similar modifications to their selections.

One unique example was item CN2, on which all 20 participants reported they would change their responses if the context were switched to a health care environment. This item, which refers to a woman asking to cut in line at the grocery store to get home to her sick child, was perceived differently when applied to a health care context in which a patient asks to cut in front of someone. In this case, all of the participants referenced rules or policies in health care that prioritize patients based on the severity of their situation. Participants discussed how they triage patients in the emergency department (P03) or use transplant waiting lists (S02) to describe how patients are screened and placed in an order that is not often modified. One pharmacist stated that there is usually a "protocol that you can fall back on" (P06) in these situations that makes their decision regarding the patient's request more straightforward.

Inductive coding of transcript data led to the identification of four item scenario features that were believed to influence the examinee's response selections when the context was switched. They include factors pertaining to the participant or *examinee*, factors pertaining to *actors* in the presented scenario, factors pertaining to the *relationship* between the examinee and actor(s), and factors pertaining to the *situation*. Definitions of the features are summarized in Table 2, and more examples and quotations are provided in Table 3. Across transcripts, participants used the phrase "it depends" 175 times, which may signify the importance of item scenario features in SJT response selections.

Participants often cited multiple features in one item that influenced their response selections, and they noted features that could affect their response differently based on the scenario. For example, item AH1 asked how an examinee should respond to a patient who is upset about the recent loss of a loved one. One pharmacist (P06) stated that, "if it were a friend, I would have been more inclined to share my own personal experiences … I'd feel more comfortable sharing personal loss and talking about it on a more personal level." In this case, the participant identified that who the actor is (ie, a friend instead of a patient) as well as their relationship with the person (eg, a personal instead of a professional relationship) had an impact on their response selection. Participants commonly identified that relationships with friends and family members come with different expectations than relationships with work colleagues or patients. For example, in item CN1, student S10 shared that it would be easier to convince a patient not to take a non-prescribed medication than it would to convince a friend not to "because you could come at it from the standpoint of I've had training in this … and there's no evidence to back this up or that's illegal."

Participants also wanted additional information about their role in the scenario and more details about the situation so they could make more informed decisions. For example, item AH2 involves a nurse asking the examinee to discuss a medication error with a patient. Student S04 wanted to know if they as an examinee were "in a position of responsibility" in the setting because that might impact how they would approach the situation. Moreover, when asked how their response would change if the scenario was in a non-health care context (eg, a restaurant), several participants felt greater ease with disclosing an error when it related to a situation such as the person providing them with the wrong food order. The situation and consequences of the situation, therefore, could impact response selections.

The prevalence of these four factors (ie, examinee, actors, relationship, situation) across the 12 SJT items were also evaluated to identify if there were patterns that linked the scenario features to other components of the item, such as the participant type (ie, student or pharmacist), and the empathy component intended to be measured (ie, affective or cognitive). References to features were aggregated and the most prevalent features identified for each item are reported in Table 3 along with examples for reference. In

Table 1. Participants' Perceptions of How Their Responses to Situational Judgment Test Items Would Change if the Features of the Scenario Were Changed

	Reported Response Selection Would Change[a]			How the Response Would Change	
	Pharmacist n (%)	Student n (%)	Total n (%)	Pharmacist	Student
Healthcare Scenario Features				**If the scenario features were non-healthcare related, the participant would instead …**	
CH1	6 (60)	9 (90)	15 (75)	Agree; notify the person	Share more personal stories; notify the person
CH2	4 (40)	0 (0)	4 (20)	File a complaint sooner; ask for alternatives	No responses
CH3	5 (50)	7 (70)	12 (60)	Confront them about lying; ask less questions	Confront them about lying
AH1	2 (20)	6 (60)	8 (40)	Share personal stories more	Share more personal stories; recommend help
AH2	3 (30)	3 (30)	6 (30)	Not explain the cause	Not explain the cause; apologize sooner
AH3	7 (70)	9 (90)	16 (80)	Transition from the topic; stop talking	Transition from the topic; Stop talking
Non-Healthcare Scenario Features				**If the scenario features were healthcare related, the participant would instead …**	
CN1	5 (50)	4 (40)	9 (45)	Tell them not to take the medication sooner	Tell them not to take the medication sooner
CN2	10 (100)	10 (100)	20 (100)	Allow a patient to cut if critical concern	Allow a patient to cut if critical concern
CN3	5 (50)	6 (60)	11 (55)	Divert conversation earlier; dismiss family	Divert conversation earlier; leave the location
AN1	3 (30)	5 (50)	8 (40)	Request more information; not readily leave	Request more information; not readily leave
AN2	4 (40)	4 (40)	8 (40)	Discuss options; not discuss experiences	Discuss options; not discuss experiences
AN3	2 (20)	4 (40)	6 (30)	Support decision sooner	Support decision earlier; recommend help

Abbreviations: A=affective empathy; C=cognitive empathy; H=healthcare setting, N=non-healthcare setting; 1, 2, 3=test item number.
[a] Differences between pharmacists and students for each item were statistically non-significant according to results of the Fisher exact test ($p>.05$). Items were randomized to participants due to time constraints, and not all 30 participants reviewed every item during the cognitive interview.

general, no discernible pattern was found that linked scenario features to the component of empathy being measured or participant type. In other words, features about the examinee, actor, relationship, and situation could be relevant regardless of the item measuring affective or cognitive empathy and whether it was a student or pharmacist answering the item. Student and pharmacist responses on the most salient feature that could influence their response selections were often the same; however, students frequently listed more than one feature per item while pharmacists listed only one.

DISCUSSION

The purpose of this study was to describe the perceived impact of item scenario features on SJT response selections, specifically within health care and non-health care contexts. The study contributes to discussions about whether SJTs are truly situational, which has been debated in other

disciplines.[22-25] Results from our study suggest that response selections may depend on item scenario features (eg, examinee position, actors in the scenario, relationships, and scenario details). This is larger than studies in other disciplines where 30%-40% of items were dependent on scenario features.[13-15] Therefore, we argue item scenario features should be considered an important factor in SJT response selections.

Additional research is warranted as there are several possible explanations for this observation. The first is to evaluate whether changes in item scenario features truly impact response selections. This research asked participants explicitly to reflect on their selections instead of incorporating actual items with different scenario features (eg, change in actors, provided details) that could be compared directly. Similar to another study, the explicit questioning about scenario features for each item could have influenced the participant's perception of the impact these features had, which

Table 2. Description of Item Scenario Features Perceived to Influence Response Selections on Situational Judgment Test

Component	Description and Examples of Scenario Features
Examinee	Position and responsibilities of the examinee (eg, healthcare provider, manager)
	Needs, wants, or expectations of the examinee (eg, responsibilities for patient care)
	How the examinee communicates with an actor (eg, how something is said)
	How long the examinee would pursue an action or outcome (eg, interest in the goal)
	Proximity of the examinee to the situation (eg, directly involved/affected, observing)
Actor	Position and responsibilities of an actor (eg, healthcare provider, family, friend)
	Needs, wants, or expectations of an actor (eg, what the patient would want)
	How an actor communicates to the examinee (eg, how rude they are)
	How an actor is anticipated to respond to an action (eg, potential outcome)
	An actor's personality (eg, openness, willingness)
Relationship	Relative position between examinee and actor (eg, boss, student, sibling)
	Whether the relationship is expected to be professional OR personal
	How long an actor and examinee have known one another (eg, relationship duration)
	How much information an actor and examinee know about one another (eg, history)
	Level of comfort between examinee and actor (eg, comfort with being honest)
	How information is shared between examinee and actor (eg, communication medium)
Situation	Severity of the consequences related to an action (eg, high-stakes, low-stakes)
	Severity of the current situation (eg, safety, emotional wellbeing, necessity)
	Legality of potential actions or lack thereof (eg, false documentation, illegal drugs)
	Actions that were previously attempted (eg, other steps taken prior to the question)
	Amount of information available or could be obtained (eg, background knowledge)
	Feasibility or capability to complete potential actions (eg, authority, resources)

may have otherwise gone unnoticed.[15] However, a benefit was that participants described what features could impact their response. Further research could evaluate how modifications of these specific features led to different response selections directly.

Another explanation as to why our results differed from those in previous reports could be the shift between scenario features related to health care and non-health care contexts. Previous research has been limited to retail, management, and military contexts. Features related to health care contexts may be more influential because of the potential consequences (ie, life or death decisions). Participants described what we perceived as a balance of personal and professional interactions that are expected from health care providers and may not be applicable in other disciplines researched. Moreover, non-health care related SJTs may be more likely to include items that involve interactions with strangers, which is not frequently the case in health care related SJT items. In this study, only a few of the questions described interactions with individuals who were (hypothetically) strangers to the participant, while most items depicted an established relationship between patients and providers.

The results of this study should be interpreted with consideration of several limitations. First, the study included a small convenience sample of students and pharmacists in one region of the United States. This sample included predominantly female participants, students in their third and fourth year, and pharmacists from hospital settings. This was intentional and acceptable because of the exploratory nature of the study and constraints anticipated with qualitative data collection and analysis. However, the validity of the results could have been more robust if a non-pharmacist comparison group or a sample with greater variability in gender, ethnicity, and practice experience had been used. It is unknown whether these features of the scenarios would be as salient or exhaustive for participants from other regions or health professions. In addition, the findings are predicated on an SJT that was appropriately designed and effectively evaluated the construct of interest. The SJT designed for this research followed best practices to promote alignment with the construct and psychometric analyses, which suggests that the SJT provided reasonable estimates of examinee empathy.[16,28] While this could be a significant concern if the research was evaluating differences in SJT performance based on scenario features, this was not considered problematic as the designed SJT served as the vessel to obtain information about response selections. Confirmatory research on SJTs should ensure greater measures of reliability and validity if there are claims about changes to scenario features and performance. This exploratory research focused predominantly on qualitative data to describe features that influence the SJT response processes; therefore, the psychometric and performance data were not a substantial component of this report.

Table 3. Examples of the Most Prominent Scenario Features Believed to Impact Response Selections Organized by Situational Judgment Test Item

Item	Pharmacist Examples of Scenario Features	Student Examples of Scenario Features
CH1	Relationship: "because you have more of a relationship with that person" (P08)	Relationship: "I'm a little more reticent to share a personal story as a healthcare provider" (S04)
CH2	Actor: "If the pharmacist … is really abrupt and abrasive then it changes how you respond" (P04)	Situation: "When it has to do with medications … then it gets to be a little higher stake" (S07)
CH3	Relationship: "I think depending on my relation to that person, I would act accordingly" (P09)	Relationship: "With a family member you just already have that trust" (S13)
AH1	Relationship: "I know this individual personally" (P03) Actor: "It's not as straightforward … depending on where the patient is in their disease state" (P04)	Relationship: "It's going to skew your decision … what is your connection with them" (S04) Actor: "it depends on the patient … the patient is usually thinking about this more" (S01)
AH2	Situation: "I can fix that. I can say "You don't want the hamburger, okay" (P01)	Situation: "I'd feel more comfortable talking about the error if it was something like food" (S02) Examinee: "Am I in a position of responsibility in this setting?" (S04)
AH3	Situation: "In a healthcare setting where there are policies and procedures to follow" (P06)	Situation: "It also depends on the hospital policy" (S13) Relationship: "It would depend on … how close of a friend it was" (S02)
CN1	Examinee: "That can come off a little harsh so that's why … to make it softer" (P05)	Situation: "You could buy caffeine pills … but Adderall is a controlled substance" (S12)
CN2	Situation: "Is this a new job presentation versus is this just your study thing" (P04)	Situation: "Standing in line at a checkout is a lot lower stakes than a healthcare setting" (S07)
CN3	Examinee: "The part of us that's a little bit gossipy or curious will say outside our context" (P14) Relationship: "Depending upon the level of relationship with that patient" (P15)	Relationship: "We're not all bunch of friends … like I have some sheen of authority" (S02) Situation: "You don't know all the facts … so that vagueness would make it difficult" (S06)
AN1	Actor: "It depends on which parent … because I like one better than the other" (P02)	Situation: "I guess I just assumed that there was some kind of worst-case scenario" (S06)
AN2	Examinee: "I was the bitch with the baby" (P01)	Examinee: "If I were their healthcare provider? Then I'll be just really clear" (S11)
AN3	Examinee: "It would be different depending on your views of higher education" (P06)	Examinee: "You're thinking about how involved you want to be in this person's situation" (S03)

Abbreviations: A=affective empathy; C=cognitive empathy; H=healthcare setting; N=non-healthcare setting; 1, 2, 3=test item number S=student; P=pharmacist, 00=participant number.

Findings from this study challenge the idea that the features of scenarios presented in SJT items do not impact participants' response selections. Therefore, the results of our study have implications for the design of future SJTs used in pharmacy education. Those responsible for SJT design and use should reflect on which scenario features should be intentionally included and the potential interpretation of the scenario features by examinees as this could impact their selections. For example, whether a physician, family member, or friend is included in a scenario may influence the examinee's response, and, in some items, this may adversely impact the examinee's selections. Additional research is warranted to systematically evaluate how modifications to the features of SJT item scenarios influence an examinee's performance without direct prompting. Doing so could identify which features (eg, the actor, examinee, relationship, or situation) have the most significant effect on the participant's responses.

Finally, this research raises questions as to whether SJTs used in health professions education should include items with scenarios that take place in both health care and non-health care settings. Participants noted there may be different "rules" that apply in those contexts. If the goal of administering an SJT is to understand how an examinee will behave in a health care context then items should be focused exclusively in those settings as the examinee's behaviors may not transfer to non-health care contexts. Responses to scenarios that take place in other contexts

may be influenced by other factors unrelated to the construct being measured. Until further research can be conducted, SJT designers in pharmacy education should be cognizant of test specifications and the potential consequences of deciding to include items with diverse contexts.

CONCLUSION

Results from this study provide evidence that scenario features of items on situational judgment tests could influence test takers' response selections. Specific features that could modify response selections include information about the examinee, the actor(s) in the SJT item scenario, relationships between the examinee and actors, and details about the situation. These features should be considered in the SJT design process. Further research is needed to determine the extent of the impact that the features of SJT item scenarios have on the performance of pharmacy school applicants, pharmacy students, and pharmacists.

REFERENCES

1. Campion MC, Ployhart RE, MacKenzie Jr WI. The state of research on situational judgment tests: a content analysis and directions for future research. *Human Performance.* 2014;27:283-310.
2. Koczwara A, Patterson F, Zibarras L, Kerrin M, Irish B, Wilkinson M. Evaluating cognitive ability, knowledge tests, and situational judgment tests for postgraduate selection. *Med Educ.* 2012;46:399-408.
3. Patterson F, Knight A, Dowell J, Nicholson S, Cousans F, Cleland J. How effective are selection methods in medical education? a systematic review. *Med Educ.* 2016;50:36-60.
4. Patterson F, Ashworth V, Zibarras L, Coan P, Kerrin M, O'Neill P. Evaluations of situational judgement tests to assess non-academic attributes in selection. *Med Educ.* 2012;46:850-868.
5. Petty-Saphon K, Walker KA, Patterson F, Ashworth V, Edwards H. Situational judgment tests reliably measure professional attributes important for clinical practice. *Adv Med Educ Pract.* 2016;8: 21-23.
6. Wolcott MD, Lupton-Smith C, Cox WC, McLaughlin JE. A 5-minute situational judgment test to assess empathy in first year student pharmacists. *Am J Pharm Educ.* 2019;83(6):6960.
7. Chan D, Schmitt N. Situational judgment and job performance. *Human Performance.* 2002;15:233-254.
8. Lievens F, Patterson F. The validity and incremental validity of knowledge tests, low-fidelity simulations, and high-fidelity simulations for predicting job performance in advanced-level high-stakes selection. *J Appl Psychol.* 2011;96:927-940.
9. De Leng WE, Stegers-Jager KM, Husbands A, Dowell JS, Born MP, Themmen APN. Scoring methods of a situational judgment test: Influence on internal consistency reliability, adverse impact and correlation with personality? *Adv Health Sci Educ Theory Pract.* 2017;22: 243-265
10. Smith KJ, Flaxman C, Farland MZ, Thomas A, Buring SM, Whalen K, Patterson F. Development and validation of a situational judgement test to assess professionalism. *Am J Pharm Educ.* 2020; 84(7):7771.
11. 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):7074.
12. American Educational Research Association, American Psychological Association, National Council on Measurement in Education. *Standards for educational and psychological testing.* Washington, DC: American Educational Research Association; 2014.
13. Lievens F, Motowidlo SJ. Situational judgment tests: from measures of situational judgment to measures of general domain knowledge. *Industrial and Organizational Psychology.* 2016;9(1): 3-22.
14. Krumm S, Lievens F, Huffmeier J, Lipnevich AA, Bendels H, Hertel G. How "situational" is judgment in situational judgment tests? *J Appl Psychol.* 2015;100:399-416.
15. Rockstuhl T, Ang S, Ng KY, Lievens F, Van Dyne L. Putting judging situations into situational judgment tests: Evidence from intercultural multimedia SJTs. *J Appl Psychol.* 2015;100:464-480.
16. Wolcott MD, Lobczowski NG, Zeeman JZ, 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.
17. Ployhart RE. The predictor response process model. In: Weekly JA, Ployhart RE, editors. Situational judgement tests: theory, measurement, and application. Mahwah, NJ: Lawrence Erlbaum; 2006. p. 83-105.
18. Pellegrino J, Chudowsky N, Glaser R. Knowing what students know: The science and design of educational assessment. Board on Testing and Assessment, Center for Education, National Research Council, Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academic Press; 2001.
19. Nichols P, Huff K. Assessments of complex thinking. In: Ercikan K, Pellegrino JW, editors. Validation of score meaning for the next generation of assessments: The use of response processes. New York, NY: Routledge; 2017. p.63-74.
20. 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.
21. Weekley JA, Ployhart RE, Holtz BC. On the development of situational judgment tests: Issues in item development, scaling, and scoring. In: Weekly JA, Ployhart, RE, editors. *Situational judgement tests: Theory, measurement, and application.* Mahwah, NJ: Lawrence Erlbaum; 2006. p. 157-182.
22. Fan J, Stuhlman M, Chen L, Weng Q. Both general domain knowledge and situation assessment are needed to better understand how SJTs work. *Industrial and Organizational Psychology.* 2016; 31(1):43-47.
23. Harris AM, Siedor LE, Fan Y, Listyg B, Carter NT. In defense of the situation: An interactionist explanation for performance on situational judgment tests. *Industrial and Organizational Psychology.* 2016;31(1):23-28.
24. Harvey RJ. Scoring SJTs for traits and situational effectiveness. *Industrial and Organizational Psychology.* 2016;31(1), 63-71.
25. Melchers KG, Kleinmann M. Why situational judgment is a missing component in the theory of SJTs. *Industrial and Organizational Psychology.* 2016;31(1):29-34.
26. McDaniel MA, Har an NS, Whetzel D, Grubb III WL. Situational judgment tests, response instructions, and validity: A meta-analysis. *Personnel Psychology.* 2007;60(1):63-91.
27. McDaniel MA, Nguyen NT. Situational judgment tests: A review of practice and constructs assessed. *International Journal of Selection and Assessment.* 2001;9(1):103-113.

28. Lievens F. Construct-driven SJTs: toward an agenda for future research. *Int J Testing.* 2017;17(3):269-276.

29. Wolcott MD. The situational judgment test validity void: Describing participant response processes [doctoral dissertation]. University of North Carolina; 2019 Retrieved from ProQuest Dissertations & Theses. Accession Number 10981238.

30. Kim SS, Kaplowitz S, Johnston MV. The effects of physician empathy on patient satisfaction and compliance. *Eval Health Prof.* 2004;27(3):237-251.

31. Riess H. The science of empathy. *J Patient Exp.* 2017;4(2):74-77.

32. Hojat M. *Empathy in patient care: Antecedents, development, measurement, and outcomes.* New York, NY: Springer; 2007.

33. Fjortoft N, Van Winkle LJ, Hojat M. Measuring empathy in pharmacy students. *Am J Pharm Educ.* 2011;75(6):109.

34. Nunes P, Williams S, Sa B, Stevenson K. A study of empathy decline in students from five health disciplines during their first year of training. *Int J Med Educ.* 2011;2:12-17.

35. Quince T, Thiemann P, Benson J, Hyde S. Undergraduate medical students' empathy: current perspectives. *Adv Med Educ Pract.* 2016;7:443-455.

36. Tamayo CA, Rizkalla MN, Henderson KK. Cognitive, behavioral, and emotional empathy in pharmacy students: Targeting programs for curriculum modification. *Front Pharmacol.* 2015;7:Article 96.

37. Leighton JP. *Using think-aloud interviews and cognitive labs in educational research: Understanding qualitative research.* New York, NY: Oxford University Press; 2017.

38. Wolcott MD, Lobczowski NG. Using cognitive interviews and think-aloud protocols to understand thought processes. *Curr Pharm Teach Learn.* 2021;13(2):181-188.

39. Willis GB. *Analysis of the cognitive interview in questionnaire design: Understanding qualitative research.* New York, NY: Oxford University Press; 2015.

40. Siebert CF, Siebert DC. *Data analysis with small samples and non-normal data* (1st ed.). New York, NY: Oxford University Press; 2018.

41. Siegel S, Castellan NJ. *Nonparametric statistics for the behavioral sciences.* New York, NY: McGraw-Hill; 1988.

42. Chi MTH. Quantifying qualitative analyses of verbal data: a practical guide. *J Learn Sci.* 1997;6:271-315.

43. Chessa AG, Holleman BC. Answering attitudinal questions: modelling the response process underlying contrastive questions. *Applied Cognitive Psychology.* 2007;21:203-225.

44. Tourangeau R, Rips LC, Rasinski K. *The psychology of survey response.* Cambridge, MA: Cambridge University Press; 2000.

45. Saldana J. *The coding manual for qualitative researchers.* Thousand Oaks, CA: SAGE; 2016.

46. Merriam S, Tisdell EJ. *Qualitative research: A guide to design and implementation.* San Francisco, CA: Jossey-Bass; 2016.

Appendix 1. Summary of Empathy SJT Item Content, Alternatives Scenarios, and SJT Item Sample

Item Label	Context	Empathy Component	Item Summary	Alternative Scenario
CH1	Healthcare	Cognitive	Patient complains that the doctor never listens to them	Friends complains someone never listens to them
CH2	Healthcare	Cognitive	Trouble getting a medication history from a pharmacist	Trouble getting information for an event
CH3	Healthcare	Cognitive	Suspect a patient is lying about their diabetes management	Suspect a friend is lying about their activities
AH1	Healthcare	Affective	Patient discusses the recent loss of a loved one	Friend discusses the recent loss of a loved one
AH2	Healthcare	Affective	Nurse asks you to discuss a medication error with family	Manager discussing a wrong food order
AH3	Healthcare	Affective	Family gets upset while you review their chemotherapy	Family member gets upset during a conversation
CN1	Non-healthcare	Cognitive	Friend is going to use medications to help them study	Patient is going to use an unauthorized medication
CN2	Non-healthcare	Cognitive	Person asks you to cut in line at a store when you're late	Patient wants to cut in line at the clinic
CN3	Non-healthcare	Cognitive	Your family questions your sibling's relationship status	Patient's family is questioning their status
AN1	Non-healthcare	Affective	Parent quickly becomes upset at a grocery store	Patient quickly becomes upset in the hospital
AN2	Non-healthcare	Affective	Relative is upset about difficulty conceiving	Patient is upset about difficulty conceiving
AN3	Non-healthcare	Affective	Best friend is planning to drop out of college	Patient is planning to stop therapy / treatment

ITEM: CN2

You go to the store to pick up a few things you forgot for a presentation. While standing in line at checkout, someone approaches you and asks if they can cut in front of you. However, there are already 5 people behind you. They mention that their children are at home sick and they are trying to get back as quickly as possible. Letting the person go in front of you will definitely make you late for your presentation.

Rank each of the following response options based on how you SHOULD respond to the scenario. Use 1 to indicate the MOST appropriate response and 5 to indicate the LEAST appropriate response. There can be no ties or duplicates.

_____ Ask the people behind you if they would mind having the person go in front of you.

_____ Acknowledge their situation and let them go in front of you.

_____ Tell them no and that they need to get in line like all the others.

_____ Ask the person what is wrong with their child and determine whether they cut can based on their response.

_____ Tell them that you are also in a rush and ask if they could cut in front of the person behind you.

A = Affective Empathy; C = Cognitive Empathy; H = Healthcare Setting; N = Non-Healthcare Setting; 1, 2, 3 = Item Number