BRIEF

The Relationship Between a Multiple Mini-Interview and Situational Judgment Test for Admissions

Wendy C. Cox, PharmD, a Michael Wolcott, PharmD, PhD, b Faustina Hahn, PharmD, a Jacqueline E. McLaughlin, PhD, MS a

a The University of North Carolina at Chapel Hill, UNC Eshelman School of Pharmacy, Chapel Hill, NC
b High Point University, School of Dental Medicine and Oral Health, High Point, NC

Corresponding Author: Wendy C. Cox, PharmD, Associate Professor, Division of Practice Advancement and Clinical Education, Associate Dean for Admissions and Accreditation, UNC Eshelman School of Pharmacy, The University of North Carolina at Chapel Hill, Chapel Hill, NC e-mail: wendy_cox@unc.edu Phone: 919-966-9990

Submitted February 7, 2022; accepted May 9, 2022; ePublished August 2022.

Objective. To explore the relationship between a multiple mini-interview (MMI) and situational judgment test (SJT) designed to evaluate non-academic constructs.

Methods. A 30-question ranked-item SJT was developed to test three constructs also measured by MMIs during the School’s admissions process. First-year pharmacy students were invited to complete the SJT, administered via Qualtrics, in Fall 2020. One hundred and four students took the SJT (82.5% response rate) with 97 (77% of possible participants) having MMI scores from the admissions process. Descriptive statistics (mean ± standard deviation), Concordance analysis, Spearman’s correlation, Cronbach’s alpha, and principal components analysis were used to explore the psychometric properties of the SJT and its relationship to MMI scores.

Results. Seventy-four percent of students identified as female (n=72) and 11.3% identified with an underrepresented racial identity (n=11). The average age was 21.8±2.1 years. Students averaged 85.5%±3.1 (out of 100 points) on the SJT and 6.1±1.0 on the MMI (out of 10 points). Principal components analysis indicated that the SJT lacked construct validity and internal reliability. However, reliability of the entire SJT instrument provided support for using the total SJT score for analysis (α=63). Correlations between total SJT and MMI scores were weak (r p <0.29).

Conclusion. Results of this study suggest that an SJT may not be a good replacement for the MMI to measure distinct constructs during the admissions process. However, the SJT may provide useful supplemental information during admissions or as part of formative feedback once students are enrolled in a program.

Keywords: situational judgment test, multiple mini-interviews, admissions, assessment

INTRODUCTION

Identifying and evaluating non-academic constructs in students applying to health sciences programs is a primary goal of many admissions committees. Non-academic constructs (also called social and behavioral constructs), such as empathy and integrity, are paramount to successful practice as a healthcare professional. Therefore, programs are exploring ways to accurately assess these constructs in applicants and in students as they progress through programs.1-5

A situational judgment test (SJT) is an assessment technique that has gained popularity in health sciences schools over the past several years as a method to measure social and behavioral aspects of students.6,7 An SJT is a written assessment tool where a case or scenario is presented and the test taker must rate the appropriateness of various responses to the scenario, written and validated by subject matter experts. Test-takers may be asked to rank order responses from most appropriate to least appropriate or to select the best response. The time to administer an SJT is minimal, and it only requires one person to administer the test to a group; however, an SJT is time consuming to develop and there is not one commercially available test in health professions education. There are also different ways to design SJTs, resulting in varying outcomes and difficulty identifying and evaluating one construct of interest.6-15

Because of the ease of administration, SJTs may be less resource-intensive than administering multiple mini-interviews (MMIs), used for similar purposes in evaluating non-academic constructs. MMIs were originally developed to evaluate non-academic qualities of applicants to medical residency programs and have gained popularity in health professions education.16-21 In these evaluations, test takes are presented with a written case and have a few minutes to formulate a response. Then they enter a room with an evaluator and provide thoughts on the case. The purpose is for the
evaluator to be able to identify and evaluate the construct of interest that the MMI is targeting. This method requires a lot of resources to plan and implement.\textsuperscript{16,22,23}

At the University of North Carolina (UNC) Eshelman School of Pharmacy, we have used MMIs as part of our admissions process since 2015 to measure non-academic constructs.\textsuperscript{23-25} The MMI has been a helpful tool in our admissions process; however, it is resource intensive, specifically the faculty and staff time required to plan and administer the MMI. Our research team hypothesized that an SJT may be used to accurately identify the constructs of interest measured in our MMI, which could be used in admissions using less resources. The objective of this study was to explore the relationship between a MMI and an SJT designed to evaluate the same non-academic constructs.

METHODS

A 30-question ranked-item SJT was developed by the authors to test three constructs also measured by MMIs during the School’s admissions process: adaptability, empathy, and integrity (Appendix 1). Prior to the administration of the SJT, the scoring system was based on the responses from eight experts. The experts were faculty and post-doctoral fellows who were practicing pharmacists. The SJT was piloted with postdoctoral fellows to verify the functionality and requests for minor edits to improve the readability.

First-year pharmacy students were invited to complete the 30-item SJT in Fall 2020. The SJT was administered via Qualtrics at the end of an orientation session. Participants were instructed to rank the five answer options in order of what they would do from most likely to least likely. Two of the questions from the empathy section were removed due to technical errors within Qualtrics. This led to a total of 28-questions on the SJT (eight empathy questions and ten adaptability and integrity questions each). The SJT score was paired with the respective MMI score. The MMI data were extracted from the admissions office, which had the scores for each of the MMI stations (1=poor to 10=exceptional). Data from the MMI model at the School have previously been shown to have strong construct validity and high internal consistency.\textsuperscript{23} The admissions data also included demographic data such as age, gender, and race/ethnicity.

Data were analyzed using descriptive statistics, nonparametric statistical tests, and psychometric analyses. Specifically, concordance analysis, Cronbach alpha, principal components analysis (PCA), correlation, and linear regression were used to explore SJT psychometric properties and its relationship to the MMI scores. For PCA, Varimax rotation with Kaiser rule (ie, Eigenvalue > 1) was used to identify and retain factors. Continuous data are presented as mean and standard deviation, and categorical data are presented as frequency (percentage) or median (interquartile range). Analyses were done within R, SPSS (version 26 – IBM), and Excel (2016 – Microsoft). This study was approved via expedited review by the University of North Carolina at Chapel Hill Institutional Review Board.

RESULTS

One hundred and four students took the SJT (82.5% response rate) with 97 participants having MMI scores from the most recent admission cycle (some had MMIs from an Early Assurance Admissions Program which were not used in this study because they were over one year old and not identical to the MMIs administered to the rest of the class). Seventy-four percent of students identified as female (n=72) and 11.3% identified with an underrepresented racial identity (n=11) with an average age of 21.8±2.1 years (Table 1).

Students averaged 85.5%±3.1 (out of 100 points) on the SJT and 6.1±1.0 on the MMI (out of 10 points). PCA indicated that the SJT lacked construct validity (ie, factored into more than the three constructs intended) and internal reliability (ie, α<.4 for each construct). Multiple PCAs were conducted after excluding items with low concordance (W<0.6), yet the SJT scores continued to factor into more than three constructs as intended. However, the reliability of the entire SJT instrument provided support for using the total SJT score for exploratory analysis (α=0.63). Correlations between total SJT and MMI scores were weak (r<0.29). Correlations of various scoring combinations (ie, top choice, top two choices, bottom two choices) were also conducted and resulted in weak to negligible correlations (Spearman’s Rho range = -0.12–0.10) (Table 2).

When analyzed by demographic groups, notable differences were found by race and gender identity for the total SJT score. Female-identifying participants scored higher on SJT items than male-identifying participants by 2.67 points (p < 0.001) when controlling for all other variables in the model. Additionally, having an underrepresented racial identity (eg, Black, Latinx, Native American) was associated with a 1.83 decrease in SJT score (p=0.03), controlling for all other variables in the model (Table 3).

DISCUSSION

More health sciences programs are using SJTs in their admissions process. It is important to understand how this assessment approach can be optimized in admissions and if it is measuring the intended constructs of interest. In this study, the total MMI score was associated with overall performance on the SJT, although the correlation was weak. This
suggests the assessments used in this study may be measuring similar constructs; however, they may be accomplishing that in a different way or providing different insights that require further exploration. For example, is there an impact of providing potential response options in SJTs that greatly differ from the responses generated by participants during an MMI and then interpreted and evaluated by a rater. This research had differing results from other SJT publications in pharmacy and health professions education—notable differences is that this SJT used a ranking response selection, whereas others often use a technique where examinees rate each response in terms of appropriateness (1= inappropriate response to 5= highly appropriate response). This may have consequences to consider in future design of SJTs; however, it should be noted that assessment methods with bespoke designs, such as MMIs and SJTs, generate data unique to the program and results should be interpreted accordingly.

An important demonstration from this research is the complexity in designing an SJT with high reliability and construct validity. The SJT created in this study failed to load into three distinct factors and demonstrated low internal consistency for each construct. This illustrates there are many factors that can influence participant response selections, which then influence performance and reliability. The construct-driven approach that concentrates on a theoretical focus on what is to be measured is often described as the optimal approach to SJT design; however, this process can be resource and time intensive and often results in a limited number of items being generated.\(^\text{10,11}\) We also illustrated when measuring multiple constructs, it can be difficult to create SJT items that readily distinguish one construct from another, which has been demonstrated with other assessment approaches like MMIs.\(^\text{25}\) The time required to develop and design SJTs can be significant. It is estimated that it took 30-40 hours to develop this SJT plus the additional time required to pilot the test with experts and refine it. In our experience, it does not take as much time to develop MMI scenarios (approximately 10 hours) but takes much longer to administer them to candidates each interview day. The time spent in SJT development is upfront, whereas most of the time required for MMIs is in administering them. Design is a particular challenge that others must be aware of in this process as it can limit the utility of the finalized instrument.

Another insight from this research was the subgroup analysis, which suggests that there may be issues in SJT fairness based on key demographic characteristics. Fairness is a critical aspect of admissions practices to ensure equitable access for all candidates. An advantage of SJTs has been evidence of enhanced fairness in scoring practices; in other words, groups are not disadvantaged based on their gender, racial identity, and other demographic aspects.\(^\text{26}\) However, the SJT in this study had statistically significant differences in performance based on gender and racial identity, which suggests there may be fairness concerns, and this may be influenced by item design. Further research would be needed to determine if the instrument accurately identified actual differences or to identify which items were biased in some way. Overall score differences ranged between 1-3 points between the subgroups, which may not have practical significance if it was determined not to impact admissions decisions. It is also important to consider that SJTs often focus on non-academic attributes (empathy, adaptability, integrity, etc.), which may be emphasized more depending on cultural background and gender identity. For example, women are often taught more explicitly about empathy and other attributes as part of societal expectations, which may influence their performance compared to men. SJT performance should be considered as highly contextual when interpreting results and clarity about whether the results may be affected by cultural or societal factors.

Part of this exploration included evaluating multiple scoring strategies (ie, single-best selection, single-worst selection, etc.) and their correlation with MMI scores to determine if the SJT had a value beyond identifying those with the highest standing on the construct. In other words, SJTs are often used to distinguish top performing candidates and those more desirable for admissions. We considered whether the SJT may have value in identifying learners who instead would be at risk for not successfully completing the program rather than identifying those who are the optimal fit. For example, we investigated the correlation between MMI score and those who selected the worst option as the best (ie, their 1st ranked response was a 5th ranked response on the key). In this research, we did not identify any other patterns of scoring that may improve the correlation. However, it also illustrates that SJT performance and psychometrics may be highly dependent on scoring practices, which has been demonstrated in previous work in medical education.\(^\text{9}\)

There are several limitations to this study. First, this study was limited to one institution and had a small sample size, especially those with underrepresented racial identities. Selection bias was present with only students accepted to our PharmD program taking the SJT. Additionally, there was more than six months between the administration of the MMI and the SJT, which could have affected the results. The scoring system also had limitations with the SJT scored in a ranked system versus a single item choice. There could also be bias with the SJT key development as different combinations of experts (hospital, community, academic pharmacists) could potentially influence the key. Furthermore, people perceive conflicts and situations differently in certain contexts, which could have affected the key and the responses. Lastly, the key may be oriented to what people believe is right as opposed to what would be the best response to a situation/conflict.
CONCLUSION

The results of this study demonstrated that the SJT lacked construct validity and the correlation between the SJT and MMI scores were weak. Given these results, an SJT may not be a good replacement for the MMI to measure distinct constructs during the admissions process; however, it may provide useful information in addition to the MMI during admissions or as part of formative feedback once students are enrolled in a program. Future research should explore the aspects of SJT design (eg, item development) and subsequent impact of using the SJT as a formative and longitudinal assessment strategy in the health professions.

REFERENCES


Table 1. Demographic Data of Study Participants

<table>
<thead>
<tr>
<th>Gender</th>
<th>Students (n=97)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>25 (25.8%)</td>
</tr>
<tr>
<td>Female</td>
<td>72 (74.2%)</td>
</tr>
<tr>
<td>Race</td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>72 (51.1%)</td>
</tr>
<tr>
<td>Asian/Asian American*</td>
<td>25 (25.8%)</td>
</tr>
<tr>
<td>Underrepresented Racial Minority+</td>
<td>11 (11.3%)</td>
</tr>
<tr>
<td>Age</td>
<td>21.7 ± 2.1</td>
</tr>
</tbody>
</table>

*Includes South Asian, Cambodian, Chinese, Filipino, Japanese, Korean, Malaysian, Pakistani, Vietnamese, and participants with mixed races with the aforementioned groups

+ includes Hispanic, Cuban, Mexican, Puerto Rican, South American, Dominican, Native American, Black, and participants with mixed races with the aforementioned groups

Table 2. Median and Interquartile Ranges of the SJT Using Different Scoring Mechanisms

<table>
<thead>
<tr>
<th>Rank</th>
<th>T2</th>
<th>SB</th>
<th>T2W</th>
<th>B2</th>
<th>SW</th>
<th>SBB</th>
</tr>
</thead>
<tbody>
<tr>
<td>Empathy</td>
<td>136 (130-140)</td>
<td>12 (11-13)</td>
<td>6 (5-7)</td>
<td>2 (1-2.85)</td>
<td>11 (10-12)</td>
<td>0 (0-0)</td>
</tr>
<tr>
<td>Integrity</td>
<td>178 (172-182)</td>
<td>16 (15-17)</td>
<td>7 (6-8)</td>
<td>0 (0-1)</td>
<td>18 (16-18.75)</td>
<td>0 (0-0)</td>
</tr>
<tr>
<td>Adaptability</td>
<td>166 (162-172)</td>
<td>14 (13-15)</td>
<td>6 (5-7)</td>
<td>2 (1-3)</td>
<td>16 (15-17)</td>
<td>0 (0-0)</td>
</tr>
<tr>
<td>Total SJT</td>
<td>478 (466.5-490)</td>
<td>42 (40-44)</td>
<td>19 (17-20)</td>
<td>4 (3-5)</td>
<td>44.5 (43-47)</td>
<td>0 (0-0)</td>
</tr>
</tbody>
</table>

RANK=partial credit based on ranking compared to key (score = 0-20 per item)
T2=identified first and/or second ranked item; no specific order (score = 0-2 per item)
SB=single best option; got first ranked item as first ranked (score = 0-1 per item)
T2W=listed a 4 or 5 ranked item in their top 1 or 2 rank INCORRECTLY (score = 0-2 per item)
B2= listed a 4 or 5 ranked item in their 4 or 5 rank CORRECTLY (score = 0-2 per item)
SW=listed a 5 ranked item in their top 1 INCORRECTLY (score = 0-1 per item)
SBB=listed a 5 ranked item in their 5 rank CORRECTLY (score = 0-1 per item)

Table 3. Linear Regression Results for Predicting SJT Score from Admissions Data and Student Demographics

<table>
<thead>
<tr>
<th></th>
<th>Model 1 (All SJT Items)</th>
<th></th>
<th>Model 2 (SJT Refined)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>B (SE)</td>
<td>Beta</td>
<td>p value</td>
</tr>
<tr>
<td>Intercept</td>
<td>79.65 (3.55)</td>
<td>.39</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Female</td>
<td>2.67 (0.62)</td>
<td>.39</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>URM</td>
<td>-1.83 (0.85)</td>
<td>-.20</td>
<td>.034</td>
</tr>
<tr>
<td>Age</td>
<td>-.02 (0.13)</td>
<td>-.01</td>
<td>.90</td>
</tr>
<tr>
<td>MMI Score</td>
<td>.56 (.27)</td>
<td>.19</td>
<td>.04</td>
</tr>
<tr>
<td>Application Review Score</td>
<td>.14 (.11)</td>
<td>.11</td>
<td>.23</td>
</tr>
<tr>
<td>R²</td>
<td></td>
<td>.29</td>
<td>.23</td>
</tr>
</tbody>
</table>

SE=Standard Error
+ Model 1 has 25 survey items, after 3 items were dropped (items with p<.002)
* Model 2 has 23 survey items, after 5 items were dropped (items with p<.01)
Appendix 1. Examples of SJT Questions from Each Category

**SJT Directions:** 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.

### Empathy

You were asked by the physician to speak with a patient’s family about the upcoming chemotherapy treatment for their 8-year old son. When you start talking about the negative side effects of the drug treatment, the mother becomes visibly upset and asks you to “stop talking about this.”

- 4  Tell that patient’s mother it is hospital policy to review all of the necessary information before beginning chemotherapy and you are required to finish
- 3  Tell the physician the family refused to complete the education and became upset
- 5  Conclude the session and document education has been complete
- 2  Request to schedule a different time to continue discussing the medication when the family would be more comfortable
- 1  Ask the mother about her concerns

### Integrity

A good friend is applying for a job where you work. The hiring manager asks you about the person and shares with you their resume. You notice your friend has lied about several things on their resume, including work experience that would make them more qualified for the job than other applicants who would likely get the job over your friend.

- 3  Tell the hiring manager about the discrepancies on the resume
- 1  Approach your friend about the discrepancies on the resume
- 4  Advise the manager not to consider your friend at this time
- 5  Tell your manager that your friend is the best choice for the position
- 2  Tell your manager that you would like to not be involved in the decision

### Adaptability

You recently accepted a promotion at work. You had always wanted to work with your new supervisor, who had an excellent reputation as a mentor, and were excited about the promotion. However, a week into your new role, your supervisor left the company and you now have a new supervisor - someone you do not know at all.

- 1  Meet with the new supervisor to learn more about their expectations
- 4  Talk to someone in Human Resources to express your concerns
- 3  Continue in your role doing the job you were hired to do
- 2  Talk with a friend in the company who knows the manager to learn about their style
- 5  Go back to your previous position