

An Examination of Ethnic Subgroup Score Differences Between Different Types of Situational Judgement Tests

By: Brent Larison

Research Advisor: Dr. Mark C. Frame

Introduction & Literature Review

- Selection of personnel for organizations is an important process that all organizations must undertake carefully in order to ensure their future success (Gatewood, Field, & Barrick, 2019).
- A key component to the selection process is the use of valid instruments and measures that can accurately predict how a job applicant will actually perform on the job. One instrument that has shown promise regarding its general validity and reliability across many settings to predict job performance is the Situational Judgment Test (SJT) (Weekly & Ployhart, 2005).
- However, SJTs being used today are not a guaranteed predictor of job performance and can lead to ethnic subgroup differences (Weekly & Ployhart, 2005). This can create legally problems if they are challenged in court. However, SJTs generally outperform cognitive ability tests when it comes to legal defensibility, since cognitive ability tests often find mean subgroup differences in scores between White and African American responses at approximately one standard deviation (Chan & Schmitt, 1997). This topic brings us to the current issue, which is determining how to reduce levels of ethnic subgroup differences in scores.
- This is an important issue which must be addressed because ethnic subgroup differences give rise to adverse impact for protected groups when they are used by organizations (Weekly & Ployhart, 2005). It has been found White participants outperform Black and Asian participants in large samples on SJTs assessing different knowledge categories, ranging from customer service to management and graduate admissions (Herde, Lievens, Jackson, Shalfrooshan, & Roth, 2020).
- Given that one organizational use of SJTs is for selection, it is key to determine whether ethnic minorities and protected groups are being placed in a disadvantage when they are administered as selection criteria. This issue will likely grow in relevance due to projected number of 405 million migrants to other countries worldwide by 2050 (Lievens, Sackett, Dahlke, Oostrom, & De Soete, 2019). We now turn our attention to the literature on contributors to ethnic subgroup differences in SJT scores.

Introduction & Literature Review

Cognitive Ability

- There are many studies which have examined factors contributing to differences in scores between minority and non-minority groups, including the assessment of cognitive ability in SJTs (Dahlke & Sackett, 2017; Whetzel, McDaniel, & Nguyen, 2008). Cognitive loading of a SJT is the extent to which the test is correlated with cognitive ability (Whetzel et al., 2008).
- A meta-analysis focused primarily on the association of cognitive ability with mean subgroup differences over a diverse array of measurements including SJTs (Dahlke & Sackett, 2017). Similar to cognitive loading, cognitive ability saturation occurs when predictors are correlated with a measure of general cognitive ability. They found significant correlations between cognitive-ability saturation and subgroup mean differences for scores between both Black and White respondents and Hispanic and White respondents (Dahlke & Sackett, 2017).
- Another meta-analysis was conducted with a similar purpose, and found comparable results when focusing specifically on SJTs (Whetzel et al., 2008). Black-White differences were found to be highest when cognitive loading was highest and personality loading was lowest.
- The researchers conducting the two aforementioned meta-analyses recommend reducing factors within a SJT that lead to cognitive loading in order to decrease ethnic subgroup differences in scores.

Introduction & Literature Review

SJT Format

- There are many factors which can lead to greater differences in scores between subgroups for SJTs (Campion, Ployhart, & MacKenzie, 2014). A consistent issue occurring in the literature has been trying to find a SJT that is both valid and does not lead to greater subgroup differences (De Soete, Lievens, Oostrom, & Westerveld, 2013). It is important to consider each component during construction of an SJT due its sensitivity to cognitive loading, potential validity and reliability issues, and potential to cause adverse impact for protected groups (Campion et al., 2014).
- Other specific components which may contribute to subgroup differences are response format (e.g., rank, rate, most/least; Arthur et al., 2014), and response mediums (Lievens et al., 2019). Researchers have even found cognitively loaded audiovisual SJTs with no ethnic subgroup score differences (Lievens et al., 2019), which is a unique and interesting finding worth investigating.
- Initial studies found lower subgroup differences in scores when a video-based SJT was administered, when compared to a text-based SJT (Chan & Schmitt, 1997). Cognitive ability was less associated with video mediums. Furthermore, when reading comprehension is controlled for in text-based SJTs, it has been found to contribute to higher cognitive loading and subgroup differences in scores (Chan & Schmitt, 1997).
- Lievens and Sackett (2006) examined differences between video-based and written SJTs, and video-based SJTs measuring interpersonal ability were significantly less correlated to cognitive ability than the written SJTs. Using SJT items assessing interpersonal skills, rather than cognitive ability, appears to be a promising solution to dealing with the current dilemma. However, Lievens and Sackett (2006) were unable to examine these effects in relation to ethnic subgroup differences, since most of their participants were white.
- This leads to our current study, which will examine aforementioned components and whether they will contribute to differences in SJT scores between ethnicities.

Current Study & Hypotheses

- There is currently no definitive solution to the dilemma of differing levels of ethnic subgroup scores in SJTs. The current study will examine components previously suggested as having potential in reducing ethnic subgroup differences in SJT scores, in an effort to help diminish the possibility of adverse impact in organizational settings.

Audio-Enhanced SJTs Versus Written SJTs

- While there is ample research suggesting there are differences in ethnic subgroup scores that can be found using different formats of SJTs, including written, web, audio, and multimedia SJTs (Chan & Schmitt, 1997; Lievens et al., 2019), the research on SJT formats has used many different formats for the situational prompts or stimuli but have not isolated the unique contribution that audio information may have when added to the situational prompts and options of traditional text based SJTs.
- We will examine SJT subgroup score differences between two types of SJTs. One type will be typed and read by participants, and the other will be typed and read by an audio player. This will be different from previous studies, because we are using the Read Aloud software in Microsoft Word to create the audio-enhanced SJTs, to help remove the obstacle of reading comprehension.
- Thus, the first research question given: Will there be a difference in ethnic subgroup SJT scores that can lead to adverse impact between normal written and read and audio enhanced formats?
- *Hypothesis 1:* There will be a difference between ethnic subgroup scores for the normal written and read SJT and the audio enhanced SJT (Chan & Schmitt, 1997; Lievens et al., 2019).

Current Study & Hypotheses

Interpersonal Ability Versus Cognitive Ability

- While there is ample evidence suggesting cognitive loading significantly leads to greater ethnic subgroup differences in SJT scores (Dahlke & Sackett, 2017; Whetzel et al., 2008), more research should be conducted to compare differences in scores between SJT formats while assessing interpersonal ability and cognitive ability (Lievens & Sackett, 2006).
- While this area in the literature has been researched considerably (Dahlke & Sackett, 2017; Whetzel et al., 2008), there are no studies which have examined differences in interpersonal and cognitive ability for the types of SJTs we will be using. Our study will be unique due to using both an academic and work domain SJT to measure cognitive ability, and we will additionally use an SJT assessing ability to facilitate teamwork to represent interpersonal ability.
- Thus, the second research question proposed: Will there be a difference in ethnic subgroup SJT scores between questions assessing interpersonal ability (e.g., personality characteristics) and cognitive ability for the normal written and read SJT and the audio enhanced SJT?
- *Hypothesis 2:* There will be a difference in ethnic subgroup SJT scores between questions assessing interpersonal ability and cognitive ability for both the normal written and read SJT and the audio enhanced SJT (Lievens & Sackett, 2006).



Methodology

Participants

- Participants will consist of undergraduate psychology students enrolled in a university in the Southeastern United States and participants will also include those from the Amazon Mechanical Turk (M-Turk). Psychology student participants will be compensated with course credit for completion of the study, and M-Turk participants will be compensated through monetary payment.
- The study is a between-subjects design and cross-sectional, with participants placed randomly into one of two groups for the two different formats of SJTs. One group will complete the normal, written and read SJTs, and the other will complete the audio enhanced SJTs.

Measures

- The measures will include three SJTs. One will be non-cognitive and assess interpersonal ability, and the other two will be used to assess cognitive ability. These three SJTs will be administered to every participant.
- The control group will complete the three SJTs in the normal, web-read format, and the experimental group will complete the three in the audio enhancement format.
- An example item from one of the SJTs is:
An exam is approaching in one of your classes. The class has been difficult but interesting for you so far. In studying for your exam, you should:
 - A. carry books with you to various places and try to study in-between other obligations
 - B. find that often times there are too many interruptions to study consistently before an exam
 - C. want to start studying early but find that other things end up getting done instead
 - D. schedule blocks of time to study in one location a week or two before the exam

Methodology

Measures (continued)

- The SJTs' scenario questions and response choices will be read and highlighted by the Read Aloud software in Microsoft Word to create the audio enhanced SJTs. Scenarios and response choices will be recorded through a video of the software highlighting and reading the words.

Academic-Oriented Situational Judgment Test

Each of the following sections presents hypothetical problem situations and asks you how you would respond. Problems in Section I are hypothetical academic situations, and problems in Section II are hypothetical job situations. Each problem has four to five alternative actions that

- Two SJTs will be used to assess cognitive ability. One will focus on an academic domain, and the other will focus on the workplace domain. These two SJTs are appropriate for assessing cognitive ability because judgement and decision making in school or work settings in general can be dependent upon cognitive ability and less dependent upon non-cognitive abilities such as personality and interpersonal ability.
- One SJT will be used to assess interpersonal ability, and it is composed of ten items which involve judgement in facilitation of teamwork activities. This SJT is appropriate for assessing interpersonal ability due to its use of judgement involving social and interpersonal situations, and cognitive loading will be less likely to be at a similar level as the two previous SJTs.
- Each of the SJTs will use a rank order response format, where participants will rank from best to worst each of the given responses for each scenario.

Methodology

Procedure

- Participants will first complete the academic domain SJT assessing cognitive ability. Second, they will complete the SJT assessing interpersonal ability. Lastly, they will complete the workplace oriented SJT assessing cognitive ability. The three SJTs will be administered through Qualtrics. Participants will be able to complete the survey online at a time of their choosing for the period of months it will be open.

Data Analysis

- Data screening procedures will be conducted prior to hypothesis testing. Once data screening is complete, the first hypothesis will be examined by obtaining the mean difference score between those who indicated themselves as African/African American and Caucasian/White for both SJT format groups.
- A *t*-test will then be conducted to determine if there is a significant difference between the mean difference scores between the two groups. Effect size will be calculated additionally for this procedure.
- The second hypothesis will be analyzed through a two-way analysis of variance (ANOVA). The first independent variable will be the ability type assessed by the SJT (i.e., cognitive versus interpersonal), and the second independent variable will be SJT format (i.e., normal written and read format versus audio enhanced SJT). The dependent variable will be the mean difference scores between the two racial groups.
- If there is a significant finding using this analysis, post hoc comparisons through the use of Tukey's HSD will be conducted to determine which independent variables were found to be significantly different from one another in ethnic subgroup differences.

Discussion

- The purpose of the present study is to examine the components previously suggested as having potential in reducing ethnic subgroup differences in SJT scores, in order to help find ways of diminishing the possibility of adverse impact in organizational settings when SJT scores are used.
- While there are many studies which have examined these aspects in the current study (Weekley & Ployhart, 2005; Whetzel et al., 2008), no study to my knowledge has examined this particular format of audio-enhanced SJTs, in addition to assessing both cognitive ability and interpersonal ability SJTs.
- Since reading comprehension has been linked to cognitive ability (Chan & Schmitt, 1997), removing this barrier with the use of the audio enhanced SJT may indeed lead to useful and unique findings.
- Limitations include the study's sample being mostly college students, which may not be representative of the population being administered SJTs in work settings. Additionally, a limitation is the research design, which is between-subjects and cross-sectional. Participants completing the normal written and read SJT will be different than those who complete the audio enhanced SJT, potentially biasing the results.
- Statistical conclusion validity is also at risk due to the inability to predict how many African/African American individuals will participate in the study; a low participation rate from this group would harm the overall statistical power of the study.

References

- Campion, M. C., Ployhart, R. E., & MacKenzie, W. I. (2014). The state of research on situational judgment tests: A content analysis and directions for future research. *Human Performance, 27*(4), 283–310.
- Chan, D., & Schmitt, N. (1997). Video-based versus paper-and-pencil method of assessment in situational judgment tests: Subgroup differences in test performance and face validity perceptions. *Journal of Applied Psychology, 82*, 143–159.
- Dahlke, J. A., & Sackett, P. R. (2017). The relationship between cognitive-ability saturation and subgroup mean differences across predictors of job performance. *Journal of Applied Psychology, 102*, 1403–1420.
- De Soete, B., Lievens, F., Oostrom, J., & Westerveld, L. (2013). Alternative predictors for dealing with the diversity–validity dilemma in personnel selection: The constructed response multimedia test. *International Journal of Selection and Assessment, 21*, 239–250.
- Gatewood, R. D., Field, H. S., & Barrick, M. R., 2019. *Human resource selection* (9th ed.). New York, New York, USA: Wessex Press.
- Herde, C. N., Lievens, F., Jackson, D. J. R., Shalfrooshan, A., & Roth, P. L. (2020). Subgroup differences in situational judgement test scores: Evidence from large applicant samples. *International Journal of Selection and Assessment, 28* (1), 45-54.
- Lievens, F., & Sackett, P. R. (2006). Video-based versus written situational judgement tests: A comparison in terms of predictive validity. *Journal of Applied Psychology, 5*, 1181-1188.
- Lievens, F., Sackett, P. R., Dahlke, J. A., Oostrom, J. K., & De Soete, B. (2019). Constructed response formats and their effects on minority–majority differences and validity. *Journal of Applied Psychology, 104*, 715–726.
- Weekly, J. A., Ployhart, R. E. (2005). *Situational judgement tests: theory, management and application* (1st ed.). Portland, Oregon, USA: Psychology Press.
- Whetzel, D. L., McDaniel, M. A., & Nguyen, N. T. (2008). Subgroup differences in situational judgment test performance: A meta-analysis. *Human Performance, 21*, 291–309.