

Factors Influencing Teacher Behavioral Assessment Preferences



UNIVERSITY OF MINNESOTA
Driven to Discover™

Faith G. Miller¹, Sandra M. Chafouleas², T. Chris Riley-Tillman³, & Gregory A. Fabiano⁴

University of Minnesota¹, University of Connecticut², University of Missouri³, University at Buffalo⁴

Abstract

The aim of this investigation was to examine the relationship between usability factors on a measure of assessment usability (URP-A) and teacher assessment preferences of three behavioral screening measures.

Background

Researchers interested in implementation science have long been concerned with identifying factors that promote or inhibit the implementation of specific practices: whether those practices involve treatments or assessment techniques. Within the last decade, the increased emphasis on evidence-based practice and accountability in school settings has contributed to the proliferation of research in this area. That is, not only is there interest in identifying effective practices but in determining how to implement them effectively. In this vein, social validity has long been of interest within an implementation science context. In particular, stakeholder perceptions and beliefs can have powerful impacts on implementation effectiveness (Kazdin, 1980; Klein & Sorra, 1996; Wolf, 1978). The social validity of a practice can have implications for the adoption and subsequent use of that methodology (Eckert, Hintze, & Shapiro, 1999).

Understanding factors that may influence assessment use is essential; this information can be used to ensure that an assessment has some likelihood of being adopted or to problem-solve why implementation issues have arisen. The Usage Rating Profile – Assessment (URP-A) was developed in order to understand the factors that may influence teachers' adoption of school-based assessments. Prior research on the URP-A has supported a measure with six factors that include: acceptability, understanding, feasibility, home-school collaboration, system climate, and system support (Miller et al., 2015).

Currently, little is known regarding how various factors impact teacher preferences of school-based assessments. That is, certain factors may influence preferences more than others. It is important to understand which factors are most influential in order to inform assessment selection. As such, the purpose of the present investigation was to examine the relationship between usability factors on the URP-A and rank-ordered assessment preferences.

Research Question

Which factor(s) best predict overall behavioral screening assessment preferences?

Method

Participants

A total of 115 elementary and middle school teachers completed the URP-A. The majority of respondents identified as white (98%), female (87%), elementary (71%) teachers with between 1-5 years of experience (60%). Public school settings were geographically located Connecticut, New York, and Missouri.

Measures

- *Usage Rating Profile- Assessment (URP-A; Chafouleas, Miller, Briesch, Neugebauer, & Riley-Tillman, 2012).*

The URP-A is a self-report measure for collecting information about the factors influencing use of an assessment methodology. The measure consists of 28 items to which participants respond regarding their level of agreement using a 6-point Likert scale. Factor scores were calculated using unweighted sum scores. Participants completed the URP-A in response to the following measures:

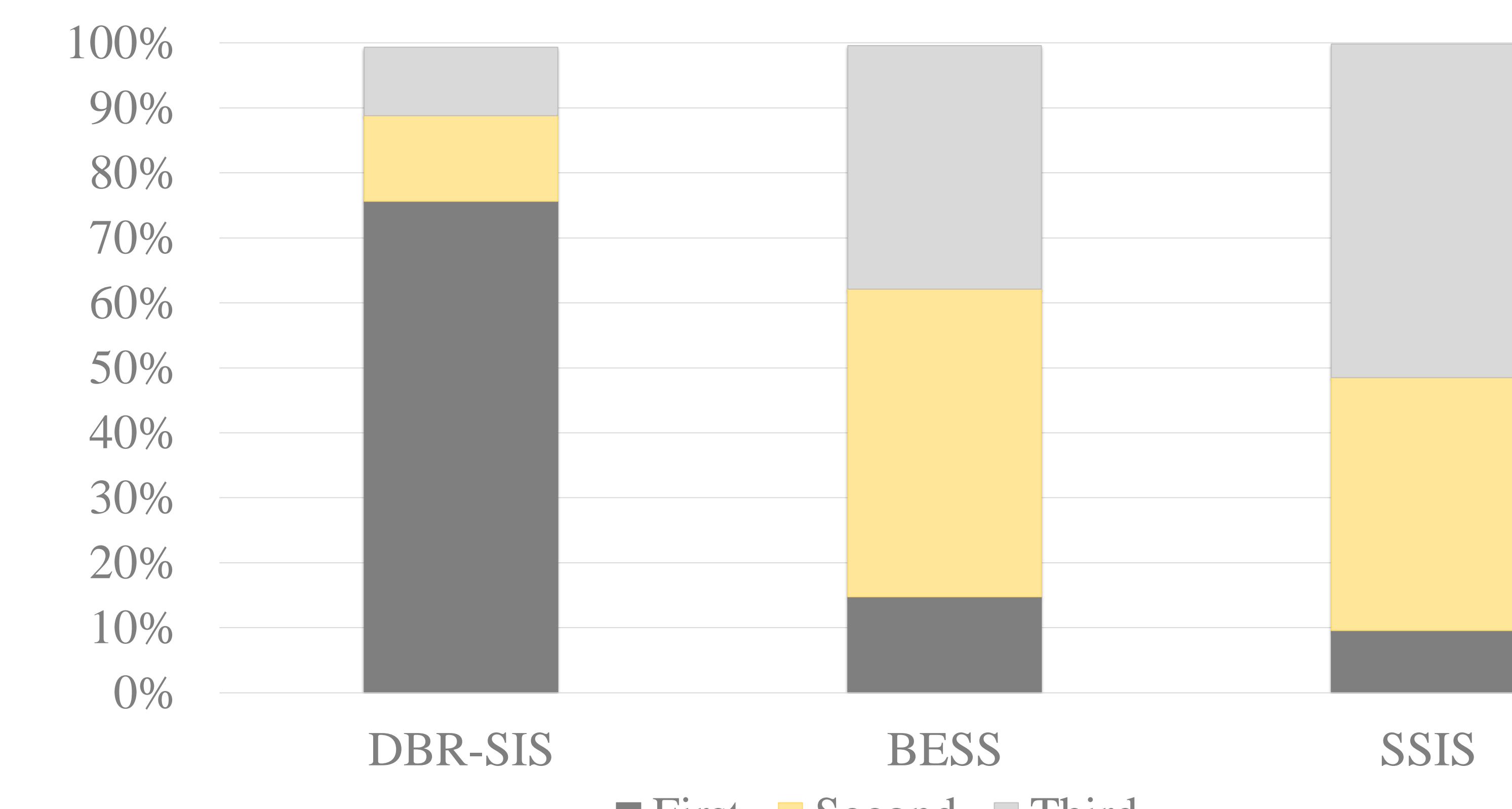
- *Direct Behavior Rating – Single Item Scale (DBR-SIS)*
DBR-SIS reflects the teacher's perception of the proportion of time a student is observed engaged in a target behavior (academic engagement, respectful, disruptive) from 0 (never) to 10 (always). Students were rated twice daily for five days.
- *Social Skills Improvement System - Performance Screening Guide (SSIS; Elliot & Gresham, 2007)*
The SSIS-PSG can be used to screen social and academic behaviors of all students in a class. This measure is comprised of four scales: Math Skills, Reading Skills, Motivation to Learn, and Prosocial Behavior.
- *Behavioral and Emotional Screening System - Teacher Form (BESS; Kamphaus & Reynolds, 2007)*
The BESS is a 27 item brief rating scale that can be useful in screening for behavioral and emotional strengths and weaknesses in children and adolescents.

Procedures

Participants completed the URP-A after administering each of the assessments over 3 time points in the 2011-12 school year. After completing the URP-A, teachers were asked to respond to the following statement: "Overall, this measure offered the best way to assess student behavior". Respondents indicated their first, second, and third choices respectively.

Results

Figure 1: Assessment Rankings



As shown in Figure 1, the majority of respondents (76%) indicated that they preferred DBR-SIS to the BESS and SSIS. Multinomial logistic regression analyses were used to predict teacher-reported assessment preferences (DBR, BESS, or SSIS) based upon responses to the URP-A. Specifically, factor scores (acceptability, understanding, feasibility, home-school collaboration, system climate, and system support) were used to predict assessment preferences. Results indicated significant improvement in the model with the inclusion of the predictors, $\chi^2 (36) = 71.86$, $p < .001$. The model also evidenced good fit according to the Pearson criterion, $\chi^2 (192) = 135.48$, $p = .99$, Nagelkerke $R^2 = .61$. The full model exhibited a correct classification rate of 86%. Likelihood ratio tests (adjusted for inflated family-wise error rate) showed that the understanding factor rating for DBR-SIS contributed significantly to the prediction of assessment preference ($p < .001$).

Conclusions

Results from this study highlight the importance of understanding the social validity of teacher-completed behavioral assessments. The understanding factor incorporates items related to the respondent's knowledge regarding the assessment and associated procedures in using it. Despite completing training in the assessment procedures for each assessment, the sample reported the DBR-SIS to be more easily understood and interpretable than the other assessments. Developers may want to attend to this feature of assessments in order to enhance social validity.