

Factorial Invariance of the Usage Rating Profile for Supporting Students' Behavioral Needs (URP-NEEDS)

Amy M. Briesch
Northeastern University

Sandra M. Chafouleas, Dakota W. Cintron,
and D. Betsy McCoach
University of Connecticut

Previous research has suggested that multiple factors beyond acceptability alone (e.g., feasibility, external supports) may interact to determine whether consumers will use an intervention or assessment in practice. The Usage Rating Profile for Supporting Students' Behavioral Needs (URP-NEEDS) was developed in order to provide a simultaneous assessment of those factors influencing use of a particular approach to identifying and supporting the social, emotional, and behavioral needs of students. As the measure was intended for use with a range of school-based stakeholders, a first necessary step involved establishing the measurement invariance of the instrument. Participants in the current study included 1,112 district administrators, 431 building administrators, and 1,355 teachers who were asked to identify the approach used within their school district to identify and support the social, emotional, and behavioral needs of students, and then to complete the URP-NEEDS in reference to this identified approach. Results supported the measurement invariance of the URP-NEEDS across stakeholder groups. In addition, measurement invariance was found across self-identified approaches to social, emotional, and behavioral risk identification within the district administrator and teacher groups.

Impact and Implications

Findings of the current study indicate that the URP-NEEDS may be useful in drawing comparisons across groups of stakeholders (e.g., teachers, administrators) regarding perceived facilitators and barriers to implementing a particular approach to identifying and supporting the social, emotional, and behavioral needs of students. Such information could be used to either inform changes to procedures (e.g., to enhance feasibility) or systems (e.g., to strengthen consultative supports), or to more efficiently target efforts to enhance user understanding and motivation in order to maximize future potential for usage.

Keywords: measurement invariance, social-emotional assessment, social-emotional risk, usability, implementation science

Although research has demonstrated the effectiveness of school-based interventions in improving mental health outcomes for youth (Rones & Hoagwood, 2000), one of the chief problems that schools continue to struggle with is the appropriate and proactive identification of students at-risk for social, emotional, and behavioral (SEB) problems. Historically, these students came to the

attention of school-based mental health professionals when their behavior exceeded a threshold of teacher tolerability, which often resulted in a referral for special education eligibility. One of the most significant criticisms of this “refer-test-place” model of service delivery was the fact that extended periods of time would pass before struggling students received the supports that they needed (Bradley, Danielson, & Doolittle, 2007). In fact, whereas most teacher referrals for academic concerns occur before the third grade, referrals for behavioral concerns are more often delayed until middle or even high school (Walker, Nishioka, Zeller, Severson, & Feil, 2000). This delay in identification and potential service provision can be highly problematic, given the tendency for behavior problems to become increasingly resistant to efforts to intervene over time (Bradley, Doolittle, & Bartolotta, 2008).

Beginning in the late 1970s, however, increased emphasis was placed on the need for preventive strategies to reduce special education eligibility rates, and one of the primary vehicles for doing so became the use of multidisciplinary consultation teams (Chalfant, Pysh, & Moultrie, 1979). The structure and purpose of these teams have shifted somewhat over the years; however, the

This article was published Online First February 28, 2019.

Amy M. Briesch, Department of Applied Psychology, Northeastern University; Sandra M. Chafouleas, Dakota W. Cintron, and D. Betsy McCoach, Department of Educational Psychology, University of Connecticut.

Preparation of this article was supported through funding by the Institute for Education Sciences, U.S. Department of Education (R305A140543). Opinions expressed herein do not necessarily reflect the position of the U.S. Department of Education, and such endorsements should not be inferred.

Correspondence concerning this article should be addressed to Amy M. Briesch, Department of Applied Psychology, Northeastern University, 404 International Village, Boston MA 02115. E-mail: a.briesch@northeastern.edu

unifying goal of these teams has been to increase teachers' efficacy to address student concerns in the classroom as opposed to referring the problems externally (Flugum & Reschly, 1994). Although there have been relatively few empirical studies examining the effectiveness of such multidisciplinary teams, evidence suggests that their use may lead to reductions in the number of special education referrals as well as improved academic and behavioral outcomes (Burns & Symington, 2002). Despite the promise of multidisciplinary teams, however, research has unfortunately shown that everyday practices may not always align with best practice. A study by Crone et al. (2016) found that middle school teams underutilized the potential data sources available to them and fewer than 40% of team meetings ended in actionable decisions. Furthermore, many teachers have reported not finding value in the strategies offered through team meetings and ultimately seeing the team simply as a necessary hurdle toward evaluation for special education (Lee-Tarver, 2006; Slonski-Fowler & Truscott, 2004). For example, only 60% of elementary school teachers in a survey by Lane, Pierson, Robertson, and Little (2004) rated the intervention procedures recommended by the team as acceptable and low acceptability ratings were also found for intervention outcomes. Unfortunately, in some cases, negative experiences with multidisciplinary consultation teams—such as feeling that their opinions were undervalued or that the support provided was insufficient—may lead teachers to avoid bringing new cases forward (Slonski-Fowler & Truscott, 2004).

One alternative to a referral-based system of identification and subsequent care that has received increased attention in recent years is the use of universal SEB screening. Universal screening involves conducting proactive assessments of all students in the school in order to identify those students demonstrating some level of risk. For example, teachers might be asked to complete a brief rating scale (e.g., Behavioral and Emotional Screening System [BESS]; Kamphaus & Reynolds, 2015) for each of their students, with those students whose scores exceed a predetermined cut-off referred for additional assessment and/or intervention. Such an approach directly addresses the concerns raised regarding teacher referral, in that one common assessment process is used proactively to identify and address problems early on. Despite the advantages of school-based SEB screening, however, research suggests that few schools are actually implementing such a proactive assessment approach. For example, findings from a nationally representative survey found that about one third of building administrators across elementary and secondary grade levels reported use of any SEB screening assessment, and that only one third of these respondents reported use in universal screening purposes (Briesch, Chafouleas, Dineen, & McCoach, 2019). These results stand in stark contrast to findings related to universal screening for academic concerns, with 80%–90% of administrators reporting use of screening tools for this purpose (Briesch et al., 2019; Bruhn, Woods-Groves, & Huddle, 2014).

Understanding the Factors That Influence Usage

Work to date has helped to elucidate a range of potential obstacles to the use of proactive models of identification and behavioral support; however, the challenge in understanding why a given approach is or is not carried out in a particular situation is

that there is rarely one simple explanation. Multiple factors may work in combination to determine whether someone chooses to adopt something, and the constellation of influences may also be unique to the individual or particular innovation (Kurtz, 1990). Although various theoretical models exist (e.g., Durlak & DuPre, 2008; Perepletchikova & Kazdin, 2005) many acknowledge that use may best be understood by simultaneously considering variables at the levels of the individual (e.g., how willing the implementer is to try something new), innovation (e.g., how much time it will take to implement), and environment (e.g., whether there is sufficient support for implementation; Sanetti & Kratochwill, 2009).

Within the literature, it has been most typical for researchers to create “home-grown” measures to assess the social validity of specific innovations (Proctor et al., 2011). In contrast, the Usage Rating Profile (URP) measures were designed to compare the usage (i.e., both initial use and sustained practice; Briesch, Chafouleas, Neugebauer, & Riley-Tillman, 2013) of different approaches and are not tied to a particular intervention (URP-Intervention Revised; Chafouleas, Briesch, Neugebauer, & Riley-Tillman, 2011) or assessment (URP-Assessment; Chafouleas, Miller, Briesch, Neugebauer, & Riley-Tillman, 2012). Within the acceptability subscale, items assess whether an individual finds the innovation to be acceptable (i.e., fair, appropriate, effective) and is personally interested in implementing it. The two innovation-level subscales assess the degree to which potential consumers understand how to carry out the intervention or assessment procedures (i.e., understanding) and perceive the innovation to require a reasonable amount of time and resources to implement (i.e., feasibility). Finally, the three factors designed to assess environmental variables ask respondents about the degree to which collaboration and communication with families (i.e., family-school collaboration) and training/consultative support (i.e., system support) are needed to support usage, as well as whether the innovation demonstrates philosophical fit within the existing context (i.e., system climate). Through each iteration of an URP measure, standard processes were utilized to establish construct validity. First, literature was reviewed to generate potential items for inclusion. Next, content area experts across related disciplines participated in a review of item relevance, importance, and fit to the proposed factor. The resulting items were then subjected to a series of psychometric evaluations. Results of factor analytic studies have suggested that both measures may be used to provide a reliable assessment of those factors believed to influence usage at the levels of the individual (i.e., acceptability), innovation (i.e., feasibility, understanding), and environment (i.e., family-school collaboration, system climate, system support; Briesch et al., 2013; Miller, Neugebauer, Chafouleas, Briesch, & Riley-Tillman, 2013). Taken together, the body of prior work has established construct validity for the URP measures.

Although there is benefit to the use of generalized terminology for many school-based innovations, at the same time, there may also be limitation to such an approach with regard to school mental health service delivery. First, mental health service delivery is a less well-established territory for schools, meaning that knowledge, skills, and attitudes may not be as widely developed by school personnel in comparison to academic domains. This concern may be especially relevant in relation to the proactive identification of students with SEB concerns given the relatively new

emergence of focus within tiered delivery systems in schools. Second, comprehensive school mental health services often rely more heavily on connections to external partnerships, suggesting external factors may play an even more critical role in usability. For example, Aarons (2004) found that individual attitudes toward use of evidence-based practices among community behavioral health providers were influenced by organizational features. In that study, openness to innovation appeared as a particularly important part of mental health programs and organizational context.

The Usage Rating Profile for Supporting Students' Behavioral Needs (URP-NEEDS; Chafouleas, Briesch, McCoach, & Riley-Tillman, 2018) was therefore created to more specifically address consumers' perceptions of usability surrounding school-based approaches to SEB risk identification and support, drawing upon past success in validating the URP suite of assessments yet expanding and modifying content to align with suggestions in the behavioral health literature. Development of the measure followed the same process used to establish construct validity with previous URP measures described above. The URP-NEEDS consists of 24 items designed to assess five underlying factors (see Figure 1). First, at the level of the individual, it has been argued that regardless of how effective or acceptable an innovation is believed to be, adoption and eventual usage are less likely to occur if key stakeholders are not oriented toward change (Aarons, 2004; Lehman, Greener, & Simpson, 2002). Therefore, acceptability was replaced by the willingness to change subscale within the URP-NEEDS, which more broadly assesses the degree to which school personnel are open and willing to try new strategies and procedures (e.g., *school personnel are willing to use new and different types of social, emotional, and behavioral strategies developed by researchers*). Second, as in the original URP, the two innovation-level subscales assess the degree to which potential consumers understand how to carry out the procedures (i.e., understanding; e.g., *school personnel are confident in their ability to carry out the social, emotional, and behavioral approach*) and perceive the SEB approach to require a reasonable amount of time and resources to implement (i.e., feasibility; e.g., *the total time required for staff to carry out the social, emotional, and behavioral approach is manageable for school personnel*). Third, the URP-NEEDS includes two factors designed to assess environmental variables. Consistent with the URP, the family-school collaboration subscale asks respondents about the degree to which collaboration and communication with families is needed to support usage (e.g., *parental collaboration is needed in order to implement this social, emotional, and behavioral approach*). The system support subscale was renamed exter-

nal supports within the URP-NEEDS given expansion of content to include the need for not only consultative but also community support (e.g., *connections to community agencies are necessary to implement the social, emotional, and behavioral approach*). The one subscale that was not retained in the URP-NEEDS was system climate, which was originally designed to assess the degree to which a classroom innovation was believed to be compatible with the larger school context. These items may have been less relevant given that approaches to SEB risk identification are typically not limited to specific classrooms but conducted across the organization (see the Appendix for a list of all items).

Purpose of Study

The URP-NEEDS measure was built from the base of prior work that has established the constructs measured by the URP to be valid indicators of usability, across different school-based innovations. The same procedures for engaging in expert validation were used in adapting the URP measure to fit an SEB context of usability assessment. Although the URP-NEEDS was designed to assess the usability of various approaches to identifying and supporting students with SEB needs, it is also critical that such a tool can be used across multiple stakeholders, from those creating the policies (e.g., superintendents) to those putting the policies into practice (e.g., teachers). Capacity for use across multiple stakeholders is important because the successful use of a school-based approach cannot necessarily be explained by a single group of consumers. For example, if an administrator views the financial cost of an approach as a significant barrier to implementation, the approach may not even be introduced into the school setting. On the other hand, if an approach that has received administrative endorsement is seen as too time-consuming, teachers may be reluctant to employ it in their classrooms. A limitation of psychometric work to establish URP measures to date is that researchers have not included specific investigation across multiple stakeholders with intent to evaluate whether constructs hold the same meaning.

For the URP-NEEDS to be recommended for use across different stakeholders, a first necessary step is to test whether the same latent traits are being measured across groups (i.e., measurement invariance; Dimitrov, 2010; Meredith, 1993). That is, evidence needs to be provided that the constructs have the same meaning across each stakeholder group. If equivalent meaning in the latent trait scores can be established across groups, then cross-group comparisons of the resulting scale scores can be made and yield

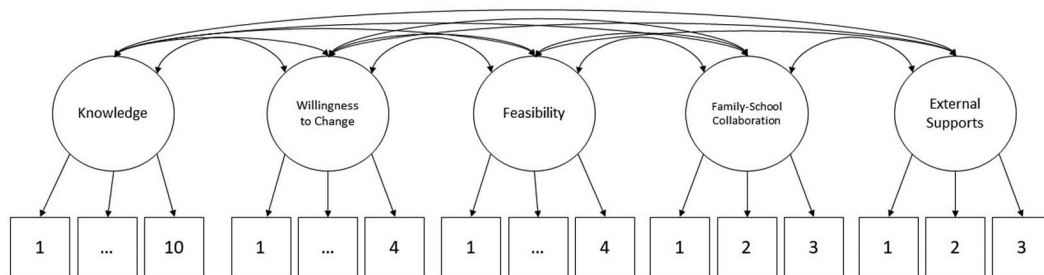


Figure 1. URP measurement model.

meaningful interpretations. Conversely, if evidence suggests that the same trait is not being measured across groups, then comparisons and analysis across groups are not supported. Therefore, the primary goal of the current study was to test the measurement invariance of the URP-NEEDS across stakeholders (i.e., district administrators, building administrators, teachers) using multiple-group confirmatory factor analysis (MG-CFA; Byrne, 2004; Jöreskog & Sörbom, 1979). These measurement invariance tests are conducted between stakeholder group (one analysis). In addition, we sought to test the measurement invariance of the URP-NEEDS across different approaches to identifying and supporting students with SEB needs given the desire to utilize the measure across contexts (e.g., referral to school-based problem-solving team, screening).

Method

Participants and Procedures

As previously noted, responses to the URP-NEEDS were gathered as part of a larger survey study. School districts were randomly sampled from the Common Core of Data Local Education Agency Universe Survey (National Center for Education Statistics, 2013–2014) in increments of 2,000 districts until invitations to participate in the study were sent to superintendents in a total of 12,132 eligible districts. If the district administrator (e.g., superintendent, director of pupil services) both agreed to participate and completed an initial survey regarding district-level practices, the researchers sent a request to the building administrators (e.g., principal, assistant principal) within two randomly selected schools (i.e., one elementary, one secondary) to complete a survey and distribute a web-based survey link to their teachers. All respondents completed the measures online and those individuals who chose to participate were entered into a drawing to win one of several gift cards. All surveys were completed between March and December of 2016. A total of 1,135 district administrators, 438 building administrators, and 1,396 teachers provided complete responses to all items on the URP-NEEDS and were therefore used for the current study.

Furthermore, to complete the URP-NEEDS, all participants were asked to identify the approach used within their school district to identify and support the SEB needs of students. The first option (heretofore referred to as “external referral”) involved referring those students exhibiting SEB problems to an outside consultant or agency for assistance. The next two options (heretofore referred to as “internal referral”) involved the implementation of intervention supports for those students exhibiting SEB problems. More specifically, respondents could select whether (a) students would be referred to an internal support team to develop and implement an intervention plan or (b) teachers would be encouraged to independently develop and put an intervention into place to see if the problem could be addressed in the classroom before referring the student for further assistance. The next option (heretofore referred to as “screening”) involved two approaches to universal screening. That is, either (a) teachers would complete a brief screening measure for all students in their classrooms and those students whose scores fell outside of the typical range would be referred for further assistance or (b) teachers would first nominate those students exhibiting SEB problems and then complete a screening measure only for the subset of nominated students to

determine who gets referred for assistance. Finally, if respondents believed that none of the available options accurately reflected those practices carried out in their school district, they were encouraged to provide a brief description of the local approach. The data were screened to identify which approaches to student identification and support were endorsed by the respondent samples (see Table 1).

Although the overwhelming majority of respondents endorsed one of the three primary approaches (i.e., external referral, internal referral, screening), roughly 5% of respondents indicated that they did not know, preferred not to answer, and/or used an approach that was different from those outlined. These responses were therefore excluded from further analysis. Among the respondents endorsing one of the three primary approaches (i.e., 1,112 district administrators, 431 building administrators, 1,355 teachers), nearly 75% reported referring students internally, 18% reported using universal screening, and 9% reported using referring students externally.

Measures

Usage Rating Profile for Supporting Students’ Behavioral Needs (URP-NEEDS). The URP-NEEDS is a 24-item self-report measure that was designed to assess the degree to which respondents find an assessment tool to be usable within their local context. Within the URP-NEEDS, usability is represented by five interrelated yet distinct components: willingness to change, feasibility, understanding, family–school collaboration, and external supports. Respondents are asked to use a 6-point Likert-type scale (i.e., 1 = *strongly disagree*, 6 = *strongly agree*) to indicate the degree to which they agree with the provided statements (e.g., *school personnel need consultative support in order to carry out the SEB approach*). Across respondent groups within the current study, internal consistency reliability estimates were found to be within the acceptable to strong range for the understanding (range = .93–.94), willingness (range = .73–.89), feasibility (range = .84–.90), and family–school collaboration (range = .70–.80) subscales; however, Cronbach’s alpha for the external supports subscale (range = .67–.76) was found to be somewhat lower (see Table 2 for within group internal consistencies). Descriptive statistics for the URP-NEEDS subscales by respondent groups can be found in Table 3.

Data Analysis

We sought to examine the measurement invariance of the URP-NEEDS across stakeholder groups (i.e., between group; district administrator, building administrator, teacher) and reported approaches to SEB risk identification (i.e., within group; internal

Table 1
Number of Stakeholders by Self-Identified Approach

Approach	District administrators	Building administrators	Teachers	Total (% of total)
External referral	146	36	93	275 (9.49%)
Internal referral	789	321	992	2102 (72.53%)
Screening	177	74	270	521 (17.98%)
Total	1,112	431	1,355	

Table 2

Internal Consistency Reliability for Subscales by Stakeholder Group and SEB Approach

Stakeholder/SEB approach	N	Understanding	Willingness	Feasibility	Family-school	External supports
District/screening	189	.93	.87	.89	.70	.72
District/internal	829	.94	.84	.86	.78	.73
District/external	153	.93	.80	.84	.78	.70
Building/screening	77	.93	.85	.90	.78	.67
Building/internal	326	.94	.86	.85	.73	.73
Building/external	39	.93	.73	.84	.71	.76
Teacher/screening	270	.94	.89	.87	.77	.74
Teacher/internal	997	.94	.85	.87	.77	.71
Teacher/external	96	.94	.86	.86	.80	.74

Note. N = sample size; SEB approach = reported approach used within their school/district to identify and support the SEB needs of students.

referral, external referral, SEB screening) using multiple-group confirmatory factor analysis (MG-CFA). Analyses were conducted using Mplus Version 8 and the identification of the MG-CFA measurement model was handled with Mplus defaults (Muthén & Muthén, 1998–2010). Given the presence of missing data, the Mplus default of full information maximum likelihood (FIML) estimation was utilized. The FIML approach is currently a highly recommended approach for handling missing data (cf. Enders, 2010), which assumes that the data are missing at random and multivariate normality.

Configural, metric, and scalar invariance tests were conducted using a sequential constraint imposition or *forward* approach (see Table 4), as opposed to a sequential constraint release or *backward* approach (cf., Horn & McArdle, 1992). Testing for configural invariance requires that the same pattern of free and fixed parameters in the CFA measurement model be equivalent across groups. The forward approach requires first testing for configural invariance, then metric invariance, followed by scalar invariance. The sequential “constraint” imposition is due to the fact that as we move from configural to metric, for example, we are imposing constraints on the parameters. For the current analysis, the baseline configural model was the model in which all the model parameters were allowed to freely vary (i.e., except for five factor loadings for identification purposes, one for each of the five subscales). In contrast, testing for metric invariance requires that the factor loadings be set equal across groups and testing for scalar invariance requires that, in addition to the factor loadings being constrained, that the item intercepts be set equal or constrained across

groups. The forward approach requires us to first establish that the configural model fits the data adequately between the two groups. Next, the metric model is fit to the data and the fit of the metric model is compared to that of the configural model. Lastly, the scalar model is fit to the data and the fit of the scalar model is compared with that of the metric model. If at any point in the sequential constraint imposition process, the fit of the model deteriorates we stop the measurement invariance testing.

Given Hu and Bentler’s (1999) recommendation that at least three goodness-of-fit indices meet criteria for the model to demonstrate good fit, several indices were examined to evaluate the fit of the configural model. First, the configural model was evaluated using the CFI, RMSEA, SRMR, and TLI fit indices. Although there do not exist universally agreed upon criteria for evaluating model fit, generally Root Mean Square Error of Approximation (RMSEA) $\leq .05$; Comparative Fit Index (CFI) and Tucker-Lewis Index (TLI) ≥ 0.95 and Standardized Root Mean Square Residual (SRMR) $\leq .06$ are considered indicative of adequate model-data fit (Hu & Bentler, 1999). In order to make comparisons of model fit across two nested models, the chi-square difference test has been proposed. However, given that model chi-square difference tests are sensitive to sample size, it has been argued that change in the comparative fit index (Δ CFI) should be included with results which are aimed at testing for invariance across groups (Cheung & Rensvold, 2002). Therefore, the Δ CFI between the sequential models was evaluated. Δ CFI values of less 0.01 indicate evidence of measurement invariance.

Table 3

Descriptive Statistics by Stakeholder Group and SEB Approach

Stakeholder/SEB approach	Understanding	Willingness to change	Feasibility	Family-school collaboration	External supports
District/screening	4.01 (1.08)	4.15 (1.01)	3.84 (1.22)	5.03 (.99)	4.88 (1.03)
District/internal	3.84 (1.14)	4.17 (1.03)	3.74 (1.22)	5.11 (.89)	4.79 (1.03)
District/external	3.53 (1.14)	3.92 (1.00)	3.50 (1.18)	5.00 (.93)	4.96 (.90)
Building/screening	4.10 (1.11)	4.36 (.96)	3.78 (1.35)	5.15 (.87)	4.86 (.91)
Building/internal	3.99 (1.16)	4.39 (.99)	3.80 (1.24)	5.20 (.89)	4.85 (1.01)
Building/external	3.56 (1.20)	4.19 (.95)	3.64 (1.09)	5.05 (.78)	4.99 (.80)
Teacher/screening	4.16 (1.15)	4.35 (1.09)	3.83 (1.28)	5.15 (.91)	4.78 (1.02)
Teacher/internal	3.97 (1.16)	4.35 (1.09)	3.60 (1.26)	5.16 (.90)	4.72 (1.02)
Teacher/external	3.86 (1.27)	4.24 (1.14)	3.59 (1.26)	5.15 (.77)	4.98 (.93)

Note. Means with standard deviations in parentheses.

Table 4
Measurement Invariance Testing Hierarchy

Model	Parameters to be constrained across groups
Configural (H_{form})	None
Metric (H_{Λ_y})	$\Lambda_y^{(g)} = \Lambda_y^{(g)}$ (equal factor loadings)
Scalar (H_{Λ_y, τ_y})	$\Lambda_y^{(g)} = \Lambda_y^{(2)}, \tau_y^{(g)} = \tau_y^{(g)}$ (equal item intercepts)

Note. Adapted from Bollen (1989).

Results

Measurement Invariance

The results of the measurement invariance tests for the between and within group analyses can be found in Tables 5 and 6. The tables report the overall model fit and the fit indices associated with the respective measurement invariance tests or models.

The results of the measurement invariance analysis suggest that the URP-NEEDS exhibits measurement invariance between the stakeholder participants (see Table 5). With the exception of the chi-square index ($\chi^2 = 2171.54$), each of the indices for the configural model indicated tenable model fit. That is, the RMSEA was $\leq .05$; the CFI and TLI were ≥ 0.95 , and the SRMR was $\leq .06$. Given the tenable model fit for the configural model, the ΔCFI was used to compare model fit between the nested models. The ΔCFI additionally was less than 0.01 for each nested hypothesis indicating that the metric model did not fit the data any worse than the configural model, and that the scalar model did not fit the data any worse than the metric model, indicating the presence of metric and scalar invariance. A similar pattern of evidence was found for the analyses conducted within district administrator and teacher groups (see Table 6), suggesting that the URP-NEEDS exhibits measurement invariance within district administrator and teacher groups, regardless of the stakeholder's self-identified SEB assessment approach. Given the small numbers of building administrators in the external referral group ($N = 36$) and the screening group ($N = 74$), conducting an MG-CFA broken out by approach within the building administrator stakeholder group was inadvisable. However, a single group CFA of the building administrators suggested that the model exhibited reasonable fit overall (CFI = .945, RMSEA = .056, SRMR = .042).

URP-NEEDS Subscale Correlations by Stakeholder Group

Additionally of note are the correlations among the URP-NEEDS subscales by stakeholder group (see Table 7). Overall, the pattern of correlations across subgroups was roughly equivalent

across stakeholder groups. Interestingly, strong correlations were identified between the understanding, willingness to change, and feasibility subscales ($r = .47-.76$), as well as between family-school collaboration and external supports subscales ($r = .64-.67$) regardless of stakeholder group. In contrast, however, weak correlations were identified between these two groups of subscales ($r = -.01-.31$).

Latent Mean Differences Within and Between Stakeholder Groups

Given that measurement invariance was found within the district administrator and teacher stakeholder groups, simple mean differences were considered for group comparisons based on Hedge's G effect sizes. The magnitude of the effect size differences between groups can be seen in Table 8. Although the majority of between group effect size results were found to be fairly small, there were a few interesting differences noted.

First, administrators who indicated that their districts used SEB screening procedures reported higher levels of understanding (ES = 0.55), willingness to change (ES = 0.29), and feasibility (ES = 0.40) than those administrators who indicated that students in the district were externally referred. Higher scores on these three subscales were also found for those administrators reporting the use of internal as opposed to external referral (understanding = 0.36; willingness to change = 0.30; feasibility = 0.27). In contrast, the differences between those administrators reporting use of SEB screening versus internal referral approaches were less pronounced (range ES = -0.02 – 0.18). Second, teachers who reported use of SEB screening procedures in their building had higher mean scores on the understanding (ES = 0.32) and feasibility (ES = 0.27) subscales, as well as lower mean scores on the external supports (ES = -0.26) subscale than teachers in buildings that used external referral. Those teachers reporting use of SEB screening also had higher mean scores on the understanding (ES = 0.20) and feasibility (ES = 0.28) subscales than teachers reporting use of internal referral. Differences between the internal and external referral groups, however, were less pronounced within the teacher sample, with the only notable difference being that teachers in the external referral group indicated a greater need for external supports than teachers in the internal referral group (ES = -0.33).

Discussion

Researchers interested in assessing aspects of usability have often created "home-grown" measures that are specific to a particular population (e.g., teachers) or procedure (e.g., prereferral teams). Although the results obtained from such measures provide useful information regarding potential facilitators and barriers to

Table 5
Measurement Invariance Results: Between Stakeholder Groups

Invariance	χ^2	df	χ^2_{diff}	df_{diff}	CFI	ΔCFI	TLI	RMSEA	SRMR
Configural	2171.54	726	—	—	.964	—	.959	.045	.034
Metric	2227.36	764	55.82	38	.963	-.001	.960	.045	.037
Scalar	2455.77	802	228.42	38	.959	-.004	.957	.046	.041

Note. $\Delta CFI < .01$ indicates invariance; all chi-square and chi-square difference tests are statistically significant at $p < .05$.

Table 6
Measurement Invariance Results: Within Stakeholder Group

Invariance	χ^2	df	χ^2_{diff}	df_{diff}	CFI	ΔCFI	TLI	RMSEA	SRMR
District administrators ($n = 1,135$)									
Configural	1449.03	726	—	—	.950	—	.944	.052	.045
Metric	1505.40	764	55.37	38	.949	-.001	.945	.051	.051
Scalar	1566.77	802	62.37	38	.948	-.001	.946	.051	.053
Teachers ($n = 1,396$)									
Configural	1531.75	726	—	—	.957	—	.951	.050	.040
Metric	1590.71	764	58.96	38	.956	-.001	.952	.049	.046
Scalar	1618.55	802	27.83	38	.956	<.001	.955	.047	.047

Note. $\Delta CFI < .01$ indicates invariance; all chi-square and chi-square difference tests are statistically significant at $p < .05$.

usage, these data are limited to descriptive purposes. In contrast, the URP-NEEDS was developed in order to assess various stakeholders' perceptions of usability surrounding different approaches to identifying and supporting students with SEB needs. Although the measure offers promise with regard to cross-informant comparisons, in order to draw meaningful conclusions from the resultant data, it was first necessary to establish the measurement invariance of the URP-NEEDS.

The most important finding within the current study was that measurement invariance was demonstrated across different groups of school-based stakeholders (i.e., building administrators, district administrators, teachers). A finding of measurement invariance helps to confirm that it is appropriate to administer the URP-NEEDS to various school personnel, and bolsters confidence that differences in subscale scores may represent true group differences as opposed to being measurement artifacts. One advantage of being able to administer the URP-NEEDS across various stakeholder groups is that it is possible to identify the unique facilitators and barriers to usage for particular respondents or respondent groups. Consider, for example, a situation in which a school

district is considering adoption of a universal SEB screening approach in all of the elementary-level buildings. In administering the URP-NEEDS to all building personnel, the central administration might find that although building principals believe that school personnel have the time and materials necessary to carry out universal SEB screening (i.e., high feasibility score), teachers view feasibility to be a substantial concern (i.e., low feasibility score). Such a direct comparison is made possible by the fact that the items within the feasibility subscale have essentially been shown to mean the same thing to both groups of individuals.

In addition, results of the current study suggest that the URP-NEEDS exhibits measurement invariance within district administrators and teachers, regardless of the stakeholder's self-identified SEB approach. This finding is important in that it means that the URP-NEEDS can be used to directly compare the perceived usability of different approaches to identifying and supporting students with SEB needs. Returning to the situation in which a school district is looking to implement a universal SEB screening approach in all of the elementary-level buildings, the district may have more than one option that they are considering (e.g., com-

Table 7
URP-NEEDS Subscale Correlations by Stakeholder Group

Subscale	Understanding	Willingness to change	Feasibility	Family-school collaboration	External supports
District ($n = 1,135$)					
Understanding	1.00				
Willingness to change	.74	1.00			
Feasibility	.64	.47	1.00		
Family-school collaboration	.20	.30	.07	1.00	
External supports	.09	.23	-.01	.66	1.00
Building ($n = 438$)					
Understanding	1.00				
Willingness to change	.74	1.00			
Feasibility	.76	.59	1.00		
Family-school collaboration	.26	.31	.15	1.00	
External supports	.22	.30	.16	.64	1.00
Teacher ($n = 1,396$)					
Understanding	1.00				
Willingness to change	.73	1.00			
Feasibility	.73	.56	1.00		
Family-school collaboration	.23	.23	.09	1.00	
External supports	.17	.19	.07	.67	1.00

Table 8
Within Group Effect Sizes on URP-NEEDS Subscales

Subscale	ES Screening—External Referral	ES Screening—Internal Referral	ES Internal Referral—External Referral
District administrators (<i>n</i> = 1,135)			
Understanding	.55	.18	.36
Willingness to change	.29	-.02	.30
Feasibility	.40	.10	.27
Family-school collaboration	.02	-.07	.09
External supports	-.12	.04	-.16
Teachers (<i>n</i> = 1,396)			
Understanding	.32	.20	.13
Willingness to change	.11	.01	.11
Feasibility	.27	.28	.01
Family-school collaboration	-.00	-.02	.02
External supports	-.26	.06	-.33

pletion of a full-length rating scale for all students nominated by their teachers, completion of a brief screening measure for all students in the building). As one piece of the district's decision-making process, teachers might be asked to complete the URP-NEEDS based on each screening option in order to gather their perceptions in an efficient and standardized manner. Within the current sample, for example, both district administrators and teachers who reported use of SEB screening indicated that they understood the procedures better (i.e., understanding) and found them to require fewer resources to implement (i.e., feasibility) than those individuals who reported that students with SEB needs were referred externally. In addition, administrators in districts using either SEB screening or internal referral procedures felt that school staff were generally more willing to change than those administrators in buildings employing external referral.

Although school psychologists may be convinced of the benefits of proactive models of identification and behavioral support, other key stakeholders in the school system may not always be. Given their training in consultation and use of a problem-solving approach, school psychologists are therefore ideally positioned to lead school- or district-level efforts to identify—and address—the potential barriers to usage at the levels of the individual, innovation, and environment (Sanetti & Kratochwill, 2009). The URP-NEEDS subscales assess both a school's internal capacity for implementation (i.e., understanding, willingness to change, feasibility) as well as the degree to which stakeholders perceive external supports to be necessary for use (i.e., family-school collaboration, external supports).

Results suggest promise for use of the URP-NEEDS in enhancing understanding of those factors that may influence use of approaches to SEB risk identification; however, limitations of both the current study and the measure should be noted. Because respondents were asked to complete the URP-NEEDS with regard to the approach utilized within their school to identify students with SEB needs, the distribution of responses was fairly skewed. Specifically, of the available options, almost three quarters of respondents reported that students exhibiting SEB problems were referred internally whereas less than 10% of respondents reported referring these students externally. Additionally, as noted previously, we could not adequately examine measurement invariance by approach within the building administrator sample. Additional

research is therefore warranted employing a larger sample of building administrators to examine the invariance across approaches within this respondent group.

Concerning the measure more broadly, although results of the current study contribute to the developing psychometric evidence base in support of the URP-NEEDS, additional research is needed such that consumers can use the measure with greater confidence. For one, the test-retest reliability of the measure has not yet been examined. Understanding whether reported perceptions of usability are consistent over a short latency period is important toward informing applied use. In addition, the internal consistency of the external supports (i.e., .67-.76) and family-school collaboration subscales (i.e., .71-.80) were found to be lower than desirable in the current sample. This suggests a potential need to examine these subscales further in order to determine whether this concern is specific to the current sample or if additional efforts are needed to strengthen the subscales (e.g., generation of additional items). Use of the URP-NEEDS is therefore recommended cautiously pending examination of these additional areas of psychometric inquiry.

Finally, further research is also needed in order to understand the degree to which each of these hypothesized factors actually predicts successful usage. Given the unique needs, priorities, and stresses of individual settings, however, it is not necessarily expected that these five factors would wield similar influence across different contexts. Therefore, the URP-NEEDS may be used to identify the primary factor—or constellation of factors—that may serve as potential obstacles to successful usage locally. These data could then be used to either inform changes to procedures (e.g., to enhance feasibility) or systems (e.g., to strengthen external supports or collaboration), or to more efficiently target efforts to enhance user understanding and motivation (e.g., willingness to change) in order to maximize future potential for usage.

References

- Aarons, G. A. (2004). Mental health provider attitudes toward adoption of evidence-based practice: The Evidence-Based Practice Attitude Scale (EBPAS). *Mental Health Services Research*, 6, 61-74. <http://dx.doi.org/10.1023/B:MHSR.0000024351.12294.65>
- Briesch, A. M., Chafouleas, S. M., Dineen, J. N., & McCoach, D. B. (2019). *School building administrator reports of screening practices*

- across academic, physical health, and behavioral domains. Manuscript in preparation.
- Bollen, K. A. (1989). *Structural equations with latent variables*. New York, NY: Wiley. <http://dx.doi.org/10.1002/9781118619179>
- Bradley, R., Danielson, L., & Doolittle, J. (2007). Responsiveness to intervention: 1997 to 2007. *Teaching Exceptional Children*, 35, 8–12. <http://dx.doi.org/10.1177/004005990703900502>
- Bradley, R., Doolittle, J., & Bartolotta, R. (2008). Building on the data and adding to the discussion: The experiences and outcomes of students with emotional disturbance. *Journal of Behavioral Education*, 17, 4–23. <http://dx.doi.org/10.1007/s10864-007-9058-6>
- Briesch, A. M., Chafouleas, S. M., Neugebauer, S. R., & Riley-Tillman, T. C. (2013). Assessing influences on intervention implementation: Revision of the usage rating profile-intervention. *Journal of School Psychology*, 51, 81–96. <http://dx.doi.org/10.1016/j.jsp.2012.08.006>
- Bruhn, A. L., Woods-Groves, S., & Huddle, S. (2014). A preliminary investigation of emotional and behavioral screening practices in K-12 schools. *Education & Treatment of Children*, 37, 611–634. <http://dx.doi.org/10.1353/etc.2014.0039>
- Burns, M. K., & Symington, T. (2002). A meta-analysis of prereferral intervention teams: Student and systemic outcomes. *Journal of School Psychology*, 40, 437–447. [http://dx.doi.org/10.1016/S0022-4405\(02\)00106-1](http://dx.doi.org/10.1016/S0022-4405(02)00106-1)
- Byrne, B. M. (2004). *Structural equation modeling with AMOS: Basic concepts, applications, and programming*. Mahwah, NJ: Erlbaum.
- Chafouleas, S. M., Briesch, A. M., McCoach, D. B., & Riley-Tillman, T. C. (2018). *Usage rating profile for supporting students' behavioral needs*. Storrs, CT: University of Connecticut.
- Chafouleas, S. M., Briesch, A. M., Neugebauer, S. R., & Riley-Tillman, T. C. (2011). *Usage rating profile—intervention (revised)*. Storrs, CT: University of Connecticut.
- Chafouleas, S. M., Miller, F. G., Briesch, A. M., Neugebauer, S. R., & Riley-Tillman, T. C. (2012). *Usage rating profile—assessment*. Storrs, CT: University of Connecticut.
- Chalfant, J., Pysh, M., & Moultrie, R. (1979). Teacher assistance teams: A model for within-building problem solving. *Learning Disability Quarterly*, 2, 85–89. <http://dx.doi.org/10.2307/1511031>
- Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, 9, 233–255. http://dx.doi.org/10.1207/S15328007SEM0902_5
- Crone, L. A., Carlson, S. E., Haack, M. K., Kennedy, P. C., Baker, S. K., & Fien, H. (2016). Data-based decision-making teams in middle school: Observations and implications from the middle school intervention project. *Assessment for Effective Intervention*, 41, 79–93. <http://dx.doi.org/10.1177/1534508415610322>
- Dimitrov, D. M. (2010). Testing for measurement invariance in the context of construct validation. *Measurement & Evaluation in Counseling & Development*, 43, 121–149. <http://dx.doi.org/10.1177/0748175610373459>
- Durlak, J. A., & DuPre, E. P. (2008). Implementation matters: A review of research on the influence of implementation on program outcomes and the factors affecting implementation. *American Journal of Community Psychology*, 41, 327–350. <http://dx.doi.org/10.1007/s10464-008-9165-0>
- Enders, C. K. (2010). *Applied missing data analysis*. New York, NY: Guilford Press.
- Flugum, K. R., & Reschly, D. J. (1994). Prereferral interventions: Quality indices and outcomes. *Journal of School Psychology*, 32, 1–14. [http://dx.doi.org/10.1016/0022-4405\(94\)90025-6](http://dx.doi.org/10.1016/0022-4405(94)90025-6)
- Horn, J. L., & McArdle, J. J. (1992). A practical and theoretical guide to measurement invariance in aging research. *Experimental Aging Research*, 18, 117–144. <http://dx.doi.org/10.1080/03610739208253916>
- Hu, L., & Bentler, M. (1999). Cut-off criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6, 1–55. <http://dx.doi.org/10.1080/10705519909540118>
- Jöreskog, K. G., & Sörbom, D. (1979). *Advances in factor analysis and structural equation models*. Cambridge, MA: ABT Books.
- Kamphaus, R. W., & Reynolds, C. R. (2015). *Behavior assessment system for children—3rd ed. (BASC-3): Behavioral and emotional screening system (BESS)*. Bloomington, MN: Pearson.
- Kurtz, S. M. S. (1990). Adherence to diabetes regimens: Empirical status and clinical applications. *The Diabetes Educator*, 16, 50–59. <http://dx.doi.org/10.1177/014572179001600112>
- Lane, K. L., Pierson, M. R., Robertson, E. J., & Little, A. (2004). Teachers' views of prereferral interventions: Perceptions of and recommendations for implementation support. *Education & Treatment of Children*, 27, 420–439.
- Lee-Tarver, A. (2006). A survey of teachers' perceptions of the function and purpose of student support teams. *Education*, 126, 525–533.
- Lehman, W. E. K., Greener, J. M., & Simpson, D. D. (2002). Assessing organizational readiness for change. *Journal of Substance Abuse Treatment*, 22, 197–209. [http://dx.doi.org/10.1016/S0740-5472\(02\)00233-7](http://dx.doi.org/10.1016/S0740-5472(02)00233-7)
- Meredith, W. (1993). Measurement invariance, factor analysis and measurement invariance. *Psychometrika*, 58, 525–543. <http://dx.doi.org/10.1007/BF02294825>
- Miller, F. G., Neugebauer, S. R., Chafouleas, S. M., Briesch, A. M., & Riley-Tillman, T. C. (2013, August). *Examining innovation usage: Construct validation of the usage rating profile—assessment*. Poster presentation at the American Psychological Association Annual Convention, Honolulu, HI.
- Muthén, L. K., & Muthén, B. O. (1998–2010). *Mplus user's guide*. 6th edition. Los Angeles, CA: Author.
- National Center for Education Statistics. (2013–2014). Local Education Agency (School District) Universe Survey Data. Retrieved from <https://nces.ed.gov/ccd/pubagency.asp>
- Perepletchikova, F., & Kazdin, A. E. (2005). Treatment integrity and therapeutic change: Issues and research recommendations. *Clinical Psychology: Science and Practice*, 12, 365–383. <http://dx.doi.org/10.1093/clipsy.bpi045>
- Proctor, E., Silmere, H., Raghavan, R., Hovmand, P., Aarons, G., Bunger, A., . . . Hensley, M. (2011). Outcomes for implementation research: Conceptual distinctions, measurement challenges, and research agenda. *Administration and Policy in Mental Health*, 38, 65–76. <http://dx.doi.org/10.1007/s10488-010-0319-7>
- Rones, M., & Hoagwood, K. (2000). School-based mental health services: A research review. *Clinical Child and Family Psychology Review*, 3, 223–241. <http://dx.doi.org/10.1023/A:1026425104386>
- Sanetti, L. M. H., & Kratochwill, T. R. (2009). Toward developing a science of treatment integrity: Introduction to the special series. *School Psychology Review*, 38, 445–459.
- Slonski-Fowler, K. E., & Truscott, S. D. (2004). General education teachers' perceptions of the prereferral intervention team process. *Journal of Educational & Psychological Consultation*, 15, 1–39. http://dx.doi.org/10.1207/s1532768xjepc1501_1
- Walker, H. M., Nishioka, V. M., Zeller, R., Severson, H. H., & Feil, E. G. (2000). Causal factors and potential solutions for the persistent under-identification of students having emotional or behavioral disorders in the context of schooling. *Assessment for Effective Intervention*, 26, 29–40. <http://dx.doi.org/10.1177/073724770002600105>
- WHO. (2016). *Depression*. Retrieved from <http://www.who.int/media/centre/factsheets/fs369/en/>

Appendix

URP-NEEDS Factors and Associated Items

Understanding

School personnel are knowledgeable about the purpose and goals of social, emotional, and behavioral screening.
 School personnel know how to use social, emotional, and behavioral screening data to document student improvements.
 School personnel understand how goals for social, emotional, and behavioral screening fit with a system of student supports.
 School personnel understand how to use social, emotional, and behavioral screening data to guide decisions about student supports.
 School personnel understand the procedures of the social, emotional, and behavioral approach.
 School personnel know how to carry out the social, emotional, and behavioral approach.
 The current social, emotional, and behavioral approach is effective for addressing a variety of problems.
 The current social, emotional, and behavioral approach offers a good way to identify a child's behavior problem.
 School personnel are familiar with what can be done to prevent or treat social, emotional, and behavioral difficulties in school.
 School personnel are confident in their ability to carry out the social, emotional, and behavioral approach.

Willingness to change

School personnel like to use new strategies to help address the social, emotional, and behavioral needs of students.
 School personnel would try a new strategy to address the social, emotional, and behavioral needs of students even if it were very different than what they are used to doing.
 School personnel are willing to use new and different types of social, emotional, and behavioral strategies developed by researchers.
 School personnel are willing to change how they operate to meet the social, emotional, and behavioral needs of students.

Feasibility

The preparation of materials needed for the social, emotional, and behavioral approach is reasonable for school personnel.
 The total time required for staff to carry out the social, emotional, and behavioral approach is manageable for school personnel.
 The materials needed for the social, emotional, and behavioral approach are reasonable for school personnel.
 The amount of time required of school personnel for record keeping related to the social, emotional, and behavioral approach is reasonable.

Family-school collaboration

Regular home-school communication is needed in order to execute the social, emotional, and behavioral approach.
 A positive home-school relationship is needed to carry out the social, emotional, and behavioral approach.
 Parental collaboration is needed in order to implement this social, emotional, and behavioral approach.

External supports

School personnel need consultative support in order to carry out the social, emotional, and behavioral approach.
 A positive relationship with community agencies is important to carry out the social, emotional, and behavioral approach.
 Ongoing assistance from external consultants is necessary to successfully use the social, emotional, and behavioral approach.

Received June 15, 2018

Revision received October 29, 2018

Accepted November 27, 2018 ■

Correction to Briesch et al. (2019)

In the article “Factorial Invariance of the Usage Rating Profile for Supporting Students’ Behavioral Needs (URP-NEEDS),” by Amy M. Briesch, Sandra M. Chafouleas, Dakota W. Cintron, and D. Betsy McCoach (*School Psychology*, Vol. 35, No. 1, 2020, pp. 51–60. <https://doi.org/10.1037/spq0000309>), in the fourth paragraph of the “Understanding the Factors That Influence Usage” section and in the “Usage Rating Profile for Supporting Students’ Behavioral Needs (URP-NEEDS)” section, the URP-NEEDS was incorrectly reported to have 23 items. This measure consists of 24 items. This item was also missing in the Appendix under the “Understanding” factor: “School personnel understand how goals for social, emotional, and behavioral screening fit with a system of student supports.” All versions of this article have been corrected.

<http://dx.doi.org/10.1037/spq0000350>