Journal of Personality Assessment

Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/hjpa20

Measuring Resilience With the RS-14: A Tale of Two Samples

Bethany J. Aiena\textsuperscript{ab}, Brandy J. Baczwaski\textsuperscript{ab}, Stefan E. Schulenberg\textsuperscript{ab} & Erin M. Buchanan\textsuperscript{c}
\textsuperscript{a} Department of Psychology, The University of Mississippi
\textsuperscript{b} Clinical-Disaster Research Center (UM-CDRC), The University of Mississippi
\textsuperscript{c} Department of Psychology, Missouri State University
Published online: 25 Sep 2014.

To cite this article: Bethany J. Aiena, Brandy J. Baczwaski, Stefan E. Schulenberg & Erin M. Buchanan (2015)
Measuring Resilience With the RS-14: A Tale of Two Samples, Journal of Personality Assessment, 97:3, 291-300, DOI:
10.1080/00223891.2014.951445

Please scroll down for article

Taylor & Francis makes every effort to ensure the accuracy of all the information (the “Content”) contained in the publications on our platform. However, Taylor & Francis, our agents, and our licensors make no representations or warranties whatsoever as to the accuracy, completeness, or suitability for any purpose of the Content. Any opinions and views expressed in this publication are the opinions and views of the authors, and are not the views of or endorsed by Taylor & Francis. The accuracy of the Content should not be relied upon and should be independently verified with primary sources of information. Taylor and Francis shall not be liable for any losses, actions, claims, proceedings, demands, costs, expenses, damages, and other liabilities whatsoever or howsoever caused arising directly or indirectly in connection with, in relation to or arising out of the use of the Content.

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden. Terms & Conditions of access and use can be found at http://www.tandfonline.com/page/terms-and-conditions
Measuring Resilience With the RS–14: A Tale of Two Samples

BETHANY J. AIENA,1,2 BRANDY J. BACZWASKI,1,2 STEFAN E. SCHULENBERG,1,2 AND ERIN M. BUCHANAN3

1Department of Psychology, The University of Mississippi
2Clinical-Disaster Research Center (UM-CDRC), The University of Mississippi
3Department of Psychology, Missouri State University

The purpose of this study was to systematically examine the psychometric properties of the 14-item Resilience Scale (RS–14; Wagnild, 2009a). Exploratory and confirmatory factor-analytic methods were employed, including an analysis of measurement invariance models by sex and race/ethnicity. Descriptive statistics, reliability, and validity data were also calculated. Analyses were based on 2 samples, one made up of individuals seeking mental health services following the 2010 Gulf oil spill (N = 1,032) and the other made up of university students (N = 1,765). For both samples, all items loaded (> .30) onto 1 factor, indicating cohesive structure for a 1-factor model explaining 53.2% of the variance in the clinical sample and 67.6% of the variance in the undergraduate sample. Further, the examination of measurement invariance indicated that the RS–14 was similarly structured for sex and race/ethnicity. Reliability coefficients exceeded .90 for both samples and also when data were examined by comparison groups. The RS–14 correlated significantly and as expected with measures of positive concepts (such as perceived meaning in life and satisfaction with life) and indexes of psychological distress (such as depression, anxiety, stress, and posttraumatic stress). These data support the utility of the RS–14 with clinical and undergraduate student samples. Implications for these data are discussed.

Most people will experience at least one traumatic event in their lifetime (Breslau, Peterson, Poisson, Schultz, & Lucia, 2004; Copeland, Keeler, Angold, & Costello, 2007; Kessler, Sonnega, Bromet, Hughes, & Nelson, 1995). How people respond to a given traumatic event, such as a disaster, will differ from person to person based on a variety of risk and protective factors (Drescher et al., 2012; Drescher, Schulenberg, & Smith, 2014; Park & Slattery, 2014; Zakour, 2012). Some people will develop a clinical syndrome, such as posttraumatic stress disorder (PTSD; Hamblen, Barnett, & Norris, 2012; Williams, McDevitt-Murphy, Fields, Weathers, & Flood, 2011), others will have only short-term disturbances in functioning, and some individuals will go on to experience posttraumatic growth (Santucci, 2012; Tedeschi & Calhoun, 1995; Triplet, Tedeschi, Cann, Calhoun, & Reeve, 2012).

A given person’s resilience is an essential feature with respect to how that person will respond during and after the occurrence of a disaster (Bonanno, 2004; Halpern & Tramontin, 2007; Zakour, 2012). Resilience refers to the ability to withstand or adaptively recover from stressors. Resilience also promotes psychological and physical well-being (Bonanno, 2004; Doherty & Clayton, 2011; Park & Slattery, 2014; Zakour, 2012). It could be further conceptualized as the adaptive use of available internal and environmental resources when a person is confronted with adverse events (Wong & Wong, 2012). The concept has been of interest to the field of disaster mental health for years, and is a key aspect of positive psychology (Bonanno, 2004; Bonanno, Galea, Bucciarel, & Vlahov, 2007; Herrman, 2012; Scali et al., 2012). As such, there is substantial and growing interest in strengths-based research and intervention efforts that encompass resilience. Consequently, research on resilience continues to proliferate, and the need for brief, psychometrically sound measures of resilience continues to grow. For instance, resilience is negatively correlated with symptoms of generalized anxiety and posttraumatic stress and positively correlated with gratitude, optimism, and positive affect (Baldwin, Jackson, Okoh, & Cannon, 2011; Fredrickson, Tugade, Waugh, & Larkin, 2003; Scali et al., 2012; Tugade & Fredrickson, 2004). Because resilience is an essential strength that promotes well-being and serves as a protective factor against a range of stressful or negative events such as disasters and other potentially traumatic events, it is vital to have reliable and valid assessment methods for a range of populations. The 14-item Resilience Scale (RS–14) is one example of an instrument designed to assess the concept. Prior to discussing the RS–14 in greater detail, we briefly summarize the literature on resilience in relation to sex and race/ethnicity.

Although many studies have examined demographic differences in PTSD following disasters, significantly fewer investigations have looked at potential sex and racial/ethnic differences in resilience. Whereas some researchers have found that men and women demonstrate comparable levels of resilience after a traumatic event (Morano, 2010), others have found that females were likely to be less resilient than men (Bonanno et al., 2007). In a large community survey, Campbell-Sills, Forde, and Stein (2009) found that women had lower self-reported resilience than men. The authors noted that these sex differences could be at least partly due to response bias, in that men might be more concerned with appearing strong and unfazed when confronted with stress.

Limited studies have examined differences in resilience in underrepresented groups. With regard to PTSD, meta-analytic data have shown that minority group membership places one at higher risk (Brewin, Andrews, & Valentine, 2000), whereas other researchers have reported that racial/ethnic differences following a disaster become nonsignificant statistically when factors such as socioeconomic status (SES) and education...
level are controlled for (Adams & Boscarno, 2005; Galea et al., 2003; Kessler et al., 1995). In terms of resilience specifically, when controlling for such factors as SES and education level in multivariate analyses, racial/ethnic differences no longer appear to be significant predictors (Bonanno et al., 2007). Although some studies are available examining sex and race/ethnicity as influential factors in resilience, more studies are warranted.

It is well-known that assessment measures might function differently considering various demographic variables (Brown, 2006). That is, psychometric properties of a measure might differ on the basis of sex or race/ethnicity. Because of the potential for such differences, it is necessary to conduct studies to better understand whether and how these variables play a role for a given measure, in this case the RS–14.

The original Resilience Scale (RS) was made up of 25 items (Wagnild & Young, 1993), but was reduced to 14 items (RS–14) to decrease completion time (Wagnild, 2009a). The RS–14 was developed by retaining items from the original RS with the highest interitem correlations and those that also measured five characteristics of resilience, termed the Resilience Core, specifically meaning and purposeful life (Items 2, 9, and 13), perseverance (Items 6 and 8), equanimity (Items 3 and 10), self-reliance (Items 1, 5, 7, 12, and 14), and existential loneliness (Items 4 and 11; Wagnild, 2009a). With regard to how these terms are defined, an individual’s life is meaningful or purposeful when he or she has clear goals and values. Perseverance refers to motivation to persist in the face of difficulty. Equanimity occurs when the individual is balanced such that he or she is able to deal with life’s stressors and maintain composure in difficult situations by remaining optimistic, taking difficulties in stride, and looking for opportunities to bounce back. Self-reliant individuals learn problem-solving skills from challenging experiences. Existential loneliness occurs when individuals are comfortable with themselves and with handling things on their own. The Resilience Core is the conceptual foundation for the RS–14, wherein being resilient is a combination of protective factors working together to produce positive outcomes to stressors (Wagnild, 2009a, 2009b; Wagnild & Young, 1990). The RS–14 employs a 7-point Likert-type response format. Item scores are summed to yield a total score ranging from 14 to 98, with higher scores suggestive of greater perceived resilience. Scoring guidelines are provided by Wagnild (2009a), with total scores categorized as very low (14–56), low (57–64), on the low end (65–73), moderate (74–81), moderately high (82–90), and high (91–98). Descriptors of general item content are provided in Table 1 for context. The interested reader is referred to Wagnild (2009a) and www.resiliencecenter.com for specific RS–14 item content, as well as measure instructions and format.

Articles have been published on the Japanese and Brazilian versions of the RS–14, as well as Chinese and Taiwanese translations (see Damáśio, Borsa, & da Silva, 2011; Nishi, Uehara, Kondo, & Matsuoka, 2010; Yang, Li, & Xia, 2012), but there is a need for additional, independent studies of the psychometric properties of the English version of the RS–14. Available data reported by Wagnild (2009a) indicate that the RS–14 yields reliable scores (coefficient alphas of .90 and greater), and as would be expected RS–14 scores are highly correlated with the original RS ($r = .97$, $p < .001$). The RS–14 has been found to be negatively correlated with measures of depression and disability and positively correlated with measures of self-esteem, social support, general psychological well-being, purpose in life, and self-reported good health (Damáśio et al., 2011; Nishi et al., 2010; Wagnild, 2009a). With regard to factor structure, principal components analyses supported a single-factor solution, with all item factor loadings greater than .40 (Wagnild, 2009a). Yang et al. (2012) examined the measurement invariance of the RS–14 in samples of U.S., Chinese, and Taiwanese college students and supported a one-factor model that demonstrated scalar invariance across cultures.

There are few independent studies of the psychometric properties of the RS–14 available in the literature. Thus, this measure would benefit from additional studies as to its reliability and association with other measures. Moreover, to our knowledge the factor-analytic support for the RS–14 has not been replicated with a disaster-affected, mental-health-service-seeking sample and compared to a college sample. Given that the measure is based on five characteristics, the Resilience Core, it is possible that the RS–14 might assess a combination of distinct, but related concepts that are conceptualized as constituting resilience. An investigation into the factor structure (or structural validity) would have implications for how the measure is scored and interpreted. Moreover, structural validity is an important, but often neglected, aspect of a measure’s psychometric properties (Steger, 2006). Replication of the initial analyses would provide essential information about the measure’s psychometric viability.

In terms of the RS–14 and demographic variables, Wagnild (2009a) reported statistically significant differences between females and males, with females scoring higher. Although these scores might be statistically significant, they are unlikely to be clinically or practically significant differences as these scores were only 2.6 points apart ($d = 0.18$). In a study conducted by Winsett, Stender, Gower, and Burghen (2010), a descriptive comparative pilot study of adolescents attending a diabetes camp, no statistically significant differences in RS–14 scores were noted by sex, although, a medium effect size was found ($d = 0.60$). Additionally, Black adolescents reported significantly higher RS–14 scores than White adolescents ($d = 0.53$; Wagnild, 2009a; Winsett et al., 2010). Overall, with respect to sex and race/ethnicity, these findings, as well as the other resilience findings reported earlier, are inconclusive with respect to drawing firm conclusions and warrant further empirical inquiry.

**THIS STUDY**

This study systematically examined the psychometric properties of the RS–14 employing data from two large samples. The first sample was made up of adults seeking mental health services in response to the Gulf oil spill ($N = 1,032$). The second sample was made up of undergraduates from a medium-sized university located in the southern United States ($N = 1,765$). Exploratory and confirmatory factor-analytic procedures were employed to assess the factor structure of the measure. In addition, we examined measurement invariance models by sex and race/ethnicity and calculated descriptive data and reliability coefficients (for the total samples and by sex and race/ethnicity). We hypothesized that for both samples, a one-factor model would be supported and that scores
TABLE 1.—14-item Resilience Scale (RS–14) factor loadings for clinical (N = 516) and college (N = 883) samples.

<table>
<thead>
<tr>
<th>Item Number</th>
<th>Item Content</th>
<th>Clinical Sample</th>
<th>College Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Ability to cope</td>
<td>0.58</td>
<td>0.72</td>
</tr>
<tr>
<td>2</td>
<td>Pride</td>
<td>0.67</td>
<td>0.85</td>
</tr>
<tr>
<td>3</td>
<td>Acceptance</td>
<td>0.65</td>
<td>0.80</td>
</tr>
<tr>
<td>4</td>
<td>Self-regard</td>
<td>0.66</td>
<td>0.78</td>
</tr>
<tr>
<td>5</td>
<td>Organized</td>
<td>0.64</td>
<td>0.77</td>
</tr>
<tr>
<td>6</td>
<td>Drive</td>
<td>0.75</td>
<td>0.87</td>
</tr>
<tr>
<td>7</td>
<td>Perseverance</td>
<td>0.77</td>
<td>0.79</td>
</tr>
<tr>
<td>8</td>
<td>Willpower</td>
<td>0.70</td>
<td>0.81</td>
</tr>
<tr>
<td>9</td>
<td>Interest/engagement</td>
<td>0.74</td>
<td>0.84</td>
</tr>
<tr>
<td>10</td>
<td>Humor</td>
<td>0.70</td>
<td>0.82</td>
</tr>
<tr>
<td>11</td>
<td>Self-efficacy</td>
<td>0.78</td>
<td>0.80</td>
</tr>
<tr>
<td>12</td>
<td>Dependable</td>
<td>0.67</td>
<td>0.80</td>
</tr>
<tr>
<td>13</td>
<td>Meaning</td>
<td>0.76</td>
<td>0.81</td>
</tr>
<tr>
<td>14</td>
<td>Resourcefulness</td>
<td>0.80</td>
<td>0.83</td>
</tr>
</tbody>
</table>

Note. Items for the RS–14 are described in the measure’s manual (Wagnild, 2009a) and can be viewed at www.resiliencecenter.com.

would be highly reliable as assessed by conventional interpretive standards (i.e., Cronbach’s alpha). With varying evidence reported in the literature, we surmised that there might be significant differences in scale functioning between men and women and also between Black and White individuals, which further motivated our measurement invariance analyses. We also investigated the measure’s validity by correlating it with measures of positive psychological variables and indexes of psychological distress. We hypothesized that the RS–14 would be positively associated with measures of positive concepts (such as perceived meaning in life and satisfaction with life) and negatively associated with measures of psychological distress (such as depression, anxiety, stress, and posttraumatic stress).

METHOD

Participants and Procedures

Sample 1. The first sample included adult clients from 10 mental health facilities on the Mississippi Gulf Coast. Participants included new clients as well as clients who had already been receiving services at the time data were collected. Facilities where these data were collected included two mental health centers, one Vietnamese community organization, one school-based counseling service, four private counseling centers, one women’s shelter, and one in-patient mental health hospital. Data were collected as part of a larger grant-funded project following the Gulf oil spill (also known as the Deepwater Horizon oil spill) that began on April 20, 2010 (Drescher et al., 2012; Drescher et al., 2014).

Following the Gulf oil spill, British Petroleum (BP) provided the Mississippi Department of Mental Health (MS DMH) with funds to address mental health needs as a result of the spill. MS DMH established a grant program, with funds primarily used for direct service provision, training, and outreach efforts. Through a contract between the MS DMH and the corresponding author, a university-based research team was assembled to evaluate the impact of the spill on Mississippi Gulf Coast residents seeking mental health services and to assess the amount of services provided by the funded sites. Clients receiving services from these organizations were asked to complete an array of measures related to psychological distress and well-being, the RS–14 among them. Eighty-seven participants were excluded for incomplete surveys, which was the only criterion for exclusion. The sample (N = 1,032) included 53.8% (n = 555) female participants, was predominantly White (69.2%; n = 714) or Black (20.8%, n = 215), with ages ranging from 18 to 79 years (M = 38.76, SD = 12.73).

Sample 2. Data were also gathered at a medium-sized university located in the southern United States. Data were part of a larger screening procedure designed to determine eligibility to participate in a wide variety of psychological studies. The screening procedure contained an array of measures, including the RS–14. Students received course credit or extra credit for their participation. Participants with incomplete surveys were excluded (n = 78). This sample (N = 1,765) was largely female (62.4%; n = 1,101), and White (78.0%; n = 1,377) or Black (16.3%, n = 287). Ages ranged from 18 to 48 years (M = 19.04; SD = 2.07).

Measures

In addition to the RS–14, the following measures of positive psychological variables and indexes of psychological distress were also administered to both samples. Measures were chosen for their conciseness and psychometric support.

Satisfaction with Life Scale. The Satisfaction with Life Scale (SWLS; Diener, Emmons, Larsen, & Griffin, 1985) is a five-item scale designed to measure a respondent’s global satisfaction with life. A 7-point, Likert-type scale is used with item anchors ranging from 1 (strongly disagree) to 7 (strongly agree). Scores on this measure range from 5 to 35, and higher scores indicate greater perceived life satisfaction. The internal consistency reliability of this scale was as follows for each sample: clinical sample = .871; college student sample = .92.

Purpose in Life Test–Short Form. The Purpose in Life Test (PIL; Crumbaugh & Maholick, 1964, 1969) is a version of the original 20-item Purpose in Life Test (PIL; Crumbaugh & Maholick, 1964, 1969). Each of the items is rated on a 7-point, Likert-type scale with differing anchors based on the individual item content. Scores on this measure range from 4 to 28, with higher scores indicating higher perceived meaning in life. The internal consistency reliability of this scale was as follows for each sample: clinical sample = .882; college student sample = .90.

21-Item Depression Anxiety Stress Scales. The 21-Item Depression Anxiety Stress Scales (DASS–21; Lovibond & Lovibond, 1995) is comprised of three 7-item scales that assess depression, anxiety, and stress separately. Items are

1. A smaller subset of these data was used in a separate investigation of self-efficacy, meaning, and life satisfaction following the Gulf oil spill (see Drescher et al., 2012).
2. See note 1.
scored on a 4-point, Likert-type scale with higher scores indicating higher levels of severity of each domain. Scores on the total scale range from 0 to 42. The internal consistency estimates were .94 (depression), .89 (anxiety), and .92 (stress) for the clinical sample and .88 (depression), .79 (anxiety), and .83 (stress) for the college student sample.

**Postraumatic Stress Disorder Checklist.** The Postraumatic Stress Disorder Checklist (PCL; Weathers, Huska, & Keane, 1991) is a 17-item measure designed to assess the symptoms of PTSD as described in the *Diagnostic and Statistical Manual of Mental Disorders* (4th ed. [DSM–IV]; American Psychiatric Association, 2000). There are three different, interchangeable versions that can be used in military settings (PCL–M), civilian settings (PCL–C), and settings in which there is a specific, identifiable stressful experience (PCL–S; Weathers et al., 1991; Weathers, Litz, Herman, Huska, & Keane, 1993; Weathers, Litz, Huska, & Keane, 1994). The PCL–S uses a 5-point Likert-type response format with options ranging from 1 (not at all) to 5 (extremely) and scores on the PCL–S range from 17 to 85, with higher scores indicating more endorsed symptoms of PTSD symptomology. The PCL–S was used with the disaster-affected sample, and the PCL–C was used with the undergraduate sample. The internal consistency reliability of this scale was as follows for each sample: clinical sample = .97; college student sample = .95.

**Data-Analytic Strategy**

The purpose of this study was to analyze the psychometric properties of the RS–14, as well as provide reliability and validity support for using the RS–14 to measure overall resilience. To assess these psychometric properties, we first analyzed each sample using exploratory factor analysis (EFA) to determine appropriate factor structure, followed by confirmatory factor analysis (CFA) of those results. Given the mixed literature on sex and race/ethnicity differences in resilience, we used multigroup confirmatory procedures to explore any differences present in both the clinical and college student samples. These procedures are now described in detail.

**Exploratory Factor Analysis.** First, EFA was used to determine if all items loaded onto one resilience factor, as this scale was originally designed to assess five areas of resilience. EFA was used in this case, as opposed to a principal components analysis, for the following reasons: (a) because resilience was expected to be an underlying latent variable, and (b) we wished to explain unique variance and item correlation, not simply test data reduction (see Preacher & MacCallum, 2003). To accomplish the primary goal of the study, scree plots and parallel analyses were examined to determine the possible number of factors, using the FACTOR program (Lorenzo-Seva & Ferrando, 2006). FACTOR is a free data analysis program that calculates both parallel analyses (to investigate the number of factors) and commonly reported fit indexes. The program can be downloaded from Lorenzo-Seva's website. It has been used widely in empirical studies, having been cited in more than 250 articles.

Potential models were tested using maximum likelihood estimation and an oblique rotation (direct oblimin) if parallel analyses supported a model with more than one factor. Items were retained if they loaded uniquely onto one factor (i.e., did not cross-load) with a recommended >.30 loading, with each factor containing at least four items to avoid becoming a unique variable (Fabrigar, Wegener, MacCallum, & Strahan, 1999). To assess model fit, the root mean square error of approximation (RMSEA; Steiger, 1990), standardized root mean residual (SRMR; Bentler, 1990), Tucker–Lewis non-normed fit index (NNFI; Bentler & Bonett, 1980), and the comparative fit index (CFI; Bentler, 1990) were calculated. The RMSEA and SRMR indicate good fit at low values (<.06, .08–.10 moderate fit), and the NNFI and CFI indicate good fit at high values (> .95; Hu & Bentler, 1999).

**Confirmatory Factor Analysis.** After exploring the RS–14 with EFA, CFA with measurement invariance tests were examined. All models were tested using SPSS AMOS 18 employing maximum likelihood estimation. First, the model from the EFA was tested for an adequate fit using the following fit indexes: chi-square ($\chi^2$) and degrees of freedom, RMSEA including 90% confidence interval (CI; Steiger, 1990), SRMR, the Tucker–Lewis Index (TLI; Tucker & Lewis, 1973), and CFI. The latter indexes are used in tandem with $\chi^2$ values because $\chi^2$ can be influenced by sample size (Jöreskog & Sörbom, 1993) and might never approximate exact fit with nonsignificant values (MacCallum, Browne, & Sugawara, 1996). Therefore, small $\chi^2$, RMSEA, and SRMR values were desired along with high values of TLI and CFI. Multigroup CFA invariance tests (MGCFA) were then examined by sex and race/ethnicity for both the clinical and college sample groups. All analyses were examined using Brown’s (2006) guidelines, along with Cheung and Rensvold’s (2002) suggestions for examining partial invariance (i.e., models fit similarly for groups minus a few key item estimates).

After all data were examined for adequate fit, each individual subgroup was analyzed separately for model fit. These groups were then combined into one model for the test of configurational invariance. Configural invariance allows the researcher to examine if factor structure (i.e., where and if items load onto the factors) is the same across groups, but does not constrain estimates to be equal. If items load significantly onto their respective factors, and model fit is adequate, the two groups are said to have the same configuration of items onto factors. If not, individual items might not be tapping latent factors for specific groups. Models are then constrained to test metric invariance. Metric invariance examines how items load onto factors forcing them to be equal across groups, which indicates if the patterns of loadings are similar for groups (i.e., or some are positive for one group and negative for another group).

Next, item intercepts are constrained across groups to examine scalar invariance, which would indicate if groups have the same baseline item average. If not, this step contains important information about how items contribute to total or average score because they are different across subgroups. Finally, strict invariance is examined by forcing error variances to be equal, indicating that the scale reliably shows the

3Descriptive data for the DASS–21 and the PCL–S were reported in a separate investigation of the impact of the Gulf oil spill on Mississippi coastal residents, with specific regard for the impact of SES (see Drescher et al., 2014). They are reported here for context.
same pattern of error variance between groups. A noninvariant model at this step would indicate that there are differences in the spread of the scores for individual items, while they have the same averages.

To examine model invariance between these steps, we used $\Delta$CFI, where values greater than .01 indicate the need to examine partial invariance as suggested by Cheung and Rensvold (2002) and Chen (2007). Partial invariance is examined at the step in which the model has a significant degradation in fit. Each estimate is relaxed individually (i.e., allowed to vary across groups) to determine which items cause model fit to be poor. Partial invariance presents differences in estimates across groups that can indicate how questionnaires assess distinct populations differently. To demonstrate differences within samples, we calculated latent means for each participant and compared these means using $t$ tests (Cheung & Rensvold, 1999). Latent means are calculated by multiplying the standardized item loading by the participant score and averaging these weighted scores. Finally, descriptive data and correlations with similar measures are provided for reliability and validity examination of the RS–14.

**RESULTS**

Because samples were sufficiently large, both the clinical and college student groups were randomly split into two data sets for the exploratory ($N$ clinical sample = 516; $N$ college student sample = 883) and the confirmatory ($N$ clinical sample = 516; $N$ college student sample = 882) factor analyses. Data were screened for assumptions of multivariate statistics on each sample separately. Although samples contained multivariate outliers examined through Mahalanobis distance, results were similar when tested without outliers, and they were therefore included in all analyses. Other assumptions (normality, linearity, homogeneity, etc.) were found to be satisfactory using guidelines by Tabachnick and Fidell (2012).

**Exploratory Factor Analysis**

A scree plot and parallel analysis indicated that a one-factor solution was the most appropriate for both the clinical and college student samples. Table 1 contains the factor loadings for both samples. For the clinical sample affected by the Gulf oil spill, the one-factor model explained 53.2% of the variance, whereas the college student model explained 67.6% of the variance. As indicated in Table 1, the one-factor model showed excellent item loadings, and each item loaded onto the one resilience factor. The development of the original RS scale used five characteristics of resilience to create scale items, but a five-factor solution was inappropriate for the RS–14 scale. First, in line with Fabrigar et al. (1999), there are not enough items to adequately fit a five-factor solution. Second, an examination of eigenvalues in Table 2 clearly indicated a one-factor solution as compared to the parallel analysis 95th percentile eigenvalues provided (Preacher & MacCallum, 2003). Fit indexes for the clinical sample showed marginal values for RMSEA (0.11), good values for SRMR (0.05), and moderate values for NNFI (0.90) and CFI (0.92). These results are mirrored in the college student sample: RMSEA (0.12), SRMR (0.04), NNFI (0.92), and CFI (0.93). Because the RMSEA values were not desirable, we tested both models without Item 1 (whose loading was appropriately high at >.30, but was the lowest of all items in the scale). Fit indexes did not change, and therefore, this item was included in the analyses to remain consistent with the literature on the scale.

**Confirmatory Factor Analysis**

Fit indexes and sample sizes for all CFA and multigroup models are presented in Table 3. Fit indexes indicated moderate fit for both the clinical and college samples as a whole. Whereas SRMR indicated good fitting models, RMSEA was marginal (0.11) for the college sample. Both TLI and CFI were found to be between 0.90 and 0.95. However, all items significantly loaded onto the one resilience factor, indicating all items were indicators of the latent factor. Standardized residuals and modification indexes did not suggest the removal or change of pathways (i.e., correlated error terms). These results combined indicate at least a moderate fitting model.

**Separate Models.** The RS–14 data for both samples were next examined for model fit by sex and race/ethnicity using MGCFA. Due to extremely large sample size differences for race/ethnicity, we randomly chose a subsample of the White participants to match the smaller Black sample size for both the clinical sample ($n = 106$) and the college student sample ($n = 144$; Wannstrom, Peterson, Asberg, Nygren, & Gustavsson, 2009). Several interesting effects arose from examining the structure of the RS–14 across these variables. First, model fits for individual groups were roughly similar, minus fit fluctuations due to smaller sample sizes. Again, the pattern where SRMR indicated a good fitting model was noted, RMSEA was either moderate or above moderate, and TLI/CFI hovered around 0.90. The goal of MGCFA was to examine the nested model structure; therefore, although some of these fit indexes were not ideal, we examined combined structure as described in the data-analytic section.

**Clinical Sample.** For sex analyses in the clinical sample, loadings for configural invariance were large for both male (all loadings .67—.83) and female (all loadings .56—.81) participants. Further model fit was invariant until strict invariance
was examined ($\Delta$CFI > .01 between scalar and strict invariance) as seen in Table 3. At this step, we examined partial invariance by relaxing constraints for items, discovering that Items 9 (Male = 1.31 [SE = 0.13]; Female = 2.18 [SE = 0.21], $f_1 = .16$) and 13 (Male = 1.24 [SE = 0.13]; Female = 2.07 [SE = 0.19], $f_1 = .16$) showed higher variances for females than males ($\Delta$CFI = −.009). This finding could indicate that females are more likely to endorse a wider range of the Likert-type scale when answering items about interest or engagement (“I keep interested in things”) and meaning or purpose (“My life has meaning”), respectively, although the average score was the same.

For race/ethnicity, we also found large loadings for items onto the resilience factor (White range: .62–.81; Black range: .56–.87). Next, partial scalar invariance was found when examining race/ethnicity across the RS–14. As the scalar invariance test showed a significant drop in fit ($\Delta$CFI = −.017), we examined each item loading individually to see if relaxing constraints would improve model fit. Item 4 (self-regard, “I am friends with myself”), when allowed to vary, increased fit to meet partial scalar invariance ($\Delta$CFI = −.006). Black individuals (intercept = 5.12, SE = 0.18) rated the self-regard item higher than White individuals (intercept = 4.22, SE = 0.11, d = 0.59), indicating that Black individuals reported more self-regard.

**College Student Sample.** Unlike the clinical sample, the college student sample showed invariance across all models by sex. At configural invariance, loadings for individual sexes (male range = .72–.88; female range = .69–.88) were high and consistent. All other steps were invariant with small drops in CFI, as seen in Table 3. This finding indicated that college students were nearly homogenous in their answers on the scale, with the same model structure, factor loadings, intercepts, and error variances when considering sex. For race/ethnicity, configural invariance loadings were strong for both White (range = .72–.87) and Black participants (range = .73–.92). Then, we found that fit significantly degrades when considering strict invariance ($\Delta$CFI > .01). Each error variance was individually relaxed and allowing Item 12 to vary across White and Black groups created partial strict invariance ($\Delta$CFI = −.009). White (variance = .48, SE = .06) participants showed less variance when answering items about dependability (“In an emergency, I’m someone people can generally rely on”) than Black participants (variance = .97, SE = .12, $f_1 = .22$).

**Both Models.** Next, latent means were calculated for each sample and group for the entire data set. For the clinical sample, the male participants showed higher ($M = 3.40, SD = 1.09$) average weighted resilience than the female participants ($M = 3.01, SD = .94$), $t(1018) = 6.11$, $p < .001$, $d = .39$. The Black participants ($M = 3.33, SD = 1.11$) showed the same latent mean as the White participants ($M = 3.19, SD = .98$), $t(927) = 1.73$, $p = .08$, $d = .14$. For the college student sample, male ($M = 4.23, SD = .99$) participants showed a lower latent mean than female ($M = 4.41, SD = .98$) participants, $t(1763) = 3.79$, $p < .001$, $d = .18$, albeit with a small effect size. In contrast to the clinical sample, Black participants ($M = 4.68, SD = 1.27$) were significantly higher in
perceived resilience than White participants ($M = 4.31, SD = .95$), $t(1695) = 5.71, p < .001, d = 0.37$.

Descriptive Data and Reliability Analyses

Given the overall support for the RS–14 items as a one-factor model, basic descriptive statistics and reliability coefficients were calculated (Table 4). The college student sample reported more resilience ($M = 74.88, SD = 17.05$) than the clinical sample ($M = 63.11, SD = 19.87, t(2795) = 15.91, p < .001, d = 0.64$). In terms of Wagnild’s (2009a) scoring guidelines, the clinical sample was low overall, whereas the college student sample scored in the moderate range. With respect to differences in scores by sex or race/ethnicity, scores did not vary more than a few points from one another within each sample (Table 4). Thus, within each sample, there are not clinically or practically meaningful differences apparent when total scores are broken down by either of these demographic variables.

For both samples, alpha coefficients were excellent by a range of interpretive guidelines (e.g., DeVellis, 2012). Coefficient alphas for RS–14 scores ranged from .93 (clinical sample) to .96 (college student sample). Alpha coefficients were excellent whether examined by the total number of participants in the sample, by sex, or by race/ethnicity (Table 4).

Validity Support

Correlations between the RS–14 and other measures were examined by samples to determine scale relationship to other known related variables. As expected, for the clinical sample the RS–14 was positively related to measures of life satisfaction and meaning in life (SWLS $r = .46$, PIL–SF $r = .67, p < .001$) and negatively related to measures of psychological distress (PCL–S $r = -.25$; DASS Depression $r = -.40$; DASS Anxiety $r = -.28$, DASS Stress $r = -.31$, all $p < .001$). Similarly, for the college student sample, the RS–14 was positively correlated with measures of life satisfaction and meaning in life (SWLS $r = .72$, PIL–SF $r = .69, p < .001$) and negatively correlated with measures of psychological distress (PCL–C $r = -.33$; DASS Depression $r = -.29$; DASS Anxiety $r = -.22$, DASS Stress $r = -.18$, all $p < .001$).

**Table 4.**—Descriptive statistics for clinical ($N = 1,032$) and college student ($N = 1,765$) samples.

<table>
<thead>
<tr>
<th>Measure</th>
<th>Clinical Sample</th>
<th>College Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$M$</td>
<td>$SD$</td>
</tr>
<tr>
<td>RS–14</td>
<td>63.11</td>
<td>19.87</td>
</tr>
<tr>
<td>Male</td>
<td>64.19</td>
<td>20.47</td>
</tr>
<tr>
<td>Female</td>
<td>62.27</td>
<td>19.42</td>
</tr>
<tr>
<td>White</td>
<td>62.45</td>
<td>19.13</td>
</tr>
<tr>
<td>Black</td>
<td>64.52</td>
<td>21.38</td>
</tr>
</tbody>
</table>

*Note.* RS–14 = 14-item Resilience Scale.

Both clinical and undergraduate student samples. We hypothesized that for both samples, a one-factor model would be supported and that scores would be highly reliable as assessed by conventional interpretive standards. We also hypothesized that there could be possible significant differences in scale functioning between men and women and also between Black and White individuals. We hypothesized that the RS–14 would correlate positively with measures of positive psychological variables and negatively with measures of negative psychological variables.

EFA's and CFAs indicated that all items loaded cleanly onto a single factor consistent with cohesive structure for a one-factor model. This model supports a resilience factor, which confirms the original factor structure proposed by Wagnild (2009a). The removal of the lowest loading item in the EFA (Item 1) did not affect EFA and CFA model results. This finding indicated that fit was not negatively impacted by potentially poor items. Both samples showed invariance across models for sex indicating that males and females report similarly on RS–14 items. However, when partial invariance was analyzed within the clinical sample, it was discovered that two items (9 and 13) showed higher variances for females than males. These items examined interest and engagement (“I keep interested in things”) and meaning in life (“My life has meaning”) and are both part of the meaning aspect of the Resilience Core. Possible explanations for these invariances could be due to sex differences in how the items are interpreted, actual levels of perceived meaning, or a greater inclination to recognize and report levels of meaning in life. Although many studies do not find or report on sex differences in meaning, there are some data to support such a finding. For instance, Edwards and Holden (2003) found that women scored higher than men on a measure of perceived meaning and life purpose, and these scores were better able to predict presence of negative life events, such as suicide attempts (higher levels of meaning in life predicted fewer suicide attempts).

When differences in race/ethnicity were examined for the RS–14 in the clinical sample, analyses of partial scalar invariance found that Black individuals reported higher levels of self-regard (“I am friends with myself”) than White individuals. This finding is consistent with previous research in that meta-analyses have found that Black individuals rate themselves higher in terms of self-esteem (Gray-Little & Hafdahl, 2000; Twenge & Crocker, 2002). Within the college sample, partial invariance found that White participants had less variability in reporting their ability to be relied on in emergency situations than Black participants (Item 12). The reasons behind this finding are unclear. However, research has shown that Black individuals might have more experience with disaster situations and may be less prepared as a result of residential area and lower SES (Eisenman et al., 2009; Elliot & Pais, 2006; Norris et al., 2002). These results should be replicated for stronger support of differences in item responding, especially because clinical and college student samples did not show similar patterns of invariance. However, these differences are meaningful to report as a preliminary analysis into scale properties.

Latent means were then calculated, and there were significant differences in both samples. In the clinical sample, males had higher levels of perceived resilience than females, and there were no significant differences across race/ethnicity. In the college student sample, female participants had higher...
levels of perceived resilience than males, and Black participants had higher reported levels than White participants. These statistically significant differences between groups might be due to differences in SES (clinical sample median SES = <$14,999; college student sample median SES = $20,000–39,000), differences in age (clinical sample $M_{age} = 38.76; college student sample $M_{age} = 19.04$), or differences in stages of life with the clinical sample being older than the college student sample and thus having different responsibilities and experiences.

When comparing mean total scores, the college student sample reported higher levels of resilience than the clinical sample. It is possible that the lower levels of resilience in the clinical sample might be a reflection of their negative response to the Gulf oil spill. Data were collected as part of a larger battery investigating the mental health effects of the Gulf oil spill on affected Mississippi coastal residents. Individuals from the clinical sample were receiving mental health services as a result of this technological disaster. Some individuals were new clients and some were ongoing clients who had already been receiving services. The effects of the spill might have negatively impacted the individuals' perceived level of resilience or perceived ability to deal with the stress of the event, as this type of disaster often results in long-term ecological and economic ramifications. Speaking to these conclusions in a definitive fashion, however, is beyond the scope of this study. With respect to sex and race/ethnicity, there were no statistically significant differences in either sample.

To assess for validity support, correlations were calculated between the RS–14 and the PIL–SF and SWLS, as well as with the three scales of the DASS–21 and the PCL (PCL–S in the clinical sample, PCL–C in the undergraduate sample). As predicted for both the clinical and college student samples, scores on the RS–14 were significantly and positively correlated with scores on other measures of positive variables, specifically indexes of meaning and life satisfaction. Additionally, scores on the RS–14 were negatively correlated with scores on measures of psychological distress, specifically indexes of depression, anxiety, stress, and posttraumatic stress. These findings are consistent with a range of studies that note resilience's significant positive relationships with other adaptive concepts, as well as a significant negative relationship with psychological distress (Baldwin et al., 2011; Fredrickson et al., 2003; Scali et al., 2012; Tugade & Fredrickson, 2004).

Coefficient alphas were excellent in each sample regardless of demographic variables, indicating high internal consistency for RS–14 scores. Even though resilience is a construct that appears to be a combination of protective factors working together to produce positive outcomes to stressors, our analyses suggest that the RS–14 adequately examines this multifaceted construct considering the available reliability and structural validity support (Herbert, Manjula, & Philip, 2012; Lyons, 1991; Rutter, 1985; Wagnild & Young, 1990).

Strengths, Limitations, and Directions for Research

This study has several major strengths, including data gathered from a clinical sample affected by a recent controversial technological disaster (data difficult to collect), large sample sizes for both study groups, and sophisticated statistical analyses. These were strengths to this study, but there were also some limitations. In our view, the primary limitation was the lack of racial/ethnic diversity within both samples, which were predominantly White. Although this homogeneity was viewed as a study limitation, the large sample sizes did permit analyses by race/ethnicity so that potential differences in item response could be better detailed. Future studies should enhance the racial/ethnic diversity of the sample, as this would improve the generalizability of the results.

This study supports the use of the RS–14 in clinical and college samples, and reliability and validity data are reported, but the specific validity data documented relate to the measure's construct validity (i.e., factor structure and correlations). As construct validity data accrues over time across investigations, the measure would benefit from additional research. We regard these results as an extension of the RS–14's initial developmental work. We encourage additional research aimed at extending these present findings. For instance, future studies should focus on studying correlations with other measures of resilience, other positive psychology variables (e.g., optimism, self-efficacy, gratitude), and aspects of psychological distress (e.g., sleeping problems, eating problems, suicidality, substance abuse). Continued research on the concept of resilience is needed in varied contexts such as in the aftermath of different types of disaster as well as different age groups, including youth and older adults.

Future research should also continue to examine the RS–14's conceptual foundation. Characteristics of the Resilience Core have a foundation in Viktor Frankl’s logotherapy, which focuses on the significance of perceived meaning in life to the human condition (Frankl, 1959/2006; Schulenberg, Drescher, & Baczwaski, 2014; Schulenberg, Hutzell, Nassif, & Rogina, 2008; Wagnild, 2009a). Thus, one would expect similarities between the PIL–SF and the RS–14. In this study they correlated at .67–.69, sharing 45–48% of the variance, depending on the sample. Some of the overlap could be accounted for by the fact that the RS–14 includes items that assess meaning in life, similar to the PIL–SF. Although there is clearly substantial overlap, over half of the variance remains unaccounted for, consistent with the idea that the measures are assessing related, albeit distinct concepts. Therefore, with regard to measures of perceived meaning in life, such as the PIL–SF (Schulenberg & Melton, 2010; Schulenberg et al., 2011) and the Meaning in Life Questionnaire (Steger, Frazier, Oishi, & Kaler, 2006), to what extent would these measures possess incremental validity in relation to the RS–14? A broader body of research regarding the RS–14 will serve to better inform researchers and clinicians in the use of this measure, with particular regard for how the measure relates to other concepts, yielding greater insight into potential applications of the measure in the assessment and treatment of a range of concerns across a range of contexts.

Funding

Data for one of the samples were collected through a contract between Dr. Schulenberg and the Mississippi Department of Mental Health. The Mississippi Department of Mental Health funded mental health organizations through a grant program made possible by funding from British Petroleum. Dr. Schulenberg was hired to evaluate the clinical services, training, and outreach provided by the organizations funded under this grant.


