

Using the Assessment Framework to measure parental mood: an investigation of the reliability of the Adult Well-Being Scale

Christopher A. Pepping*, Sharon Dawe† and Paul H. Harnett‡

*Griffith University, Brisbane, Queensland, †Research Professor, Australian Centre for Child Protection, University of South Australia, Adelaide, South Australia, and ‡University of Queensland, Brisbane, Queensland, Australia

Correspondence:

Sharon Dawe,
School of Applied Psychology,
Griffith University,
Mt Gravatt Campus, 176 Messines
Ridge Road, Mt Gravatt, Qld. 4122,
Australia
E-mail: s.dawe@griffith.edu.au

Keywords: anxiety, assessment
framework, children, depression,
families, parent

Accepted for publication: October
2013

ABSTRACT

The adoption of evidence-based practice in social work has been widely promoted in recent years and with this, a growing emphasis on the evaluation of practice using well-validated and reliable measurement processes. The Department of Health's 'Framework for the Assessment of Children in Need and their Families' in the UK includes quantitative measures that form part of a systematic assessment of the needs of children and their families that includes assessment of parenting capacity and parental emotional state. The measure selected to assess parental mood was originally known as the Irritability, Depression and Anxiety Scale, and has been renamed within the Assessment Framework as the Adult Well-Being Scale. This instrument is designed to assess depression, anxiety, and inward and outward irritability. However, there has been relatively little contemporary evaluation of the reliability and validity of the measure, and the extent to which it measures the four constructs it is designed to assess. This research therefore conducted extensive analyses of the reliability, validity and underlying factor structure of the Adult Well-Being Scale. The four subscales did not demonstrate sound psychometric properties. At best a total score may be used as an indicator of 'overall psychological distress'.

INTRODUCTION

There has been a growing interest in the way in which evidence-based practice (EBP) is incorporated into social work. This has resulted in a proliferation of research and practice guidelines that investigate the conditions in which EBP is most likely to be implemented (Gray *et al.* 2013) and considered discussion of the limitations and advantages of such implementation (Knight 2013). As the field moves towards a majority view that practitioners should have the opportunity to analyse, evaluate and then synthesize evidence of practice effectiveness (Counoyer 2004; Knight 2013), government policy and practice guidelines have emerged that are intended to support practitioners in the adoption of EBP. The Department of

Health's 'Framework for the Assessment of Children in Need and their Families' in the UK provides a systematic assessment for the evaluation of the needs of children and their families (Department of Health 2000a). It reflects a key component of EBP that encourages the evaluation of practice through the use of objective measures of outcome (Baker & Ritchey, 2009). As part of a comprehensive assessment process, psychometric instruments were selected to provide reliable and valid measurement of key domains of concern in child and family functioning with eight of these published in the Family Pack of Questionnaires and Scales (Department of Health 2000a).

The use of such quantitative measures has typically been the domain of clinical psychology, and to a lesser

extent, psychiatry, whereas psychometric measures have not been used routinely in social work practice. However, there is a growing acknowledgement that dynamic factors associated with family functioning can be measured using well-validated psychological measures (Department of Health 2000b). Further, these measures can and should be used in routine social work practice, and 'structured professional judgment should now be recognised by practitioners alongside the need for a new "mindset" about the use of standardised instruments' (Barlow *et al.* 2012, p. 13). Information collected at assessment, including findings from psychometric instruments, provide the basis for discussion with the family about current concerns and areas of strength and difficulties. This may include, for example, difficulties around managing child behaviour based on the Strengths and Difficulties Questionnaire (SDQ) as well as difficulties in parental mood, currently measured using the Adult Well-Being Scale (AWBS). This, in turn, leads the identification of clear goals and a supportive intervention aimed to assess whether change is possible. The use of scales, in addition to observation, enables an evaluation whether there is capacity to change (Harnett 2007).

Importantly, the use of multiple sources of information during an assessment is recommended within the Assessment Framework, as reliance on any one source has clear limitations and is unlikely to provide a comprehensive or balanced view of the situation (Department of Health 2000b). However, the basic premise underscoring the use of measures is that they are reliable and valid measures of the domain of functioning (e.g. parenting skills) or construct in question (e.g. parental depression). Thus, it is important that practitioners, irrespective of discipline, have some understanding of the relative merits of the measures they are using, including reliability and validity.

Although the psychometric properties of many of the measures included within the Assessment Framework have been extensively investigated, e.g. the Alcohol Use Disorders Inventory and the SDQ, there has been relatively little investigation of the reliability and validity of a key measure of parental mental health: the AWBS. This measure originally published as the Irritability, Depression and Anxiety Scale (Snaith *et al.* 1978) was included in the Family Pack of Questionnaires and Scales as there is a well-established link between parental mood, such as anxiety and depression, and parenting practices (Bonds *et al.* 2002; Johnson *et al.* 2006), and parental psychopathology and poor child outcomes (Haskett

et al. 1995; Spence *et al.* 2002). The AWBS has four subscales: depression, anxiety, inward irritability and outward irritability. Items for the measure were initially selected from the Hostility and Direction of Hostility Questionnaire (Caine *et al.* 1967) and the Buss–Durkee Inventory (Buss & Durkee 1957), and were selected based on the relevance of the item to the construct of irritability, and additional items were added by the researchers (Snaith *et al.* 1978).

The original validation study provided some evidence of convergent validity, i.e. the extent to which each subscale correlated with other instruments that measure the same or related constructs. There were significant correlations between the depression and anxiety subscales and the well-established subscales of the Hamilton Scales ($r = 0.75$ and 0.70 , respectively; Hamilton 1967). In relation to outward irritability, there was a significant correlation between the scale and Snaith's own 9-point scale assessment, which was combined with reports from nurses and relatives regarding the client's outward hostility ($r = 0.79$). For inward irritability, there was a significant correlation with the client's own report of anger directed towards himself or herself ($r = 0.84$; Snaith *et al.* 1978).

Three further studies have explored the convergent validity of the AWBS, finding correlations with theoretically related constructs (Snaith & Taylor 1985; Aylard *et al.* 1987). For example, Snaith & Taylor (1985) reported a significant correlation ($r = 0.72$) between the depression subscale of the AWBS and the Montgomery–Åsberg Depression Rating Scale (Montgomery & Åsberg 1979), and between the anxiety subscale of the AWBS and the Hamilton Anxiety Scale ($r = 0.58$; Hamilton 1967). Further, Lester (1988) found significant correlations between the depression ($r = 0.57$) and anxiety ($r = 0.50$) subscales of the AWBS with the Beck Depression Inventory (Beck *et al.* 1961). Finally, Snaith & Taylor (1985) examined the correlations of the irritability subscales of the AWBS with a single item that tapped into irritability from the Beck Depression Inventory (Beck *et al.* 1961), and the inward ($r = 0.49$) and outward ($r = 0.51$) irritability subscales of the AWBS. They also found significant correlations between irritability determined by clinical interview and the inward ($r = 0.43$) and outward ($r = 0.69$) irritability subscales of the AWBS in women with postpartum depression, and between the interview and the inward ($r = 0.39$) and outward ($r = 0.56$) irritability subscales of the AWBS in individuals with obsessional psychopathology.

Although this provides some support for the convergent validity of the AWBS, more convincing evidence that each of the scales are valid measures of the constructs they were intended to represent would come from an investigation of the underlying factor structure of the instrument. Factor analysis will determine whether the individual items making up a scale group together to form distinct subsets of the items, or a factor (Tabachnick & Fidell 2007). To date, the one study examining the factor structure of the AWBS found that items did not load onto four distinct factors, but rather displayed many mixed loadings of individual items between the four subscales in a sample of undergraduates (Lester 1988). This suggests that the AWBS may not be measuring four specific factors, but rather generalized emotional disturbance. Thus, although there are some initial promising findings for the AWBS, it is notable that there has been little contemporary evaluation of the psychometric properties of the scale. In particular, there has been little investigation of the reliability, and the underlying factor structure of the scale would provide confidence that each of the scales measure parental depression, parental anxiety and parental irritability. Further, apart from suggesting that irritability is a feature associated with many disorders, it is difficult to extrapolate exactly what this may mean for parenting capacity. Conceivably these scales may measure emotional regulation, certainly a critical construct that is clearly linked with parenting capacity and good child outcome (Bariola *et al.* 2011).

Finally, there has not been any analysis of the internal reliability of the AWBS generally considered to be a starting point in psychometric analysis of measures. Internal reliability refers to the extent to which items from a particular instrument or subscale correlate (Dawe *et al.* 2002). Thus, items that assess one construct should all be scored in a similar way: if a participant scored high on one 'depression' item, then they should do likewise on another 'depression' item. Internal reliability is then calculated using a split-half reliability approach or an investigation of the internal consistency using Cronbach's alpha (α). Split-half reliability splits an instrument (or the subscales of the instrument) in half, the instrument or subscale is then scored separately and the scores from one-half of the instrument are then correlated with scores from the other half (Dawe *et al.* 2002). Higher correlations between the two halves of the same instrument (or subscale) reflect higher split-half reliability. Cronbach's alpha assesses the extent to which the items assessing the same construct correlate with each

other, and scores can range from 0 (completely unreliable) to 1 (perfectly reliable). However, the minimal alpha coefficient to be considered adequate is 0.7 (Nunnally 1994).

The purpose of this brief report was to investigate the internal reliability, the convergent validity and the construct validity using factor analysis of the AWBS. This is particularly important as there is a strong literature linking parental depression and anxiety to poor child outcome. Further, having some understanding of the constructs related to the irritability scales will also help in understanding the meaning of the scores and provide context for treatment planning.

METHOD

Participants were 329 undergraduate students enrolled in an introductory psychology course (241 women and 88 men, ranging in age from 16 to 55; $M = 21.53$ years, standard deviation = 6.59). Participants completed an online questionnaire for course credit. The study received ethical approval from the authors' institution, and participants read an information sheet and then moved to a consent screen to indicate agreement with participation in the anonymous survey.

The Depression, Anxiety and Stress Scale-21 (DASS-21; Lovibond & Lovibond 1995), a 21-item scale, is a psychometrically valid measure of depression, anxiety and stress (Lovibond & Lovibond 1995; Henry & Crawford 2005). It has excellent internal consistency, $\alpha = 0.92$, 0.81 and 0.88 for depression, anxiety and stress, respectively (Clara *et al.* 2001). Importantly, the DASS-21 is sensitive to meaningful clinical change (Ng *et al.* 2007), which indicates that it is a useful measure for routine clinical work. It was hypothesized that the depression and anxiety scales of the DASS and AWBS, respectively, would be highly correlated, thus demonstrating convergent validity of the AWBS.

The Manifest Anxiety Scale-Short Form (Bendig 1956) is a 20-item scale designed to measure trait anxiety. It has good psychometric properties and adequate internal consistency, $\alpha = 0.79$ (Thibodeau 2011). The Manifest Anxiety Scale-Short Form should correlate highly with the anxiety subscale of the AWBS, i.e. convergent validity, but less so with the depression subscale, i.e. divergent validity.

The constructs of inward and outward irritability are less well established. However, the measures selected to ascertain convergent validity with these constructs were developed to measure aspects of each. The Multidimensional Anger Inventory (Siegel 1986) was

included as a measure of inwardly and outwardly directed affect. This is a 38-item scale that assesses anger as a multidimensional construct. For the purposes of the present research, only the inwardly directed and outwardly directed anger dimensions were included. Inwardly directed anger refers to the tendency to suppress or limit expressions of anger, whereas outwardly directed anger refers to the tendency to engage in verbal or physical forms of aggression (Siegel 1986). Internal consistency for the combined inventory is high ($\alpha = 0.88$) and internal consistency for the anger in subscale was acceptable ($\alpha = 0.72$), but was poor for the anger out subscale ($\alpha = 0.50$; Siegel 1986). These subscales were used to examine the convergent validity of the AWBS. Specifically, the tendency to direct negative affect inward (inwardly directed anger) should be expected to correlate highly with the inward irritability subscale of the AWBS, whereas the tendency to direct negative affect outward (outwardly directed anger) should correlate highly with the outward irritability subscale of the AWBS.

The Difficulties in Emotion Regulation Scale (DERS; Gratz & Roemer 2004) was used as a global measure of emotion regulation. This is a 36-item scale that consists of six subscales of difficulties in emotional regulation: non-acceptance, goals, impulse control, awareness, strategies and clarity. The scale has excellent construct validity and high internal consistency, $\alpha = 0.93$ for the total score (Gratz & Roemer 2004). Given that unhealthy expression of anger is a component of difficulties in emotion regulation, emotion regulation difficulties as measured by the DERS should correlate positively with the inward and outward irritability subscales of the AWBS to demonstrate convergent validity. Given the considerable overlap between the construct of impulsivity and outwardly expressed irritability, a measure of irritability should correlate highly with outward irritability to demonstrate convergent validity. The measure chosen to measure impulsivity was the Barratt Impulsiveness Scale-11 (Patton *et al.* 1995), which is a 34-item scale that assesses behavioural and biological correlates of impulsivity. The scale has good internal consistency reported in the literature ($\alpha = 0.82$; Patton *et al.* 1995).

RESULTS

Reliability

Initially, the reliability of each of the subscales and of the total score was explored. Split-half reliability was

Table 1 Internal consistency (Cronbach's alpha) for convergence measures

Scale	Cronbach's alpha (α)
AWBS depression	0.55
AWBS anxiety	0.73
AWBS irritability (in)	0.81
AWBS irritability (out)	0.67
AWBS total score	0.87
DASS depression	0.91
DASS anxiety	0.86
DASS stress	0.87
Trait anxiety	0.86
Anger (in)	0.71
Anger (out)	0.66
BIS (total)	0.82
DERS (total)	0.94

AWBS, Adult Well-Being Scale; BIS, Barratt Impulsiveness Scale; DASS, Depression, Anxiety and Stress Scale; DERS, Difficulties in Emotion Regulation Scale.

poor for the depression subscale ($r = 0.56$), and moderate for the anxiety ($r = 0.66$) and outward irritability ($r = 0.67$) subscales. Adequate split-half reliability was found for the inward irritability subscale ($r = 0.84$) and for the total score ($r = 0.81$). Internal consistency as measured by Cronbach's alpha was poor for depression and moderate to acceptable for outward irritability, anxiety, inward irritability and the total score (see Table 1). The internal reliability for all other measures was also undertaken (see Table 1).

Convergent validity

To investigate the convergent validity of the four subscales of the AWBS, the correlations between the subscales and other theoretically and empirically relevant measures were calculated (see Table 2). Of particular note was the high correlation between the depression subscale of the AWBS and the anxiety subscale of the AWBS, suggesting that these two scales were not measuring distinct constructs. There were also significant correlations between depression AWBS and DASS depression, DASS anxiety and trait anxiety. Similar patterns were also found with the anxiety subscale of the AWBS. Convergence between measures was somewhat poor for both inward and outward irritability. The inward irritability subscale of the AWBS was only modestly correlated with the inwardly directed anger subscale of the Multi-Dimensional Anger Inventory, although highly correlated with all three subscales of the DASS and with the trait anxiety measure. The outward irritability

Table 2 Correlation matrix for Adult Well-Being Scale and related measures ($n = 329$)

	1	2	3	4	5	6	7	8	9	10	11	12	13
1 AWBS depression	—												
2 AWBS anxiety	0.61**	—											
3 AWBS irritability (in)	0.56**	0.65**	—										
4 AWBS irritability (out)	0.40**	0.40**	0.46**	—									
5 AWBS total score	0.80**	0.85**	0.85**	0.68**	—								
6 DASS depression	0.63**	0.61**	0.71**	0.33**	0.72**	—							
7 DASS anxiety	0.55**	0.68**	0.62**	0.36**	0.70**	0.70**	—						
8 DASS stress	0.57**	0.69**	0.58**	0.41**	0.71**	0.73**	0.72**	—					
9 Trait anxiety	0.53**	0.66**	0.60**	0.40**	0.70**	0.63**	0.60**	0.67**	—				
10 Anger (in)	0.32**	0.42**	0.32**	0.52**	0.49**	0.30**	0.30**	0.45**	0.48**	—			
11 Anger (out)	-0.11*	-0.17**	-0.09	0.12*	-0.09	-0.14**	-0.14**	-0.16**	-0.20**	0.00	—		
12 BIS (total)	0.25**	0.31**	0.34**	0.28**	0.37**	0.33**	0.28**	0.27**	0.35**	0.35**	0.04	—	
13 DERS (total)	0.53**	0.59**	0.65**	0.46**	0.70**	0.67**	0.61**	0.67**	0.75**	0.49**	-0.21**	0.41**	—

* $P < 0.01$, ** $P < 0.001$.

Anger (in), Multidimensional Anger Inventory, inwardly directed anger; Anger (out), Multidimensional Anger Inventory, outwardly directed anger; AWBS, Adult Well-Being Scale; BIS, Barratt Impulsiveness Scale; DASS, Depression, Anxiety and Stress Scale; DERS, Difficulties in Emotion Regulation Scale.

subscale of the AWBS showed a very low correlation with the outwardly directed anger subscale of the Multi-Dimensional Anger Inventory, and a low correlation with impulsivity.

Factor structure

An initial exploratory principal component analysis was performed on the 18-item AWBS to ascertain if the items fell broadly into groups of related items that corresponded to the proposed subscales. With the present sample of 329 participants and four subscales of the AWBS, the sample size was adequate for factor analysis. Sample size and factorability were acceptable (Kaiser–Meyer–Olkin measure [KMO] = 0.87), and the initial correlations within the 18 items were sufficiently high to warrant further factor analysis (Bartlett's test of sphericity; $P = 0.000$). Exploratory factor analysis revealed that the items fell into five factors with eigenvalues >1 , explaining 59.89% of the variance in total. The factors accounted for 33%, 9%, 7%, 6% and 6% of the variance for factors 1, 2, 3, 4 and 5, respectively.

As anxiety and depression measures are often correlated with one another, it was important to ensure that this was taken into consideration in the next phase of testing the factor structure, and therefore oblimin rotation was performed. Items were considered to load onto a factor with loadings of 0.30 or greater. Ideally, if items fall neatly onto clear factors, there would be no cross loadings, but rather a single score per item. Factor 1 corresponded to several of the anxiety items, with some split loadings for inward irritability. Factor 2 corresponded with the outward irritability items and had one split loading for a depression item. Factor 3 loaded onto an anxiety item and several depression items. However, again, there were several split loadings. Factor 4 had two split loadings for depression. Factor 5 loaded onto several inward irritability items, and a depression and outward irritability item. Thus, the factor structure was not acceptable, as there was not a clean solution of factors that corresponded with the four subscales of the AWBS (see Table 3).

A second exploratory principal component factor analysis with oblimin rotation was therefore performed, but this time we forced a four-factor solution given that the AWBS consists of four subscales. The factors accounted for 33%, 9%, 7% and 6% of the variance for factors 1, 2, 3 and 4, respectively. Table 4 displays the factor loadings from the factor analysis. The four-factor solution provided ease of

Table 3 Adult Well-Being Scale items and their loadings of the exploratory factor analysis

Items	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
I feel tense or 'wound up' (anxiety)	0.854	0.028	0.167	-0.037	0.153
I have an uncomfortable feeling like butterflies in the stomach (anxiety)	0.803	0.045	-0.005	-0.071	-0.068
I get scared or panicky for no good reason (anxiety)	0.525	-0.029	0.162	0.133	-0.277
I can sit down and relax quite easily (anxiety)	0.131	-0.053	0.753	0.016	-0.041
I can go out on my own without feeling anxious (anxiety)	0.032	0.243	0.284	-0.137	-0.045
I can laugh and feel amused (depression)	0.049	0.034	0.742	0.228	-0.076
I feel cheerful (depression)	0.233	-0.045	0.631	-0.050	-0.155
I'm awake before I need to get up (depression)	-0.056	0.137	0.429	0.726	0.029
I've kept up my old interests (depression)	0.025	0.301	0.319	-0.580	0.018
My appetite is ... (depression)	-0.072	-0.146	0.245	-0.258	-0.533
Lately I have been getting annoyed with myself (inward)	0.399	0.124	0.043	-0.057	-0.367
The thought of hurting myself occurs to me (inward)	0.011	0.069	-0.002	-0.021	-0.842
I feel like harming myself (inward)	0.163	0.023	0.079	0.175	-0.677
I get angry with myself or call myself names (inward)	0.394	0.095	-0.024	0.054	-0.541
I am patient with other people (outward)	-0.157	0.730	0.171	-0.178	0.126
I lose my temper and shout and snap at others (outward)	0.228	0.701	-0.143	0.087	-0.016
People upset me so that I feel like slamming doors or banging about (outward)	0.145	0.651	-0.052	0.069	-0.098
I feel I might lose control and hit or hurt someone (outward)	-0.131	0.527	0.004	0.125	-0.477

Note: Boldface denotes the column where the particular item loads most strongly.

Table 4 Adult Well-Being Scale items and their loadings of the exploratory four-factor analysis

Items	Factor 1	Factor 2	Factor 3	Factor 4
I feel tense or 'wound up' (anxiety)	0.680	-0.045	0.125	-0.170
I have an uncomfortable feeling like butterflies in the stomach (anxiety)	0.795	-0.002	-0.021	-0.180
I get scared or panicky for no good reason (anxiety)	0.685	-0.041	0.123	0.114
I can sit down and relax quite easily (anxiety)	0.114	-0.060	0.749	0.118
I can go out on my own without feeling anxious (anxiety)	0.034	0.250	0.320	-0.076
I can laugh and feel amused (depression)	0.070	0.034	0.690	0.332
I feel cheerful (depression)	0.293	-0.042	0.646	0.043
I'm awake before I need to get up (depression)	-0.065	0.119	0.247	0.740
I've kept up my old interests (depression)	-0.041	0.314	0.462	-0.488
My appetite is ... (depression)	0.288	-0.080	0.351	-0.094
Lately I have been getting annoyed with myself (inward)	0.624	0.134	0.062	-0.042
The thought of hurting myself occurs to me (inward)	0.601	0.153	0.060	0.130
I feel like harming myself (inward)	0.634	0.075	0.073	0.269
I get angry with myself or call myself names (inward)	0.751	0.120	-0.020	0.081
I am patient with other people (outward)	-0.272	0.734	0.210	-0.126
I lose my temper and shout and snap at others (outward)	0.217	0.685	-0.180	0.032
People upset me so that I feel like slamming doors or banging about (outward)	0.194	0.649	-0.074	0.057
I feel I might lose control and hit or hurt someone (outward)	0.205	0.581	0.008	0.227

Note: Boldface denotes the column where the particular item loads most strongly.

interpretation, with fewer split loadings. Factor 1 corresponded with inward irritability and anxiety items, and factor 2 with outward irritability. Factor 3 consisted of two anxiety items and several depression items; however, two of these were split loadings. Factor 4 corresponded with three depression items, two of which were split loadings. This indicates that

the only factor with a clean solution was the outward irritability scale.

Given that one of the key measures we used to explore the validity of the AWBS was the DASS-21, we examined the factor structure of this instrument to ensure that the data in the present sample were valid and consistent with the three-factor structure of this

Table 5 Depression, Anxiety and Stress Scale-21 items and their loadings of the exploratory three-factor analysis

Items	Factor 1	Factor 2	Factor 3
I felt that life was meaningless (depression)	0.959	0.010	-0.122
I felt that I had nothing to look forward to (depression)	0.826	0.068	0.015
I felt I wasn't worth much as a person (depression)	0.802	0.077	0.051
I was unable to become enthusiastic about anything (depression)	0.761	0.004	0.088
I couldn't seem to experience any positive feeling at all (depression)	0.754	0.057	0.078
I felt downhearted and blue (depression)	0.624	0.012	0.264
I was aware of the action of my heart in the absence of physical exertion (anxiety)	-0.014	0.821	-0.040
I experienced breathing difficulty (anxiety)	-0.124	0.800	0.125
I was aware of dryness of my mouth (anxiety)	-0.045	0.719	-0.085
I experienced trembling (e.g. in the hands; anxiety)	0.208	0.607	0.032
I was worried about situations in which I might panic and make a fool of myself (anxiety)	0.207	0.561	0.063
I felt I was close to panic (anxiety)	0.273	0.458	0.211
I felt scared without any good reason (anxiety)	0.202	0.546	0.048
I felt that I was using a lot of nervous energy (stress)	-0.023	0.486	0.410
I found it hard to wind down (stress)	-0.159	0.021	0.856
I found it difficult to relax (stress)	0.033	0.079	0.762
I felt that I was rather touchy (stress)	0.043	0.035	0.689
I was intolerant of anything that kept me from getting on with what I was doing (stress)	0.103	-0.064	0.682
I found myself getting agitated (stress)	0.295	-0.052	0.617
I tended to overreact to situations (stress)	0.122	0.132	0.559
I found it difficult to work up the initiative to do things (depression)	0.226	0.106	0.303

Note: Boldface denotes the column where the particular item loads most strongly.

measure. We performed a principal components factor analysis with oblimin rotation, again, because the factors were correlated. Sample size and factorability were acceptable ($KMO = 0.96$), and Bartlett's test of sphericity reached statistical significance, $P = 0.000$, therefore indicating that factor analysis was appropriate. A three-factor solution was obtained, and all but 2 of the 21 items loaded onto the appropriate factor. One item from the stress subscale had a split loading with stress and anxiety, and one depression item loaded onto stress rather than depression. Overall, the three factors explained 60.93% of the variance. Table 5 displays the factor loadings from this factor analysis. These findings demonstrate a clear three-factor solution for the DASS-21 in the present sample, with almost perfect loadings onto the appropriate factors.

DISCUSSION

The Assessment Framework (Department of Health 2000a) provides extensive guidance around the assessment process. It is based on an ecological approach with a focus on the child's developmental needs, the parents' or caregivers capacities to respond appropriately, and the wider family and environmental factors. Principles of assessment, contextual considerations and process issues around engaging families are clearly

articulated. As part of a comprehensive assessment that includes interview and observational assessments, the Assessment Framework incorporates the use of measures that support practitioners to gather information relating to each of the domains of the Assessment Framework. Implicit in this is the assumption that the selected instruments are reliable and valid measures of the construct being measured. Although several early studies investigated the psychometric properties of the measure and found somewhat mixed support for the quality of the AWBS (e.g. Snaith & Taylor 1985; Lester 1988), there has been no contemporary analysis of the psychometric properties of the instrument, which is clearly needed if it is to be included within the Assessment Framework. This report therefore aimed to investigate the internal reliability, validity and underlying factor structure of the AWBS.

The internal consistency of the AWBS was investigated using both split-half reliability and Cronbach's alpha. Higher alpha coefficients demonstrate that the individual items correlate more strongly with each other, and items that form a single domain or subscale should correlate strongly. It was notable that all, but one of the alpha coefficients obtained for the comparative measures were in the acceptable range (see Table 1). However, the alpha coefficients obtained for the depression were particularly low (0.55), indicating that these items are not endorsed in a consistent

manner by respondents. The anxiety and inward irritability subscales demonstrated acceptable internal consistency, as did the total score for the AWBS. The split-half reliability of the subscales was poor for depression, moderate for anxiety and outward irritability, and adequate for inward irritability and the total score. Thus, the depression subscale consistently demonstrated poor reliability, and the anxiety and outward irritability showed little evidence of reliability. The inward irritability subscale appears to be the most reliable of the four subscales, demonstrating adequate internal consistency using Cronbach's alpha and split-half reliability.

The convergent and discriminant validity of the AWBS were investigated through an examination of the correlations between the AWBS dimension scores and constructs from well-established measures. The depression subscale of the AWBS correlated highly with the depression subscale from the DASS-21, as well as the measure of difficulties in emotion regulation (DERS), which provides evidence of adequate convergent validity. The anxiety subscale of the AWBS demonstrated high correlations with the DASS-21 anxiety subscale, as well as the trait anxiety scale, again, suggesting that this subscale has adequate convergent validity. Only a modest correlation was found between the inward irritability subscale of the AWBS with the inwardly directed anger subscale, which raises questions about the convergent validity of inward irritability subscale. Further, it correlated very highly with all three subscales of the DASS-21, with trait anxiety and with the DERS, which suggests that the inward irritability subscale may be measuring general emotional disturbance or negative affect. Only low correlations were found between the outward irritability subscale of the AWBS and the outwardly directed anger, and impulsivity. Interestingly, outward irritability correlated most highly with inwardly directed anger. Again, this raises questions about the validity of the outward irritability subscale and suggests that it may not be tapping into processes specific to outward irritability, but may instead be measuring general emotional disturbance. Further evidence of this is that the total score of the AWBS correlated more highly with all subscales of the DASS-21, the DERS and with impulsivity than did any of the individual AWBS subscales. Again, this suggests that the individual subscales are not tapping into the specific constructs they are designed to assess.

A high value for Cronbach's alpha indicates good internal consistency of the items in the scale. However, it does not mean that the scale or subscale is

unidimensional, and thus the underlying dimensionality was further investigated using factor analysis. First, it is important to note that although anxiety and depression are clearly distinct disorders, distinguishing between these two constructs using either clinician report or psychological measures has been difficult. Many studies have found that these two constructs are highly correlated and tend to tap a common factor of general distress or negative affectivity (Henry & Crawford 2005). However, when measures incorporate specific components of anxiety such as physiological arousal (e.g. 'I was aware of dryness of my mouth' and 'I experienced trembling [e.g. in the hands]' from the DASS-21) and for depression low positive affectivity (e.g. 'I couldn't seem to experience any positive feeling at all' and 'I felt downhearted and blue' from the DASS-21), there is a clearer differentiation between these two constructs (Henry & Crawford 2005; den Hollander-Gijsman *et al.* 2010). The factor analysis of the AWBS did not reveal four distinct subscales (see Tables 3 and 4), but rather demonstrated considerable overlap between the subscales, with the exception of outward irritability. This again suggests that the instrument is tapping into general negative affect or emotional disturbance, rather than the specific constructs the scale purports to measure. The factor analysis of the DASS-21, however, revealed three distinct factors that corresponded almost perfectly with the three subscales (depression, anxiety and stress) that the instrument is designed to measure.

It is critically important that instruments used in child and family work have sound psychometric properties, as the information provided by such measures can complement a comprehensive assessment and assist in helping to understand the challenges facing families, areas of difficulty such as parental mood or child management difficulties. Further, repeated use of measures across time provides evidence for family capacity to change.

It is notable that none of the four subscales of the AWBS demonstrated acceptable psychometric properties across all three of the domains of reliability, convergent validity and construct validity based on the underlying factor structure. However, the total score of the AWBS did show somewhat better psychometric properties, particularly with regard to internal consistency and convergent validity. That is, not only did the factor structure not correspond with the four independent subscales, but the individual subscales demonstrated lower reliability than the total score, and also did not correlate as highly with related measures

as did the total score. Perhaps the use of the total score of the AWBS as a measure of general emotional disturbance in parents warrants consideration.

However, given that specific measures that tap into depression and anxiety are preferable, the DASS-21 may provide a more valid assessment of these constructs. The DASS-21 is a reliable and valid instrument that assesses the three underlying constructs it is designed to assess (Lovibond & Lovibond 1995; Henry & Crawford 2005), is easy to score and is also sensitive to change (Ng *et al.* 2007), which allows practitioners to administer the measure before and after interventions in order to monitor the effectiveness of interventions. Further, because the DASS-21 accurately assesses the underlying constructs it is designed to assess, it may be a better tool for generating discussion between practitioners and caregivers, around areas of difficulty in relation to depressed mood and feelings of anxiety and stress as the information provided is more accurate and is therefore less likely to provide misleading information that may negatively impact on these discussions. In addition to displaying sound psychometric properties, the DASS-21 is a freely available instrument (<http://www2.psy.unsw.edu.au/dass/download.htm>).

CONCLUSIONS, LIMITATIONS AND FUTURE DIRECTIONS

This paper argues that the addition of routine use of standardized measures complements information obtained during interview and observation. Such an approach is receiving growing support both in the UK and internationally, with a strong drive to evidence-informed practice in the field. Further, using such measures as part of a process in which parental capacity to change is under consideration is also receiving increasing support. Nonetheless, it is important to ensure that routine measures of symptoms are not confused with diagnostic status. This is an important distinction for the field. An elevated score on the attention and concentration scale of the SDQ does not lead to a diagnosis of attention deficit hyperactivity disorder (American Psychiatric Association 2013). Rather it provides an indication of the severity and frequency of a number of behaviour problems that can be associated with such a diagnosis, but can also reflect a range of other difficulties in a child's life, including inconsistent parenting and disruption due to multiple caregivers and/or exposure to trauma. Similarly, an elevated score on depression provides a score on symptom severity that is associated with a

diagnosis of depression, but note that the time frame for the completion of the DASS is in the last week, a time frame that is not consistent with Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (American Psychiatric Association 2013) criteria for depression. Nor is there a diagnostic category associated with 'stress'.

In addition to this general note of caution, there are several limitations of the present research that need to be acknowledged. First, the test-retest reliability of the AWBS was not investigated in the present research, which would have provided another source of information about the reliability of the measure over time. However, as Snaith *et al.* (1978) suggested, the test-retest reliability of the AWBS may not be a useful analysis, as the measure was originally designed to be sensitive to meaningful clinical change in mood state and therefore could reasonably be expected to show change over time. Second, the sample used in the present research was undergraduate students rather than parents of children at risk. This does raise an important issue related to the development and validation of psychological measures. For some measures or constructs, use of an undergraduate sample would be of concern (e.g. child abuse risk would not be reliably tested in young population of non-parents), for other measures not so. Constructs such as anxiety and depression do not differ across age and social status, and thus testing the reliability of an instrument focusing on these latter construct is less problematic in a student population.

Implicit in the forgoing discussion is the assumption that social work practice needs reliable and valid measures within EBP. However, defining EBP has not been without problems, and integration into routine practice has not been without reservations. The objections and counter arguments to the adoption of EBP have been prominent in the field for some years (e.g. Gibbs & Gambrill 2002); yet despite considerable academic enthusiasm, implementation has been patchy at best (Knight 2013). The provision of a set of measures readily available online and with scoring and interpretation guidance was an important first step in helping UK social workers access measures that could be used in practice. However, it is becoming clearer that simply providing practitioners with resources is not sufficient. The barriers to implementation of EBP include the development of a research culture in which practitioners feel they have the skills and knowledge base to make informed judgements about quality of research (Gray *et al.* 2013). We propose that the current study helps this by providing an assessable

analysis of a familiar measure using constructs that are integral to psychometric evaluation of all measures. It also highlights the importance of ensuring the psychometric robustness of measures prior to selection into database systems whenever possible.

In summary, the inclusion of a measure within the Assessment Framework that taps parental mood is sensible, and the measure selected for this purpose was the AWBS. However, the AWBS has had little contemporary assessment of its psychometric properties, and the present research was therefore concerned with this issue. The individual subscales of the AWBS (depression, anxiety, inward irritability and outward irritability) did not demonstrate acceptable psychometric properties, and factor analytic procedures showed that the instrument did not assess the four distinct constructs that it purports to measure. The total score of the AWBS was more psychometrically sound, and therefore, the use of the total score may be adequate as a measure of general emotional disturbance in parents. However, other measures such as the DASS-21 would provide a more valid and reliable assessment of depression and anxiety that would more accurately complement the assessment process. Although the current form of the AWBS requires some revision, the scale does have potential to provide valuable clinical information with further conceptual clarification and item development around the constructs of 'irritability'.

REFERENCES

- American Psychiatric Association (2013) *Diagnostic and Statistical Manual of Mental Disorders*, 5th edn. American Psychiatric Publishing, Arlington, VA.
- Aylard, P.R., Goodling, J.H., McKenna, P.J. & Snaith, R.P. (1987) A validation study of three anxiety and depression self-assessment scales. *Journal of Psychosomatic Research*, **31**, 261–268.
- Baker, L. & Ritchey, F. (2009) Assessing practitioner's knowledge of evaluation: initial psychometrics of the practice evaluation knowledge scale. *Journal of Evidence-Based Social Work*, **6**, 376–389.
- Bariola, E., Gullone, E. & Hughes, E.K. (2011) Child and adolescent emotion regulation: the role of parental emotion regulation and expression. *Clinical Child and Family Psychology Review*, **14**, 198–212. doi: 10.1007/s10567-011-0092-5.
- Barlow, J., Fisher, J. & Jones, D. (2012) *Systematic review of models for analysing significant harm*. Department of Education (UK). Research Report DFE-RR199.
- Beck, A.T., Ward, C. & Mendelson, M. (1961) An inventory for measuring depression. *Archives of General Psychiatry*, **4**, 561–571.
- Bendig, A.W. (1956) The development of a short form of the manifest anxiety scale. *Journal of Consulting Psychology*, **20**, 384.
- Bonds, D.D., Gondoli, D.M., Sturge-Apple, G. & Salem, L. (2002) Parenting stress as a mediator of the relation between parenting support and optimal parenting. *Parenting: Science and Practice*, **2**, 409–435.
- Buss, A.H. & Durkee, A. (1957) An inventory for assessing different kinds of hostility. *Journal of Consulting Psychology*, **21**, 343–349.
- Caine, T.M., Foulds, G.A. & Hope, K. (1967) *Manual of the Hostility and Direction of Hostility Questionnaire*. University of London Press, London.
- Clara, I.P., Cox, B.J. & Enns, M.W. (2001) Confirmatory factor analysis of the depression-anxiety-stress scales in depressed and anxiety patients. *Journal of Psychopathology and Behavioral Assessment*, **23**, 61–67.
- Cournoyer, B. (2004) *The Evidence-Based Social Work Skills Book*. Pearson, Boston, MA.
- Dawe, S., Loxton, N.J., Hides, L., Kavanagh, D.J. & Mattick, R.P. (2002) *Review of diagnostic screening instruments for alcohol and other drug use and other psychiatric disorders*. National Drug Strategy Monograph, Commonwealth Department of Human Services & Health. Available at: http://www.google.com.au/url?sa=t&rct=j&q=&esrc=s&source=web&cd=6&ved=0CEEQFjAF&url=http%3A%2F%2Fwww.nada.org.au%2Fmedia%2F14712%2Fscreening_assessment_review.pdf&ei=zDGsUK2oD6LGmQX0yoCgBA&usg=AFQjCNHpxAI4IP5NWdNmj_OScfMw6eGw4Q (accessed 10 November 2012).
- den Hollander-Gijsman, M.E., De Beurs, E., Van der Wee, N.J.A., Van Rood, Y.R. & Zitman, F.G. (2010) Distinguishing between depression and anxiety: a proposal for an extension of the tripartite model. *European Psychiatry*, **25**, 197–205.
- Department of Health (2000a) *Framework for the Assessment of Children in Need and Their Families*. Stationery Office, London.
- Department of Health (2000b) *Framework for the Assessment of Children in Need and Their Families: The Family Pack of Questionnaires and Scales*. Stationery Office, London.
- Gibbs, L. & Gambrill, E. (2002) Evidence-based practice: counterarguments to objections. *Research in Social Work Practice*, **12**, 452–475.
- Gratz, K.L. & Roemer, L. (2004) Multidimensional assessment of emotion regulation and dysregulation: development, factor structure, and initial validation of the difficulties in emotion regulation scale. *Journal of Psychopathology and Behavioral Assessment*, **26**, 41–54.
- Gray, M., Joy, E., Plath, D. & Webb, S.A. (2013) Implementing evidence-based practice: a review of the empirical research literature. *Research on Social Work Practice*, **23**, 157–166.
- Hamilton, M. (1967) The development of a scale for primary depressive illness. *The British Journal of Social and Clinical Psychology*, **6**, 278–296.
- Harnett, P.H. (2007) A procedure for assessing parents' capacity for change in child protection cases. *Children and Youth Services Review*, **29**, 1179–1188.

- Haskett, M.E., Myers, L.W., Pirrello, V.E. & Dombalis, A.O. (1995) Parenting style as a mediating link between parental emotional health and adjustment of maltreated children. *Behavior Therapy*, **26**, 625–642.
- Henry, J.D. & Crawford, J.R. (2005) The short-form version of the depression anxiety and stress scales (DASS-21): construct validity and normative data in a large non-clinical sample. *The British Journal of Clinical Psychology*, **44**, 227–239.
- Johnson, J.G., Cohen, P., Kasen, S. & Brook, J.S. (2006) Personality disorders evident by early adulthood and risk for anxiety disorders during middle adulthood. *Journal of Anxiety Disorders*, **20**, 408–426.
- Knight, C. (2013) Social workers' attitudes toward peer-reviewed literature: the evidence base. *Journal of Teaching in Social Work*, **33**, 177–195.
- Lester, D. (1988) Examination of a scale to assess irritability. *Personality and Individual Differences*, **9**, 935–936.
- Lovibond, P.F. & Lovibond, S.H. (1995) The structure of negative emotional states: comparison of the Depression Anxiety Stress Scales (DASS) with the Beck Depression and Anxiety Inventories. *Behaviour Research and Therapy*, **33**, 335–343.
- Montgomery, S.A. & Åsberg, M. (1979) A new depression scale designed to be sensitive to change. *The British Journal of Psychiatry*, **134**, 382–389.
- Ng, F., Trauer, T., Dodd, S., Callaly, T., Campbell, S. & Berk, M. (2007) The validity of the 21-item version of the Depression Anxiety and Stress Scales as a routine clinical outcome measure. *Acta Neuropsychiatrica*, **19**, 304–310.
- Nunnally, J.C. (1994) *Psychometric Theory*, 3rd edn. McGraw-Hill, New York.
- Patton, J.H., Stanford, M.S. & Barratt, E.S. (1995) Factor structure of the Barratt impulsiveness scale. *Journal of Clinical Psychology*, **51**, 768–774.
- Siegel, J.M. (1986) The multidimensional anger inventory. *Journal of Personality and Social Psychology*, **51**, 191–200.
- Snaith, R.P. & Taylor, C.M. (1985) Irritability: definition, assessment and associated features. *The British Journal of Psychiatry*, **147**, 127–136.
- Snaith, R.P., Constantopoulos, A.A., Jardine, M.Y. & McGuffin, P. (1978) A clinical scale for the self-assessment of irritability. *The British Journal of Psychiatry*, **132**, 164–171.
- Spence, S.H., Najman, J.M., Bor, W., O'Callaghan, M.J. & Williams, G.M. (2002) Maternal anxiety and depression, poverty and marital relationship factors during early childhood as predictors of anxiety and depressive symptoms in adolescence. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, **43**, 457–469.
- Tabachnick, B.G. & Fidell, L.S. (2007) *Using Multivariate Statistics*, 5th edn. Allyn and Bacon, London.
- Thibodeau, R. (2011) Individual differences in depressive symptoms are associated with impaired incentive, but not aversive motivation. *Personality and Individual Differences*, **50**, 376–380.