

The Assessment of Emotional Distress Experienced by People with Mild Intellectual  
Disabilities: A Study of Different Methodologies.

Paul Bramston and Gerard Fogarty  
University of Southern Queensland

Contact:

Dr P Bramston

Psychology Department

University of Southern Queensland

Toowoomba, Qld 4350

Email: [bramston@usq.edu.au](mailto:bramston@usq.edu.au)

Full reference: Bramston, P., & Fogarty, G. (2000). The assessment of emotional distress experienced by People with mild intellectual disabilities: A study of different methodologies. *Research in Developmental Disabilities*, 21, 487-500.

## The Assessment of Emotional Distress Experienced by People with Mild Intellectual Disabilities: A Study of Different Methodologies

The systematic study of emotional disorders among people with an intellectual disability is a relatively recent phenomenon with most studies reporting the presence of affective symptomatology. A number of studies in the 1980s found that people with an intellectual disability experience at least as high, and perhaps higher, rates of emotional disturbance than the general population. On the basis of a survey of the research literature, Benson and Reiss (1984) estimated emotional disorder prevalence rates among people with an intellectual disability to be somewhere between 15 and 20 percent of the population. Lund (1985) put the figure as high as 30 percent. Iverson and Fox (1989) reviewed the literature and found incidence rates of emotional disorders varied greatly, ranging from 13 to 47 percent. In their own sample of 165 intellectually disabled adults, 35% reported at least one psychopathological disorder, although Iverson and Fox noted that reliable measurement was difficult with this population.

The varying incidence rates reported for these disorders may be largely a consequence of the variation in assessment techniques and measures used by researchers. When dealing with a population where measurement has always been difficult, different methodologies and measures are not just a matter of convenience or preference - as they sometimes are with the general population. There is a need for research into the limitations of assessment methods when dealing with people who have intellectual disabilities. Concerns over the validity of these techniques provided the motivation for the present study which examined the impact of three assessment methods - self-ratings, ratings by significant others (informant ratings), and ratings by trained clinicians - in the domains of anger, depression, and stress. We begin with a review of each of these three methodologies before describing instances of their application in the three domains.

The most common way to detect emotional disorders is probably through a visit to a clinician such as a family doctor or a psychologist who uses a clinical interview as the basis for a global diagnosis. This approach generally accesses high clinical expertise but, for people with an intellectual disability, it often involves a clinician with limited experience in the disability field. The second approach, having 'a significant other' provide information on behalf of the person, is more common among younger children and people with an intellectual disability. The significant other, also known as informant rating, is a convenient approach but the accuracy of the ratings depend heavily on the experience and insight of the rater, as well as on how well he/she knows the person being rated. Little is known about the comparability of ratings made by different informants (Kazdin & Petti, 1982). In the third approach, emotional disorder is assessed directly by asking the individuals to report on their symptoms. The use of this technique among people with an intellectual disability is especially problematic given that they frequently have expressive and receptive language deficits and perhaps an inability to grasp the task demands in testing situations (Heal & Sigelman, 1995).

A search through the literature on anger, depression, and stress shows that researchers have used all three methods of assessment. Benson and Ivins (1992), for example, studied anger across different ranges of intellectual disabilities using an adapted form of the Children's Inventory of Anger (Finch, Saylor, & Nelson, 1983). They used

self-report ratings on a 4-point Likert scale with the items read aloud to the participants. Reiss and Rojahn (1993) investigated the joint occurrence of depression and aggression in people with an intellectual disability using the Reiss Screen for Maladaptive Behaviour (Reiss, 1988). This instrument uses ratings on behalf of the person with an intellectual disability by a significant other (informant rating) on alphabetically listed symptoms. Anger among a group of people with a mild intellectual disability was assessed by Baker and Bramston (1997) using self-reported anger scores on an adaptation of the State-Trait Anger Inventory (STAXI: Spielberger, 1988). These authors concluded that their results supported the validity of self-reported anger by people with mild to moderate intellectual disability.

Research into depression among people with an intellectual disability is characterised by this same readiness to use different assessment techniques. Matson, Barrett, and Helsel (1988) administered the Children's Depression Inventory (CDI: Kovacs, 1985) within a clinical interview to compare the depression rates of groups with and without an intellectual disability. Meins (1993) used an informant rating version of the CDI to assess the depression levels of 798 people with an intellectual disability, the majority of whom had a mild or moderate disability. Direct care staff were used as informants and all 24 items on the scale were found to be relevant to a group with mild/moderate disability, leading the author to conclude that the informant version of the CDI is a suitable screening instrument for depression in adults with an intellectual disability. Benavidez and Matson (1993) used self-report to compare three depression scales, one of which was the CDI, on groups of adolescents with and without an intellectual disability.

In research on stress among intellectually disabled persons, Zetlin (1993) used clinical interviews, whilst Bramston and colleagues (Bramston & Bostock, 1994; Bramston & Fogarty, 1995; Fogarty & Bramston, 1997; Bramston, Fogarty, & Cummins, 1999) relied upon the self-report Lifestress Inventory, designed specifically for this population (see description later).

The use of different assessment methods is not a problem in itself but it is important to examine the convergent validity of these three main assessment techniques when used with disabled persons. The small amount of research that exists on this topic suggests that there may not be a great deal of convergence. For example, in a study involving intellectually average children, a comparison between self-report and informant rating showed a correlation of only .05 on the CDI (Kazdin & Petti, 1982). A study by Benson and Ivins (1992) compared self-report with informant ratings of anger, depression, and self-concept. On the anger scale, the correlation between self-report and informant rating was .13 and on the depression scale it was .26, suggesting there was little overlap between the assessment methods. Nearly twice as often the informants viewed the participant as angrier than they self-reported. Anger and depression were found to correlate .14 on self-report and .13 on informant rating. More recently, Benavidez and Matson (1993) compared the responses of people with and without an intellectual disability to three different depression measures. On the CDI, they found correlations of .48 between self-report and informant rating for people with an intellectual disability and .19 for those without an intellectual disability.

Clearly, in research designed to assess convergence among measurement techniques, there are worrying signs that convergence is weak. There is an obvious need

to determine equivalence across the various measurement techniques, particularly as we begin to further explore emotionality within people with an intellectual disability. The aim of the present study was to assess the degree of convergent validity among clinician ratings, informant ratings, and self-report ratings for people with mild/moderate intellectual disability across the affective domains of anger, depression, and stress. The multitrait-multimethod (MTMM) approach was used to address this question.

### Method

#### Participants

Questionnaires were administered to 147 consenting adults with a mild or moderate intellectual disability (88 male and 59 female). The participants were employees of one of five sheltered workshops situated in both urban and rural areas of South East Queensland. Ages ranged from 18 to 63 with a mean of 33 years. All participants lived with their families or in supported community based homes. Participants were selected on the following criteria: (a) individuals had been assessed as functioning in the mild to moderate range of intellectual disability by agencies (IQ between 40 and 69 on the WAIS-R); (b) they exhibited longstanding adaptive behaviour deficits which resulted in them requiring sheltered employment, and (c) they exhibited sufficient verbal skills to enable them to take part in a simple conversation.

#### Self-Report Measures

Participants responded to three instruments: an adaptation of the State-Trait Anger Expression Inventory, Research Edition (Spielberger, 1988), the Children's Depression Inventory (Kovacs, 1985), and the Lifestress Inventory (Bramston, Fogarty & Cummins, 1999). All three scales were administered in interview format. Where necessary, item wording was changed so that the item could be read as a question. For example, a statement such as "I am quick tempered" was changed to "Are you quick tempered?". Probe questions were used by the interviewer to confirm understanding and to seek further information. Pictorial representations of the response categories used in each scale were also available. Thus, pictures of buckets without water, partially filled, and completely full were available to those participants who preferred to respond by pointing to one of the pictures. A description of each self-report scales follows.

The State-Trait Anger Expression Inventory, Research Edition (STAXI: Spielberger, 1988) was adapted by Baker and Bramston (1997) to assess anger in people with an intellectual disability. In this adapted version of the STAXI, a four point Likert scale was used where 1 indicated no anger, 2 a little, 3 a fair bit, and 4 a lot of anger. This Likert Scale was substituted for the one generally used with the STAXI because it is readily understood by people with a mild or moderate intellectual disability (e.g., Bramston & Fogarty, 1995; Fogarty & Bramston, 1997). Anger self-report scores were obtained by calculating an average rating across all 20 items. The Cronbach alpha reliability estimate for this scale in the present study was .89.

The CDI is a 27-item scale modelled after the Beck Depression Inventory (Kovacs, 1981). It has a self-report version and was selected for this study because it has been used successfully among adults with an intellectual disability (e.g., Benavidez & Matson, 1993). Each item on the CDI offers respondents three statements of depressive symptoms from which they must choose the one which best describes how they feel. These alternatives are weighted 0 to 2 with higher scores indicating higher levels of

depression. Depression self-report scores were obtained calculating an average rating across the 27 items. The Cronbach alpha reliability estimate for this scale in the present study was .79.

The Lifestress Inventory has 33 items describing life events that have been identified in previous research as causing stress for people with an intellectual disability. The interviewer asks whether each event has happened recently (past few weeks). If the event has not been experienced, a score of 0 is recorded for that item. If it has been experienced, the participant is then asked to indicate the amount of stress caused by the event. The response categories were 1 to indicate that the event caused no stress, 2 a little stress, 3 a fair bit, and 4 a lot of stress. Two scores can be obtained from this instrument. The first is a frequency score indicating how many stressors have been encountered. The second is an impact score, obtained by averaging the scores on the 33 items. Both scores tend to be highly correlated. The impact score was used in the present study. As recommended by Fogarty and Bramston (1997), the 0 (stressor not encountered) and 1 (stressor encountered but not experienced as stressful) scoring categories were combined when calculating the impact scores. The Lifestress Inventory has been shown to be both reliable and valid in the assessment of perceived stress among people with an intellectual disability (Bramston & Fogarty, 1995; Fogarty and Bramston, 1997). The Cronbach alpha reliability estimate for the Lifestress Inventory in the present study was .86.

#### Clinical Interview

A qualified female psychologist interviewed participants individually and questioned them closely on levels of perceived depression, anger, and stress. This procedure was designed to follow the standard procedures set down for psychological consultations in a community health centre. At the end of the 15 minute consultation, the psychologist rated the participant for the degree of depression, anger, and stress perceived. Ratings used a scale from 1-5, with 1 indicating no sign of the emotion and 5 indicating extremely high levels of the particular emotion.

#### Ratings by a Significant Other

As in Meins' (1993) study, the sheltered workshop supervisor who knew the participant best was used as the rater. The supervisor was given a brief description of anger, depression, and stress. Anger was described as feelings of irritation, annoyance, fury, and rage; depression as feelings of sadness, pessimism, and hopelessness; and stress as feelings of irritability, anxiety, inadequate coping, and frustration. The supervisor was then asked to rate each participant over the past few weeks on each. The same 1-5 rating scale was used as in the clinical interview.

#### Procedure

The order of the self-report, clinical interview, and ratings by supervisor was varied across the sample to control for order of administration effects. A trained tester administered the Lifestress Inventory, the CDI, and the STAXI, also in counterbalanced order. The clinical interview and ratings by the supervisor similarly counterbalanced the order in which they rated depression, anger, and stress. Data collection took place over a one month period at the place of employment with all three measures for each participant being taken on the same day.

## Results

Data screening showed that there were no out of range values and very little missing data. Summary statistics and correlations for all nine variables are shown in Table 1.

Table 1

Summary Statistics and Correlations Among all Nine Variables (N = 147)

Variables	Correlations									
	Self-Report			Supervisor			Psychologist			
	<u>M</u>	<u>SD</u>	1	2	3	4	5	6	7	8
1. AngScale	1.36	.39								
2. DepScale	1.43	.29	.33							
3. StrScale	1.33	.38	.51	.45						
4. AngSuper	1.78	1.21	<b>.10</b>	<u>.13</u>	<u>-.08</u>					
5. DepSuper	1.72	1.08	<u>.08</u>	<b>.19</b>	<u>-.00</u>	.74				
6. StrSuper	1.94	1.11	<u>.15</u>	<u>.23</u>	<b>.07</b>	.62	.62			
7. AngPsych	1.43	.70	<b>.34</b>	<u>.30</u>	<u>.45</u>	<b>.05</b>	<u>.18</u>	<u>.31</u>		
8. DepPsych	1.67	.89	<u>.26</u>	<b>.28</b>	<u>.52</u>	<u>-.07</u>	<b>.08</b>	<u>.15</u>	.54	
9. StrPsych	1.48	.75	<u>.20</u>	<u>.35</u>	<b>.40</b>	<u>.11</u>	<u>.22</u>	<b>.29</b>	.71	.62

Note: Correlations above .16 (approx.) significant ( $p < .05$ ).

Correlations in boldface indicate convergent validity and underlined values are discriminant validity coefficients. The other correlations indicate common method effects.

Although the focus of this paper will be on the relations among the variables, it is interesting to also compare the ratings given by the psychologist with those given by the supervisor. These ratings were based on a scale that ranged from 1 to 5, with a score of 5 indicating an extreme degree of anger, stress, or depression and a rating of 1 indicating extremely low levels of these emotions. The ratings given by the psychologist tended to be lower and less variable across individuals than the ratings given by the supervisors. Both the psychologist and the supervisor used the full range of categories but the supervisor tended to be more liberal in the use of the higher categories. This tendency is reflected in the higher average ratings and larger standard deviations for supervisor ratings. T-tests for paired samples showed that the difference in ratings was significant for stress ( $t_{146} = 4.91, p < .01$ ) and anger ( $t_{146} = 3.07, p < .01$ ) but not for depression.

The correlations in Table 1 are arranged in the traditional MTMM format (see Campbell & Fisk, 1959). That is, the MTMM matrix is divided into triangular submatrices of relations among different traits assessed with the same method (heterotrait-monomethods: HTMM), square matrices of relations among measures assessed with different methods: HTHM), and relations among the same traits assessed with different methods (convergent validities). If these three different methods are

suitable for assessing anger, depression, and stress among people with an intellectual disability then the convergent validity indices in the MTMM matrix should exhibit certain properties. Specifically, Campbell and Fiske (1959) proposed four guidelines to help assess these indices.

The first guideline is that convergent validity indices should be substantial. These indices are shown in boldface in Table 1. It can be seen that four of the nine were not significant, all four involving relations between supervisor ratings and the other methods of assessment. The remaining convergent validity indices were significant but hardly "substantial", the largest being .40. Thus, support for this guideline was weak. The second guideline is that convergent validities should be higher than HTHM correlations. The average convergent validity correlation in Table 1 was .20, which was the same as the average HTHM correlation, indicating weak support for this guideline as well. The third guideline is that convergent validity coefficients should be higher than HTMM correlations. In the present instance, the average HTMM correlation (.57) was much higher than the average validity coefficient (.20), indicating lack of support for this guideline as well. The fourth guideline is that the pattern of correlations among different traits should be similar for different methods. This is not the case in Table 1. The highest correlation for the "scale" and "psychologist" methods was between the depression and stress constructs whereas for the "supervisor" method it was between anger and depression. The mismatch between the methods is highlighted by the fact that the correlation between anger and depression, whilst it was the highest for the supervisor ratings, was actually the lowest for the scale and psychologist ratings. On the bases of these rather simple criteria, one would have to conclude that evidence for convergent validity is lacking in these data. None of the guidelines specified by Campbell and Fiske (1959) was satisfied.

Marsh and Grayson (1995) suggested further testing of the MTMM data by use of various confirmatory factor analytic (CFA) models. Many variants of these MTMM models have been proposed in the literature but here we followed the approach advocated by Byrne (1998). The first model to be tested in this approach is what is called the general CFA model, a relatively unrestrictive baseline model against which all others can be compared (Byrne, 1998, p. 199). Its specification includes both trait and method factors and allows for correlations among traits and among methods. Figure 1 shows the form that it takes in the present instance.

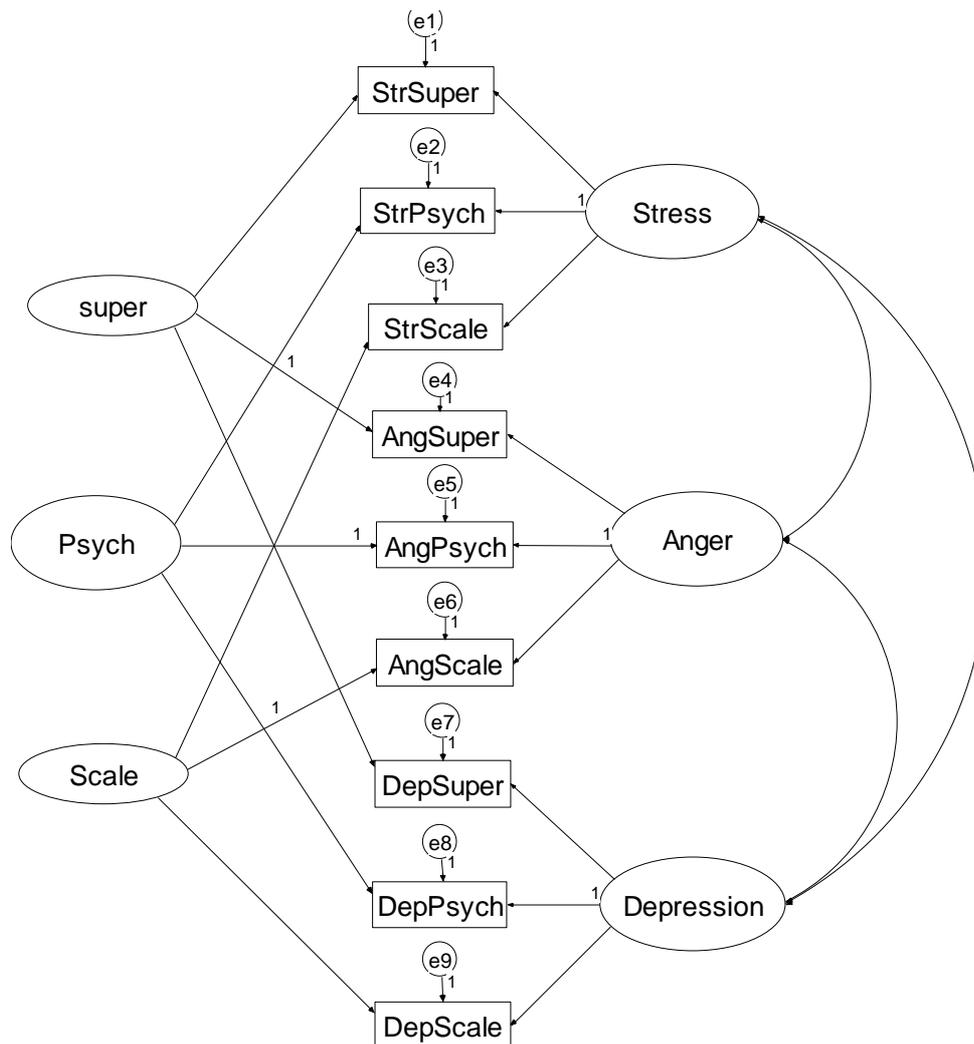


Figure 1 Multitrait-multimethod model reflecting both trait and method factors

Using the Amos (Arbuckle, 1997) structural equation modelling package, it was not possible to obtain convergence with this model, a problem often noted in the literature. Defining starting values that assigned low parameter estimates to trait loadings solved the convergence problem but led to an inadmissible solution. A number of authors have described features of data that can create such difficulties in the analysis of MTMM matrices (e.g., Marsh, 1989; Grayson & Marsh, 1994). These authors acknowledged the complicated structure of MTMM matrices and recommended testing various subsets of the general CFA model depicted in Figure 1. To conserve space, we will not describe the various models fitted to these data. They included the five models suggested by Byrne (1998): 1) correlated traits/correlated methods; 2) no traits/correlated methods; 3) perfectly correlated traits/freely correlated methods; 4) freely correlated traits/uncorrelated methods; 5) the correlated uniqueness model. Suffice it to say that most led to either inadmissible solutions or the solution failed to converge. The former

problem occurred with models that described trait factors, the latter problem occurred when the requirement for trait factors was relaxed and attempts were made to fit correlated uniqueness models (Marsh & Grayson, 1995; Byrne, 1998). A model that was successfully fitted to the data allowed for three correlated methods factors and no trait factors at all, that is, model two in Byrne's list. This model, with resulting parameter estimates, is shown in Figure 2.

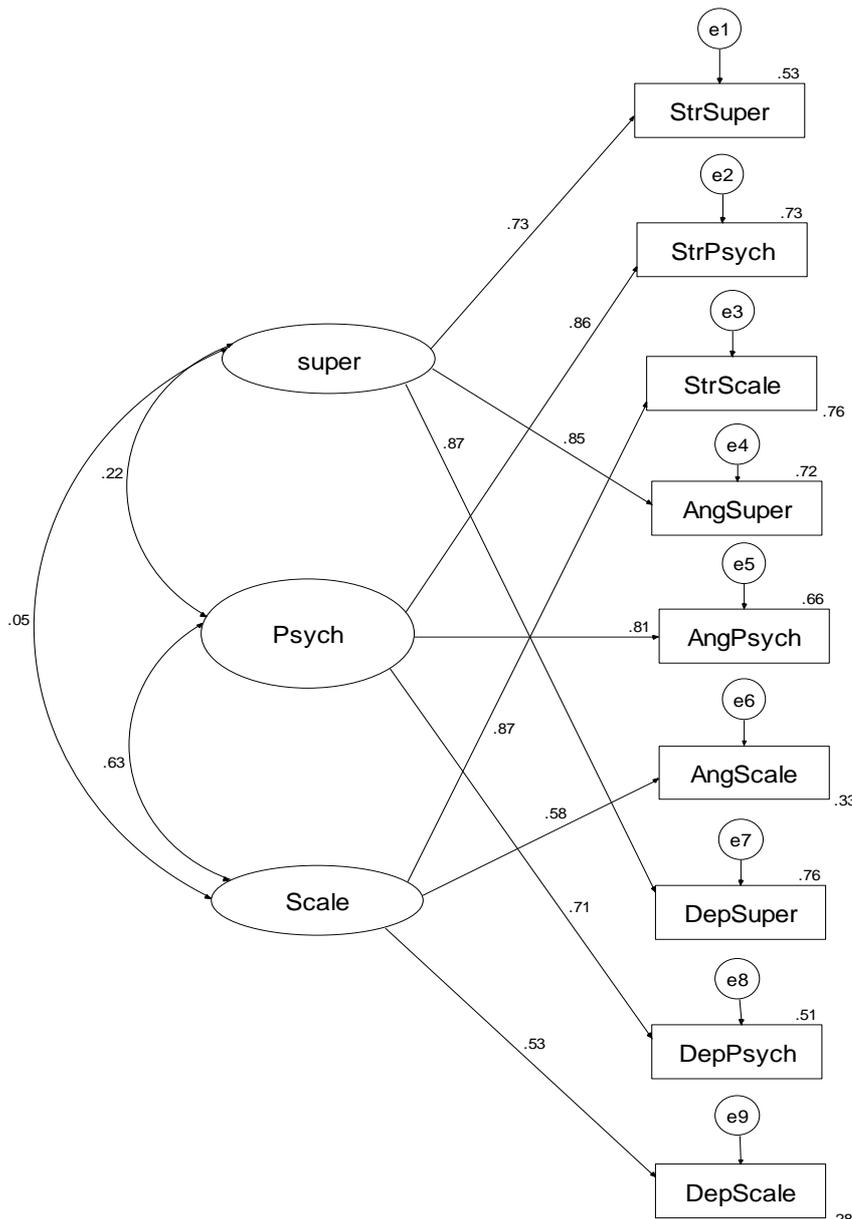


Figure 2

Multitrait-multimethod model reflecting only method factors

For present purposes, the Non-Normed Fit Index (NNFI) recommended by McDonald and Marsh (1990) and the Root Mean Square Error of Approximation (RMSEA) recommended by Browne and Cudeck (1993) were considered as well as the usual  $\chi^2$  measure of goodness of fit. The NNFI varies along a 0-1 continuum in which values greater than .90 are taken to reflect an acceptable fit. Browne and Cudeck suggested that an RMSEA value below .05 indicates a close fit and that values up to .08 are still acceptable. Fit indices for the model shown in Figure 2 were on the outer limits of what might be deemed acceptable. The Chi square value of 59.06 with 24 degrees of freedom had a probability of less than .01, the NNFI value was an acceptable .90, and the RMSEA at .10 was just above the .08 cutoff. With one of the three criteria satisfied, the fit could only be described as marginal.

Marsh and Grayson (1995) cautioned that it can be extremely difficult to overcome the technical difficulties of obtaining convergence and admissible solutions with MTMM matrices and that a heavy emphasis should be placed on substantive interpretations and theoretical framework (see also Byrne, 1998, p. 227). The first stage of data analysis in the present study involved direct inspection of the MTMM and indicated very little evidence of convergent validity. The second stage involved testing various CFA models where the only model that returned an admissible solution that came close to good fit was one that involved method factors and no trait factors. Taken together, these findings suggested that there was little support in the present data for the notion that three separate traits of anger, depression, and stress were being measured by the self-report scales, the supervisor ratings, and the psychologist ratings. Rather, the primary source of variance appears to have been the methods by which the ratings were obtained. Given the difficulty of obtaining satisfactory solutions with CFA, as a final check on the most likely sources of variance in the multitrait-multimethod matrix, exploratory factor analysis was employed. Principal factor analysis employing root one criterion and direct oblimin rotation yielded a three-factor solution that corresponded to the model shown in Figure 2. All three supervisor ratings defined a single method factor, the psychologist ratings defined a second factor, and the self-report ratings defined the third factor. The solution exhibited almost perfect simple structure. The three method factors accounted for 58% of the variance. It was concluded that most of the variance in the multitrait-multimethod matrix could be attributed to method factors.

#### Discussion

This study compared three popular assessment methodologies, referral to a clinician, ratings by a significant other, and self-report in the domains of anger, depression and stress. The overlap between the three methodologies was consistently low. For example, the correlation between self-report and supervisor ratings on the CDI was .19, the same figure obtained by Benavidez and Matson (1993) for people without an intellectual disability (they found a correlation of .48 among people with an intellectual disability). The present study extended these findings by showing that a) lack of convergence is not confined to self-report and informant ratings but also includes clinical interview and b) the lack of convergence among methods applies to stress and anger as well as depression. In addition, this study found some evidence of convergence between self-report and clinical interview ratings that merits further investigation.

The most likely reason for the absence of greater concordance among the results from different assessment approaches is the different perspectives of the respondents. Emotions such as fear, guilt and loneliness, for example, may be difficult for another person to perceive. In the present study, the almost complete lack of convergent validity of the supervisor ratings is probably attributable to the difficulty a third person must always encounter when asked to judge emotions experienced by others. The psychologist in this study was also required to do this but had the benefit of professional training in assessment of emotions and their likely behavioural manifestations. The ratings by the supervisors were also significantly higher than the other scores for both anger and stress, suggesting that supervisors may be inclined to report more pathology than indicated by other approaches. This replicates the finding of Benson & Ivins (1992) who found informants consistently rated anger higher than levels reported by the participants themselves. Further research could profitably explore the extent of this tendency, particularly since so much research in this area relies exclusively on ratings by significant others (Einfeld & Tonge, 1996a, 1996b).

Another reason posited for low levels of overlap between self-report and other methods of assessment among children is their fear of honestly reporting symptoms for which they have previously been criticised or punished (Kazdin & Petti, 1982). The people with an intellectual disability in this study worked and lived under reasonably close supervision of workshop supervisors, parents or houseparents and may well have felt reluctant to report some symptoms for fear of the repercussions of admitting to such feelings. This phenomenon was observed by Baker and Bramston (1997) in their study of anger amongst people with mild intellectual disabilities.

An additional factor impacting on convergence and divergence rates is the likelihood that the presence of an intellectual disability can make it more difficult to notice symptoms of emotional disorders. This masking or diagnostic overshadowing has been widely acknowledged in the research literature (Matson, 1993) and is thought to often blur diagnostic distinctions in this population.

The implication of these findings is that one needs to be very careful about attaching too much significance to ratings obtained from a single source, especially if that source constitutes work supervisors. The supervisors in this study were very experienced in dealing with people with an intellectual disability but it appears that they may not have been able to distinguish among the emotions of anger, depression, and stress. The presence of high correlations across disorders among the supervisor ratings ( $r = .66$ ) suggests that whenever a person is judged to be angry, there is a tendency for the person to also be judged as depressed and stressed. This same tendency is true of the psychologist's ratings ( $r = .62$ ) but here there is overlap between the psychologist's ratings and self-report ratings. The average correlation among the three emotions was lowest ( $r = .43$ ) with the self-report measures suggesting the participants themselves, when responding to inventories, were best able to discriminate between overlapping but conceptually distinct emotional states.

With such a low level of agreement among the different assessment methods, the question naturally arises as to which is the most reliable method of assessment. Our evidence suggests that supervisors tend to make more global assessments where individuals are rated much the same on related constructs, possibly because of the diagnostic overshadowing referred to above or simply because they are too busy to

notice. To a busy observer, the behaviours that characterise anger, depression and stress may appear much the same. The in-depth clinical interviews, recommended by Kazdin and Petti (1982) to circumvent this problem, proved in this study to be more closely correlated with self-report. These interviews were designed to mirror real-life consultations, i.e., they were brief and used a clinician who did not know the participants. Future research could trial the efficacy of a more in-depth interview with a familiar clinician with some history of rapport with the participants.

A further possible solution is the combination of more than one measure and methodology as part of the same assessment, as suggested by Kazdin and Petti (1982). However, the present study shows that combining two or more approaches may not be as helpful as Kazdin and Petti believed. When the elements to be combined are only weakly related, it is difficult to say which ones should have more influence in diagnosis. More research of the kind we have conducted here is needed before we can make judgements about the relative merits of self-report, clinical interview, and informant rating methods of assessment. We conclude by suggesting some forms this research might take.

The present study attempted to examine three methodologies as they are generally applied in current clinical practice. In doing this we may have created a situation where convergence is minimal because of short assessment time, supervisors untrained in assessment, and psychologist not known to the participants. These factors could be manipulated in future studies to determine their impact on convergence. That is, provide some training in assessment to supervisors and allow the clinicians more contact with the participants. We anticipate that convergence will be higher under these conditions. A more obvious improvement to the design of the present study involves the use of multiple clinicians and supervisors so that inter-rater reliability can be gauged. If there is no evidence of inter-rater reliability for a particular technique, then there can be no convergence between this technique and any other assessment method. The laws of reliability prevent convergence under these conditions. In retrospect, a study of this kind should have preceded the multimethod study reported here. In our defence, given the widespread acceptability of informant ratings, we were not expecting the findings that emerged.

It is also important that such studies be conducted on the wider population. Until that is done, we cannot be sure that these findings have wider application. Kline, Lachar, and Gdowski (1992) used a similar design to the one employed here to test for method effects when parents, teachers, and clinicians were asked to assess children on the traits of aggression, depression, and cognitive deficits. Although convergence did occur for the three traits, there was also strong evidence of method effects.

On the basis of our own work, we favour the use of self-report instruments with items and methods of presenting these items designed to be relevant to people with an intellectual disability. We are less certain of what we can say about ratings by work supervisors and psychologists using, as they did in this study, assessment strategies typical of today's busy practitioner.

## References

- Arbuckle, J.L. (1997). Amos user's guide. Chicago: Small Waters.
- Benavidez, D.A., & Matson, J.L. (1993). Assessment of depression in mentally retarded adolescents. Research in Developmental Disabilities, 14, 179-188.
- Benson, B.A., & Ivins, J. (1992). Anger, depression and self-concept in adults with mental retardation. Journal of Intellectual Disability Research, 36, 169-175.
- Baker, W. & Bramston, P. (1997). Attributional and emotional determinants of aggression in people with mild intellectual disabilities. Journal of Intellectual & Developmental Disability, 22, 169-185
- Bramston, P., & Bostock, J. (1994). Measuring perceived stress in people with intellectual disabilities: The development of a new scale. Australia and New Zealand Journal of Developmental Disabilities, 19, 149-157
- Bramston, P. & Fogarty, G. (1995). Measuring stress in the mildly intellectually handicapped: The factorial structure of the Subjective Stress Scale. Research in Developmental Disabilities, 16, 117-131.
- Bramston, P., Fogarty, G., & Cummins, R. A. (1999) The nature of stressors reported by people with an intellectual disability. Journal of Applied Research in Intellectual Disabilities, 12, 1-10.
- Browne, M.W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K.A. Bollen & J.S. Long (Eds.): Testing structural equation models (pp. 136-162). Newbury Park: Sage Publications.
- Byrne, B.M. (1998). Structural equation modelling with LISREL, PRELIS, and SIMPLIS: Basic concepts, applications, and programming. Mahwah, NJ: Lawrence Erlbaum Associates.
- Campbell, D.T., & Fiske, D.W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. Psychological Bulletin, 56, 81-105.
- Einfeld, S.L., & Tonge, B.J. (1996a). Population prevalence of behavioural and emotional disturbance in children and adolescents with mental retardation: I Rationale and methods. Journal of Intellectual Disability Research, 40(2), 91-98.
- Einfeld, S.L., & Tonge, B.J. (1996b). Population prevalence of behavioural and emotional disturbance in children and adolescents with mental retardation: II Epidemiological findings. Journal of Intellectual Disability Research, 40(2), 99-109.
- Finch, A.J., Saylor, C.F., & Nelson, W.M. (1983). The Children's Inventory of Anger: a self-report inventory. Paper presented at the American Psychological Association, Anaheim, California, USA, August 1983.
- Fogarty, G.J., Bramston, P. (1997). Validation of the Lifestress Inventory. Research in Developmental Disabilities, 18, 435-456.
- Grayson, D., & Marsh, H.W. (1994). Identification with deficient rank loading matrices in confirmatory factor analysis: Multitrait-multimethod models. Psychometrika, 59 (1), 121-134.
- Heal, L.W., & Sigelman, C.K. (1995). Response biases in interviews of individuals with limited mental ability. Journal of Intellectual Disability Research, 39, (4), 331-340.

- Iverson, J.C., & Fox, R.A. (1989). Prevalence of psychopathology among mentally retarded adults. Research in Developmental Disabilities, *10*, 77-83.
- Kazdin, A.E., & Petti, T.A. (1982). Self-report and interview measures of childhood and adolescent depression. Journal of Child Psychology and Psychiatry, *23*, (4) 437-457.
- Kline, R.B., Lachar, D., & Gdowski, C.L. (1992). Clinical validity of a Personality Inventory for Children (PIC) typology. Journal of Personality Assessment, *58*, 591-605.
- Kovacs, M. (1981). Rating scales to assess depression in school aged children. Acta Paedopsychiatrica, *46*, 305-315.
- Kovacs, M. (1985). The Children's Depression Inventory (CDI). Psychopharmacology Bulletin, *21*, 995-998.
- Lund, J. (1985). The prevalence of psychiatric morbidity in mentally retarded adults. Acta Psychiatrica Scandinavica, *72*, 563-570.
- Marsh, H.W. (1989). Confirmatory factor analysis of multitrait-multimethod data: Many problems and a few solutions. Applied Psychological Measurement, *13*, 335-361.
- Marsh, H.W., & Grayson, D. (1995). Latent variable models of multitrait-multimethod data. In R.H. Hoyle (Ed.), Structural equation modelling: Concepts, issues, and applications (pp. 177-198). Thousand Oaks, CA: Sage.
- Matson, J.L. (1983). Depression in mentally retarded persons: Research findings and future directions. Australia New Zealand Journal of Developmental Disabilities, *9*(4), 185-190.
- Matson, J.L., Baret, R.P., & Helsel, W.J. (1988). Depression in mentally retarded children. Research in Developmental Disabilities, *9*, 39-46.
- McDonald, R.P., & Marsh, H.W. (1990). Choosing a multivariate model: Noncentrality and goodness-of-fit. Psychological Bulletin, *107*, 247-255.
- Meins, W. (1993). Assessment of depression in mentally retarded adults: reliability and validity of the Children's Depression Inventory (CDI). Research in Developmental Disabilities, *14*, 299-312.
- Reiss, S. (1988). Test manual for the Reiss Screen for Maladaptive Behavior. International Diagnostic Systems Inc., Worthington, OH.
- Reiss, S. & Rojahn, J. (1993). Joint occurrence of depression and aggression in children and adults with mental retardation. Journal of Intellectual Disability Research, *37*, 287-294.
- Saylor, C.F., Finch, A.J., Spirito, A., & Bennett, B. (1984). The Children's Depression Inventory: A systematic evaluation of psychometric properties. Journal of Consulting and Clinical Psychology, *52*, 955-967.
- Spielberger, C.D. (1988). State-Trait Anger Expression Inventory. Odessa, Florida: Psychological Assessment Resources.
- Zetlin, A.G. (1993). Everyday stressors in the lives of Anglo Hispanic learning handicapped individuals. Journal of Youth and Adolescence, *22*, 327-335.