"Running Head:" Construct Validity of MAS

The Motivation of Stereotypic and Repetitive Behaviour: Examination of Construct Validity of the Motivation Assessment Scale

Annette V. Joosten and Anita C. Bundy

University of Sydney

Abstract

Construct validity of the Motivation Assessment Scale (Durand & Crimmins, 1988: 47) was re-examined using Rasch analysis on data from 67 participants (246 MASs). Participants were children with a dual diagnosis of autism and intellectual disability and children with only intellectual disability. The results failed to provide adequate support for construct validity of the MAS as a unidimensional scale or for the originally proposed four factor structure. Point biserial correlations for two sensory motivator items were negatively correlated indicating the scale is not unidimensional. Data from the same two sensory items plus another sensory item failed to fit the measurement model; and person fit of 91.5% was below the desired 95%. Principal component analysis supported a second dimension. The differential item analysis (DIF) indicated that diagnosis may influence hierarchy of item difficulty.

KEY WORDS: motivation, stereotypic, repetitive, construct validity, Rasch analysis

INTRODUCTION

Stereotypical or repetitive behaviours are reported in 37-95% of individuals with autism and/or intellectual disability (Filipek et al., 1999: 472) although they are generally reported as occurring less frequently or less intensely in children with intellectual disability alone than in those children who also have autism (Bodfish, Symons, Parker, & Lewis, 2000: 237; King & Lynn, 1998: 523; Nijhof, Joha, & Pekelharing, 1998: 3).

Stereotypical behaviours include simple body movements such as rocking and flapping arms or more complex rhythmical repetitive patterns of movement. Some stereotypical behaviours are object oriented or sensory focused, for example repetitive vocalisations.

Stereotypical behavior is not compatible with learning new skills; hinders communication; and reduces interaction with the environment (Bright, Bittick, & Fleeman, 1981: 167; Koegel & Covert, 1972: 381; Nijhof et al., 1998; Storey, Bates, McGhee, & Dycus, 1984: 510). The presence of stereotypic behaviour is reported to increase the likelihood of living in a more restricted environment and to stigmatise people who engage in such behaviour (Bonadonna, 1981: 12; Durand & Carr, 1987: 119). The more restricted environment the greater the level of self-stimulatory behaviour observed (Bright et al., 1981: 167).

Over time there has been increasing awareness that effective intervention programs depend on knowledge of what motivates these behaviours (Horner, Carr, Strain, Todd, & Reed, 2002: 425; Spreat & Connelly, 1996: 528). The more precise the assessment the better outcomes from intervention (Didden, Duker, & Korzilius, 1997: 388).

Motivation is frequently assessed indirectly by a third party (e.g., parent, teacher) who knows the child well enough to respond to questions about his or her behaviour.

Indirect assessment often is completed through interviews with a structured set of items eliciting information about the behaviour and setting (Sigafoos, Kerr, & Roberts, 1994: 334). The answers to these questions allow interventionists to hypothesise about antecedents and consequences of the behaviour (Schreibman, 1994: 14)

Indirect assessment using an interview or rating scale is designed to be an easy-to-administer, efficient and practical method of assessment (Sigafoos et al., 1994: 334). The most widely used indirect functional assessment is the Motivation Assessment Scale developed by Durand and Crimmins (1988) (Howlin, 1998a : 311; Paclawskyj, Matson, Rush, Smalls, & Vollmer, 2001: 485) .

While the simplicity of indirect assessment is appealing, Floyd, Phaneuf and Wilczynski (2005) raised some concern about the validity of instruments that rely on raters forming judgements based on retrospective accounts. They were also concerned that no guidelines exist to evaluate the quality of the results obtained (Floyd et al., 2005: 59). Although several studies have investigated the reliability of the MAS only a small number have investigated its validity (Floyd et al., 2005: 63). Researchers who have used statistical means for assessing validity have concluded that further investigation of the psychometric properties was required (Emerson & Bromley, 1995: 396; Floyd et al., 2005: 66; Howlin, 1998a: 311).

The original 4- factor structure of the MAS was based on the assumption that the MAS items could be grouped into four categories of motivation. Validity of the MAS has largely been accepted because of its good face validity. Validity of the MAS also has been examined statistically by its authors by comparing teacher's ratings on the scale with the observations in a number of analogue assessment conditions. The results provided evidence for the validity the MAS (Durand & Crimmins, 1988: 48). However, this research involved very small sample sizes.

Three studies (Bihm, Kienlen, Ness, & Poindexter, 1991: 1235; Duker & Sigafoos, 1998: 131; Singh et al., 1993: 65) have been conducted to investigate the factor structure of the MAS to determine if the 16 items did indeed form four meaningful factors that corresponded to the four motivators proposed by Durand and Crimmins (1988: 47).They also investigated which items loaded highly on each factor structure (Singh et al., 1993: 65). The results have varied.

While for the most part Bihm et al., (1991) and Singh et al., (1993) (with institution dwelling participants) confirmed Durand and Crimmins' (1988) findings, both Duker and Sigafoos (1998) and Singh et al., obtained different results. Duker and Sigafoos reported that the factor structure in their study of 86 participants was ambiguous; it differed from that proposed by Durand and Crimmins and by Bihm et al., possibly due to different behaviours that were assessed in each study. They hypothesised that differences in topography or types of behaviours may have resulted in difficulty establishing validity.

Similarly, in their school sample of 96 participants Singh et al., found neither 3-, 4-, nor 5- factor structures were meaningful. They proposed that this outcome was related to behaviour frequency, which was much less in their school sample. They also speculated that the level of intellectual disability in their school sample may have been higher than that in their institution-dwelling sample. However, information on this was not available.

The findings of Durand and Crimmins (1988) and Singh et al.,(1993) suggest that motivation for stereotyped behaviour may differ for different people. While many factors may contribute, the effect of diagnosis may be great enough that it precludes establishing a factor structure that applies to all groups. That is, certain motivations may be more common in children with certain diagnosis. For example, children with autism may be more likely to engage in sensory motivated behaviours than children with intellectual

disability but further investigation into the motivation of these behaviours is needed (Vig & Jedrysek, 1999: 235).

An alternative to the classical means for examining construct validity with factor analysis is through the use of Rasch analysis where evidence for construct validity suggests that the items on a scale measure a single latent trait (Bond & Fox, 2001: xx) where easy items are easy for all people and more capable people are more apt to get high scores on difficult items. When a test has evidence for construct validity, items (and people) are ordered logically along the hierarchy according to the extent to which they probe (or manifest) the underlying trait. To examine unidimensionality, Rasch provides goodness of fit statistics that reveal the degree to which the people and items respond in the manner expected by the model (Wright & Stone, 1979: 66). Items that fail to fit the model are those with unacceptably high fit statistics indicating that they diverge too much from the pattern expected by the analysis (Bond & Fox, 2001: 26).

The purpose of this study was to use Rasch analysis to investigate construct validity of the MAS (i.e. the extent to which the MAS is a unidimensional measurement of motivation of stereotypic behaviours). If MAS items were not found to reflect a unidimensional scale, then we would use principal components analysis to revisit the original 4-factors. Failing confirmation of the 4-factor structure, we planned to investigate the influence of diagnosis on the motivation for stereotyped behaviours. That is, was there a tendency for the motivations for stereotyped behaviours to differ according to whether or not children with intellectual disability also had autism?

Current Study

We sought to answer the questions:

(1) Do the MAS items positively correlate to the overall test? (2) Do data from at least 95% of items and children fit the Rasch model? (3) How closely does the empirical

variance match the modelled variance? (4) Is there any evidence that items on the MAS differ significantly for children with autism and intellectual disability compared to children with intellectual disability only?

METHOD

Participants

This study included 67 participants aged 5-18 years, who attend a Day Specialist School in a moderate sized suburban Australian town. Enrolment at the school requires evidence of intellectual disability based on the results of two psychometric tests administered by a psychologist (e.g., The Weschler Intelligence Scale for Children (WISC-111/1V) and the Vineland Adaptive Scale). A dual diagnosis of autism and intellectual disability is established by the psychologist (or an Autism Assessment and Diagnostic Team) if the children also meet the criteria for Autistic Disorder on the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV) 1994. All children live at home with parents or carers and attend school on a daily basis. Table 1 presents specific information about the gender and diagnosis of the participants.

[place Table 1 about here]

Instrument

The MAS (Durand & Crimmins, 1988) is a 16-item questionnaire that takes approximately 10 minutes to complete. The items are grouped equally into subscales that reflect four sources of motivation to: (a) gain social attention, both negative and positive; (b) gain access to tangible objects or in response to the withdrawal of tangible reinforcers; (c) escape from or to avoid people or activities and (d) experience sensory feedback or

stimulation. If a behaviour occurs in more than one context or a child has a range of stereotyped or repetitive behaviours, multiple MAS forms are completed. This enables examination of the possibility that children have multiple motivators for stereotypical behaviours. Each item is rated on a 7-point rating scale with responses ranging from 'never (0)' to 'always (06)'.

Procedure

Teachers who knew the children well (i.e., had taught the children for at least 1 school year) identified children who engaged in stereotypic or repetitive behaviours in the classroom and the playground at school. Over a 6-week period, the teachers and the principal investigator completed one MAS form for each identified behaviour for each child. As most children engaged in more than one repetitive behaviour, a total of 246 MAS scales were completed. Children with ASD and intellectual disability accounted for 132 observations and children with intellectual disability only for114.

Data Analysis

Winsteps 3.58 (Linacre, 2005) was used to analyse the data. Prior to any other analyses, point biserial correlations were calculated to determine whether they were all positively oriented.

Goodness of fit statistics were then calculated to determine the degree of the data from the children and the items along the unidimensional construct (i.e., does the scale work similarly for all children and do data from all items fit the expected model?). Two pairs of fit statistics (infit and outfit) are reported in the form of both mean square and standardised values. Infit statistics describe the persons' interactions with items around their ability level and give more insight into the items performance than do outfit statistics, which are sensitive to outlying scores (Bond & Fox, 2001: 176). Fit statistics that are too large (MnSq >1.4; t > 2) indicate more variation than modelled, (i.e. the response string is

more haphazard than expected) (Bond & Fox, 2001: 177) are of particular concern. Data from items or people that failed to fit because they are overly variable or erratic scores require further investigation.

A principal components analysis was also generated by Winsteps 3.58 (Linacre, 2005), allowed further examination of unidimensionality of the MAS. The evidence would suggest that MAS is unidimensional items if the empirical variance closely matches the modeled variance and if the unexplained variance from the first factor is less than 3 Eigenvalue units (Linacre, 2005).

Finally, a differential analysis (DIF) also generated by Winsteps 3.58 (Linacre, 2005), enabled examination of possible contrasting item-by-item profiles. In this case, we contrasted the data for children with dual diagnosis of autism and intellectual disability with data from children with intellectual disability.

RESULTS

Point biserial correlations for two sensory items, S04 and S03, were negatively correlated with the overall test suggesting that the MAS is not a unidimensional scale and that sensory motivators form a different dimension to the other items. Thus, it was not surprising that goodness of fit statistics for those items plus one more sensory item (S02) were outside the acceptable range (MnSq >1.4; t > 2). Thus, data from only 81.5% of items conformed to the measurement model, considerably less than the desired 95%.

Similarly, fit statistics for data from participants were unexpectedly high for 27 of the 246 observations responses. Of the 27 observations with unexpectedly high fit statistics 18, were from children with the dual diagnosis of autism and intellectual disability.

Principal components analysis confirmed further evidence that the MAS does not reflect a unidimensional construct. While the empirical variance of 52.5% closely matched

the modelled variance of 54.8%, the unexplained variance from the first factor was 5.8 Eigenvalue units, far greater than the desired value of <3 units.

The first factor reflected a subset of observations with *unexpectedly high* scores on sensory motivators and one attention motivator worded like a sensory motivator and *unexpectedly low* scores on motivators: to obtain a tangible object, to escape or to gain attention. Thirteen of these 20 observations were from children with dual diagnosis of autism and intellectual disability. These observations were in contrast with others that had the opposite profile (low scores on sensory and one attention item but high scores for obtaining objects, escaping, or gaining attention. Twelve of the 23 children on whom these observations were based were children with intellectual disability only.

The differential analysis (DIF) conducted to contrast different response profiles according to diagnostic group revealed that items S03 and E04 were significantly higher for children with autism and items E01, T01 and T04 for children with intellectual disability, as shown in Table 2 and Figure 1.

[Place Table 2 and then Figure 1 about here]

DISCUSSION

The purpose of this investigation was to re-examine evidence for construct validity of the MAS. Specifically, we questioned the 4-factor structure because of the inconsistency of the results of earlier studies (Bihm et al., 1991; Duker & Sigafoos, 1998; Singh et al., 1993). Using Rasch analysis, we tested the hypothesis that motivation for stereotyped behaviour is a unidimensional construct. Since we could not establish a single construct, we revisited the original 4-factor solution and failed to find support for that either. Thus, we investigated the influence of diagnosis.

The results of our DIF analysis showed that the stereotyped behaviours of children with autism were more apt to stem from sensory and escape and less apt to come from other motivators (gaining attention or tangible objects). These findings may suggest that the hierarchy of difficulty may differ for children who have autism and intellectual disability versus for difficulty for children with only intellectual disability.

Our results may help to explain the findings of other researchers who have found differences as a result of group. Singh et al.,(1993) for example, supported Durand and Crimmins' original 4-factor structure with institution-dwelling subjects but not with school-based subjects.

There is good reason to suspect that the motivations for stereotyped behaviour in children with autism (even if they also have intellectual disabilities) might differ from those of children with intellectual disability but no autism. Children with autism often are thought to have abnormal central nervous system processing, which in turn contributes to either over-or- under arousal. Various authors (Guess & Carr, 1991: 300; Schneck, 2000: 142) have proposed that repetitive and stereotyped behaviours provide a calming influence to an over-aroused nervous system and have an alerting affect on an under-aroused system. Thus, a powerful motivator for children with autism engaging in repetitive behaviour may be that it makes them feel good (or better). Over time, some children with autism may learn that there are other benefits to stereotyped behaviours. However, because social impairments are characteristic of the disorder, children with autism are unlikely to engage in stereotyped behaviours to gain the attention of others.

In contrast, children with intellectual disability but not autism enjoy attention and may learn early on that repetitive behaviours bring others to them. Thus, they may use repetitive behaviours primarily to gain social benefits. They do not experience the same reactions to sensation and are unlikely to use repetitive behaviours as a means to become

calmer or more alert. Children with intellectual impairment but no evidence of autism also may use stereotyped behaviours for other rewards (e.g., to obtain objects or to escape from demands imposed by others).

CONCLUSIONS

Our findings provide some evidence that diagnosis may be a factor in the failure of previous researchers to establish construct validity of the MAS because it affects motivation for stereotypic behaviours. However, it seems unlikely that differential diagnosis is the only reason for failure of researchers to confirm the 4-factor model for the MAS proposed originally by (1988). Other factors, such as the context in which the behaviours occur should also be investigated in further studies.

REFERENCES

American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders (4 th edition). Washington, DC.*

Bihm, E. M., Kienlen, T. L., Ness, M. E., & Poindexter, A. R. (1991). Factor structure of the Motivation Assessment Scale for persons with mental retardation. *Psychological Reports, 68*, 1235-1238.

Bodfish, J. W., Symons, F. J., Parker, D. E., & Lewis, M. H. (2000). Varieties of repetitive behavior in autism: Comparisons to mental retardation. *Journal of Autism & Developmental Disorders, 30*(3), 237-243.

Bonadonna, P. (1981). Effects of a vestibular stimulation program on stereotypic rocking behaviour. *The American Journal of Occupational Therapy, 35*(12), 775-781.

Bond, T. G., & Fox, C. M. (2001). *Applying the Rasch Model: Fundamental Measurement in the Human Sciences*. New Jersey: Lawrence Erbaulm Associates.

Bright, T., Bittick, K., & Fleeman, B. (1981). Reduction of self-injurious behaviour using sensory integration techniques. *The American Journal of Occupational Therapy, 35*, 167-172.

Didden, R., Duker, P. C., & Korzilius, H. (1997). Meta-analytic study on treatment effectiveness for problem behaviours with individuals who have mental retardation. *American Journal of Mental Retardation, 101*, 387-399.

Duker, P. C., & Sigafoos, J. (1998). The Motivation Assessment Scale: Reliability and construct validity across three typographies of behaviour. *Research in Developmental Disabilities, 19*(2), 131-141.

Durand, V. M., & Carr, E. G. (1987). Social Influences on "self-stimulatory" behavior: Analysis and treatment applications. *Journal of Applied Behavior Analysis,, 20*(2), 119-132.

Durand, V. M., & Crimmins, D. B. (1988). *The Motivation Assessment Scale*. Topeka, KS: Monaco & Associates.

Emerson, E., & Bromley, J. (1995). The form and Function of challenging behaviours. *Journal of Intellectual Disability Research, 39*(5), 388-398.

Filipek, P., Accardo, P. J., Baranek, G. T., Cook, E. H. J., Dawson, G., Gordon, B., et al. (1999). The screening and diagnosis of autistic spectrum disorders. *Journal of Autism & Developmental Disorders, 29*(6), 439-476.

Floyd, R. G., Phaneuf, R. L., & Wilczynski, S. M. (2005). Measurement properties of indirect assessment methods of functional behavioural assessment: a review of research. *School Psychology Review, 34*(1).

Guess, D., & Carr, E. G. (1991). Emergence and maintenance of stereotypy and self-injury. *American Journal on Mental Retardation, 96*(3), 299-319.

Horner, R. H., Carr, E. G., Strain, P. S., Todd, A. W., & Reed, H. K. (2002). Problem behavior interventions for young children with autism: A research synthesis. *Journal of Autism & Developmental Disorders, 32*(5), 223-446.

Howlin, P. (1998a). Practitioner Review: Psychological and Educational treatments for Autism. *Journal Child Psychology and Psychiatry, 39*(3), 307-322.

King, B. H., & Lynn, D. (1998). self-injurious behavior in mental retardation. *Current opinion in Psychiatry, 11*, 523-526.

Koegel, R. L., & Covert, A. (1972). The relationship of self-stimulation to learning in autistic children. *Journal of Applied Behavior Analysis,, 5*, 381-387.

Linacre, J. M. (2005). A User's guide to Winsteps Ministeps: Rasch-model computer programs. Chicago.

Nijhof, G., Joha, D., & Pekelharing, H. (1998). Aspects of stereotypic behavior among autistic persons: A study of the literature. *The British Journal of Developmental Disabilities, 44,1*(86), 3-13.

Paclawskyj, T. R., Matson, J. L., Rush, K. S., Smalls, Y., & Vollmer, T. R. (2001). Assessment of the convergent validity of the Questions about Behavioral Function Scale with anoalgue functional analysis and the Motivation Assessment Scale. *Journal of International Disability Research, 45*(6), 484-494.

Schneck, C. M. (2000). The Efficacy of a Sensorimotor Treatment Approach by Occupational Therapists. In R. A. Huebner (Ed.), *Autism: A Sensorimotor Approach to Management* (pp. 139-178). Austin: PRO-ED INC.

Schreibman, L. (1994). General Principals of Behavior Management. In E. Schopler & G. B. Mesibov (Eds.), *Behavioral Issues in Autism* (pp. 11-39). New York: Plenum Press.

Sigafoos, J., Kerr, M., & Roberts, D. (1994). Interrater reliability of the Motivation Assessment Scale:Failure to replicate with aggressive behavior. *Research in Developmental Disabilities, 15*(5), 333-342.

Singh, N. N., Donatelli, L. S., Best, A., Williams, D. E., Barrera, F. J., Lenz, M. W., et al. (1993). Factor Structure of the Motivation Assessment Scale. *Journal of Intellectual Disability Research, 37*, 65-74.

Spreat, S., & Connelly, L. (1996). Reliabilty analysis of the motivation assessment scale. *American Journal of Mental Retardation, 100*(5), 528-532.

Storey, K., Bates, P., McGhee, N., & Dycus, S. (1984). Reducing the self-stimulatory behavior of a profoundly retarded female through sensory awareness training. *The American Journal of Occupational Therapy, 38*(8), 510-516.

Vig, S., & Jedrysek, E. (1999). Autistic features in young children with significant cognitive impairment:Autism or mental retardation? *Journal of Autism & Developmental Disorders, 29*(3), 235-248.

Wright, B. D., & Stone, M. H. (1979). *Best Test Design*. Chicago: Mesa Press.

Table 1. Participants

Gender	Number	Diagnosis
Female	9	ID
Female	5	ASD+ID
Male	29	ID
Male	24	ASD+ID
Total	67	

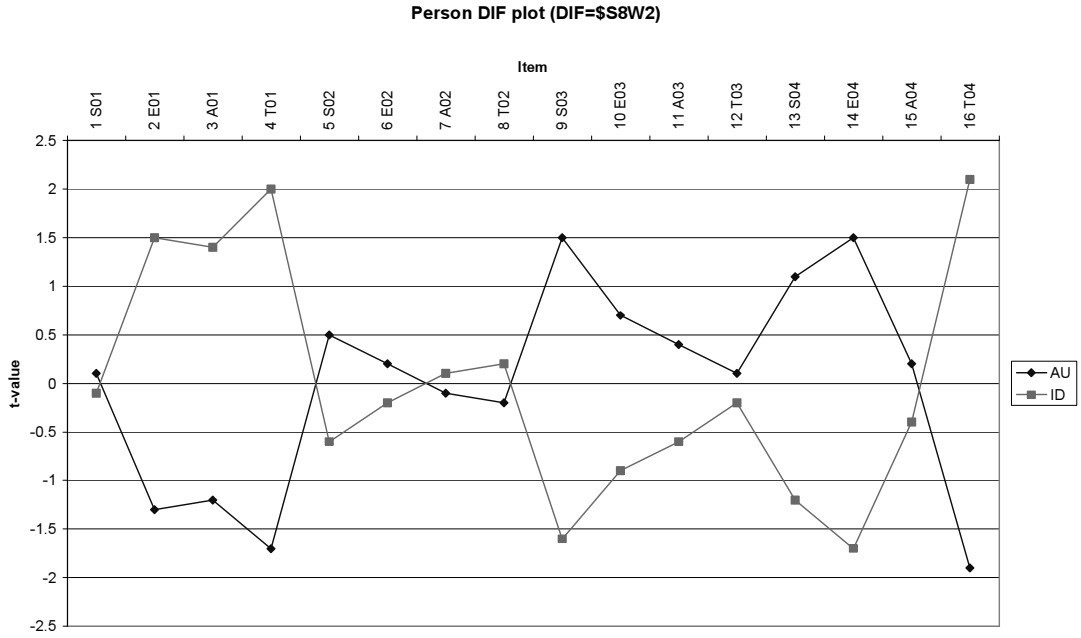
ASD+ID – Autism spectrum disorder and intellectual disability
ID – Intellectual disability

Table 2. Differential Item Analysis (DIF)

Item	ASD + ID dif measure	ID dif Measure	Probability
S (sensory) 03	42.2	40.0	.03
E (escape) 04	51.4	49.6	.03
E (escape) 01	49.0	50.6	.05
T (tangible) 01	51.6	53.9	.01
T (tangible) 04	50.2	52.6	.005

Figure Caption

Figure 1. Differential Item Analysis (DIF)



Correspondence concerning this article should be addressed to:

Annette V. Joosten, Faculty of Health Sciences, University of Sydney

Anita C. Bundy, Faculty of Health Sciences, University of Sydney