Research in Developmental Disabilities 34 (2013) 1804–1814

Contents lists available at SciVerse ScienceDirect

Research in Developmental Disabilities

Validity and reliability of the *Behavior Problems Inventory*, the *Aberrant Behavior Checklist*, and the *Repetitive Behavior Scale – Revised* among infants and toddlers at risk for intellectual or developmental disabilities: A multi-method assessment approach☆

Johannes Rojahn [a,*], Stephen R. Schroeder [b], Liliana Mayo-Ortega [b,c], Rosao Oyama-Ganiko [c], Judith LeBlanc [b,c], Janet Marquis [b], Elizabeth Berke [a]

[a] George Mason University, United States
[b] University of Kansas, United States
[c] Centro Ann Sullivan del Peru, Peru

ARTICLE INFO

Article history:
Received 17 December 2012
Received in revised form 18 February 2013
Accepted 22 February 2013
Available online 16 March 2013

Keywords:
Aberrant Behavior Checklist
Behavior Problems Inventory
Repetitive Behavior Scale – Revised
Psychometric properties
Reliability, Validity
Multitrait–multimethod
Neurodevelopmental disorders
At risk
Infants
Toddlers
Children

ABSTRACT

Reliable and valid assessment of aberrant behaviors is essential in empirically verifying prevention and intervention for individuals with intellectual or developmental disabilities (IDD). Few instruments exist which assess behavior problems in infants. The current longitudinal study examined the performance of three behavior-rating scales for individuals with IDD that have been proven psychometrically sound in older populations: the *Aberrant Behavior Checklist (ABC)*, the *Behavior Problems Inventory (BPI-01)*, and the *Repetitive Behavior Scale – Revised (RBS-R)*. Data were analyzed for 180 between six and 36 months old children at risk for IDD. Internal consistency (Cronbach's α) across the subscales of the three instruments was variable. Test–retest reliability of the three BPI-01 subscales ranged from .68 to .77 for frequency ratings and from .65 to .80 for severity ratings (intraclass correlation coefficients). Using a multitrait–multimethod matrix approach high levels of convergent and discriminant validity across the three instruments was found. As anticipated, there was considerable overlap in the information produced by the three instruments; however, each behavior-rating instrument also contributed unique information. Our findings support using all three scales in conjunction if possible.

Individuals with intellectual or developmental disabilities (IDD) are at a heightened risk for developing chronic and severe behavior problems during the course of their lives. Reliable and valid assessment of such behaviors is an important element in empirically verifying successful prevention and intervention.

In the past decade, there has been an increased interest in early identification and preventive intervention of behavior problems, including aggression, self-injurious behavior (SIB), and stereotyped behavior, among infants and toddlers at risk for IDD (Sigafoos, Lancioni, Didden, & O'Reilly, in press; Schroeder & Courtemanche, 2012). Yet, there are only a few existing

☆ This study is part of an ongoing international collaborative research project between the University of Kansas, Centro Ann Sullivan del Peru, and George Mason University which was funded by the Fogarty International Research grant no. HD 060500. The authors wish to gratefully acknowledge the effort of the participating families and the pro bono work of many professions in Peru.

* Corresponding author at: Department of Psychology, George Mason University, 10340 Democracy Lane, Suite 202, Fairfax, VA 22033, United States. Tel.: +1 703 993 4241.
E-mail address: jrojahn@gmu.edu (J. Rojahn).

0891-4222/$ – see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.ridd.2013.02.024

assessment instruments that were specifically developed and validated for behavior problems in children below the age of three, such as the recently published *Baby and Infant Screen for Children with Autism Traits, Part 3* (BISCUIT-3) (Matson et al., 2009). There are, however, several well-validated behavior-rating scales that were originally developed for and validated in adult populations with IDD which have also been used successfully among younger populations.

The current study is a psychometric analysis involving three behavior-rating scales with a population of infants and toddlers in Peru. In this study, we analyzed several psychometric areas of the BPI-01 using these challenging behavior instruments. Firstly, we examined the test–test reliability of the BPI-01. Secondly, using a multitrait–multimethod approach (Campbell & Fiske, 1959), we examined the confirmatory and discriminant validity of the BPI-01, in comparison to the ABC and the RBS-R. Campbell and Fiske (1959) described the multitrait–multimethod approach as a tool to examine convergent and discriminant validity in order to establish overall validity of an empirical instrument. This approach involves presenting a matrix that includes all possible intercorrelations when measuring specific traits through several methods. Lastly, we examined the sensitivity and specificity of the BPI-01 in comparison to the RBS-R and the ABC.

1. Method

1.1. Participants and recruitment

Participants were recruited by the Centro Ann Sullivan del Peru (CASP) via radio, television, and newspaper announcements in Lima, Peru and the rest of the country for a family-based early intervention program. The advertisements solicited children between 6 and 36 months old showing signs of aberrant behaviors. Around 1000 families called CASP in response to the advertisements. Parents where then connected with a trained CASP Triage and Information Coordinator who further explained the study and inclusion criteria. At this point, parents decided whether their child was appropriate for the study and 341 families agreed to visit CASP with their child for a 15–30 min screening interview.

Children were chosen for clinical trials based on their signs of SIB, aggression, and stereotyped behavior and their parents' answers on the *Parental Concerns Questionnaire* (PCQ). The PCQ is a 15-item yes/no questionnaire based upon risk factors of aberrant behaviors that were developed by our study (Mayo-Ortega et al., 2012). Trained veteran parents administered the PCQ. Each of these parents had had a child enrolled at CASP and had received around 185 h of training per year for several years. We feel that this is particular strength of the screening process because the veteran parents could empathize with the families and encourage them. Participants had to meet several inclusion criteria gleaned from the research literature (Dawson, 1996; Dunlap et al., 2006; Rojahn, Schroeder, & Hoch, 2008) that would classify them as at-risk for behavior problems in IDD (Mayo-Ortega et al., 2012). The inclusion criteria involved genetic disorders associated with IDD (i.e., Down, Smith-Magenis, Prader-Willi, Rett, and Fragile-X syndrome), family history of brain disorders, common comorbid medical conditions associated with IDD (i.e., congenital rubella, tuberous sclerosis, brain trauma, and stroke), pre-and perinatal disorders causing serve or profound ID, psychiatric factors (i.e., family history of autism, mood disorders, compulsive disorders, anxiety disorders, and BPD), neurochemical metabolic factors (i.e., Hyperserotonemia, dopamine depletion in the nigrostriatal pathways of the basal ganglia, dysfunctional opioid peptide system, elevated ammonia or lactic acid levels, organic aciduria, fatty acid disorder, aminoacidopathy, and mitochondrial disorder), neuropsychological factors (i.e., lack of orientation to social stimuli, joint attention, motor coordination, position, and/or motor imitation, and presence of stereotyped behavior), and communication deficits in either receptive or expressive communication (Mayo-Ortega et al., 2012).

Of these 341 families that responded and visited the CASP facility for a screening interview, 262 met the inclusion criteria, and 234 were administered the first round of in-depth interdisciplinary assessment measures. This first round involved developmental pediatric exams with consultations in neurology, genetics, and nutrition, questions about vision, hearing and dental health, cognitive (Bayley, 2006) and communication (Wetherby & Prizant, 2002) assessments, and an autism screening with the *Child Autism Rating Scale* (CARS; Schopler, Reichler, & Renner, 1988). The CARS was part of the interdisciplinary evaluation after screening and their scores did not affect inclusion in the study. Of the 180 participants who completed the study, 74 had CARS scores of 15–53 ($M = 35$). These children were then reassessed six months later and one year later.

In the end, 180 participants completed the BPI-01 (Rojahn, Matson, Lott, Esbensen, & Smalls, 2001) for all three time points and were included in this study. Participants were excluded because they either failed to meet inclusion criteria or did not return to CASP to complete subsequent rounds of data collection. These participants consisted of 110 boys and 68 girls (2 missing data), ranging in age from 4 to 48 months ($M = 27.4$; $SD = 10.1$).

None of the children in this study had a formal diagnosis of an Autism Spectrum Disorder (ASD) but they were all identified as at risk for ASD. Of the 180 to complete the study, 21 were un-testable on the Bayley (2006) due to severity of their disability or noncompliance. A total of 100 children were more than one standard deviation below the mean and 59 had scores that were below average but within one SD below the mean. We were hesitant to label any of these children as having an IDD because they were so young. Instead they were classified as at risk for IDD.

1.2. Measures of behavior problems

The key behavioral assessments utilized in this study were the ABC, the BPI-01, and the RBS-R.