Research in Autism Spectrum Disorders 5 (2011) 1430–1440

Contents lists available at ScienceDirect

Research in Autism Spectrum Disorders

Journal homepage: http://ees.elsevier.com/RASD/default.asp

Assessing implementation of comprehensive treatment models for young children with ASD: Reliability and validity of two measures

Kara Hume [a,*], Brian Boyd [b], Matt McBee [a], Drew Coman [c], Anibal Gutierrez [c], Evelyn Shaw [a], Laurie Sperry [d], Michael Alessandri [c], Samuel Odom [a]

[a] Frank Porter Graham Child Development Institute, University of North Carolina, Chapel Hill, 517 South Greensboro Street, Carrboro, NC 27510, USA
[b] Department of Allied Health Sciences, University of North Carolina, Chapel Hill, UNC-CH Bondurant Hall, Chapel Hill, NC 27599-7120, USA
[c] Department of Psychology, University of Miami, P.O. Box 248185, Coral Gables, FL 33124-0751, USA
[d] School of Education and Human Development, University of Colorado, Denver, P.O. Box 173364, Campus Box 106, Denver, CO 80217-3364, USA

ARTICLE INFO

Article history:
Received 5 October 2010
Accepted 3 February 2011
Available online 31 March 2011

Keywords:
Autism spectrum disorder
Comprehensive treatment models
Treatment implementation
Treatment integrity

ABSTRACT

Treatment implementation is an under-studied and under-reported aspect of intervention studies involving individuals with autism spectrum disorder (ASD). One primary area of concern is the lack of reliable and valid implementation measures, which allows a conclusive association to be drawn between the intervention and participant outcomes. This study examined the psychometric properties of two implementation measures developed for comprehensive treatment models serving preschoolers with ASD (i.e., LEAP and TEACCH). Both of the measures were completed in classrooms using LEAP or TEACCH instructional approaches as well as in classrooms in which a business-as-usual or non-model specific treatment approach was used. Across four months of one school year, a maximum of 4 observations were conducted in each of the 34 classrooms involved in the study. Results indicated that both implementation tools are reliable and valid, and that particular subscales of these measures allowed for discrimination of the three types of classrooms from each other. This step of psychometrically validating implementation measures as part of conducting efficacy studies may yield more robust associations between implementation and intervention effects.

1. Introduction

To promote positive developmental outcomes for children with autism spectrum disorders (ASD), over the past decades researchers have developed comprehensive treatment models of service (National Research Council, 2001). A prominent question in the field has been about the relative efficacy of these models (Odom, Boyd, Hall, & Hume, 2010). Any examination of efficacy, however, is built on the presumption that models are well implemented (Durlak, 2010), which, in turn, requires that implementation, be systematically assessed. Systematic assessments of implementation in scientific studies of efficacy must be reliable and valid, and to date, there are been few reports of the psychometric characteristics of implementation measures designed for programs for students with ASD. The purpose of this study was to analyze the reliability and validity of implementation measures for two prominent CTMs for children with ASD. Implementation was defined in the current study as "*the extent to which the critical components of an intended program are present when that program is enacted*" (Century, Rudnick, & Freeman, 2010).

* Corresponding author. Tel.: +1 919 843 2291; fax: +1 919 966 1786.
E-mail address: kara.hume@unc.edu (K. Hume).

1750-9467/$ – see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.rasd.2011.02.002

In their examination of educational practices for children with autism, a committee convened by the National Academy of Sciences identified a set of comprehensive models of intervention, which they defined as a set of practices designed to achieve a broad learning or developmental impact on the core deficits of ASD (National Research Council, 2001). Further, they occurred over an extended period of time (e.g., one year or multiple years), are intense in their application (e.g., 25 h or more per week), usually have multiple components targeting skills across multiple developmental domains, and many have strong parent involvement or training components. CTMs have been in existence for over 30 years and new models continue to be created. In a recent review, Odom et al. (2010) identified 30 CTMs that have been developed over the last three decades and are still in operation. Examples of historic CTMs are the UCLA Young Autism Project (which we will call the Lovaas model) (Lovaas, 1987), Treatment and Education of Autistic and Communication Handicapped Children (TEACCH) (Mesibov, Shea, & Schopler, 2005), the LEAP model (Hoyson, Jamieson, & Strain, 1984), and the Denver Model (Rogers et al., 2006).

A key movement in the field of educational and psychological interventions for students with ASD and other disabilities has been to establish the evidence-base for practices (National Autism Center's National Standards Project, 2009; National Professional Development Center on ASD, 2007), with a primary source of evidence for CTMs being supplied by efficacy studies. Yet, there remains a scarcity of high-quality, model-specific or comparative evaluation information about CTMs (Hume & Odom, in press). More than half of the 30 models reviewed by Odom et al. (2010) had no evidence of efficacy published in a peer-reviewed journal. In a recent critical review of CTMs for young children with ASD and their families, Rogers and Vismara (2008) evaluated the current research on comprehensive treatments for young children with ASD, finding limited evidence of efficacy for all but the Lovaas model, with some limited support for Pivotal Response Treatment (PRT) (Koegel, Koegel, Harrower, & Carter, 1999). Even the Lovaas model has been questioned. In recent evaluation report by the What Works Clearinghouse (2010), evaluators rated the Lovaas model as having small effects. Importantly, this evaluation was based on only two articles (from the more than 50 publications reviewed) that met the methodological inclusion criteria.

To respond to the need for high-quality treatment efficacy research, NIH convened a panel of investigators convened by NIH to determine the key needs around designing efficacy research, and they concluded that conducting comprehensive treatment research in the community settings and analyzing the relative effects of different CTMs were important (Lord et al., 2005). In a follow-up to that panel discussion, Smith, Scahill, et al. (2007) proposed a process for developing a program of research that began with manualization of treatment procedures and establishing fidelity/implementation protocols. In their review of comprehensive treatment programs, Odom et al. (2010) found as a group, CTMs were strongest in the operationalization (i.e., providing manualized descriptions of the content and procedures involved in model implementation) of their models; however, the actual measurement of implementation was relatively weak in comparison. Only one CTM had a fidelity measure with preliminary psychometric data, while 25% had informal or no methods of measuring fidelity.

The limited evidence of implementation measurement is also evidence in the ASD literature on focused intervention practices (i.e., single intervention designed to address outcome more limited in scope than would be addressed in CTMs). Wolery and Garfinkle (2002), in their review of the intervention literature from 1970 to the early 1990s, found that only 13% of the studies including students with ASD reported procedural fidelity information. When reviewing only single case design studies involving young children with ASD published from 1990 to 2003, Odom et al. (2003) found that 32% of the studies included implementation measures. A more recent finding indicated that only 18% of intervention studies for students with ASD published between the years of 1993–2004 assessed and reported treatment fidelity data (Wheeler, Baggett, Fox, & Blevins, 2006).

Both the limited use of implementation measures and the even more limited documentation of the reliability and validity of these measures is particularly problematic for several reasons. First, failure to ensure the integrity or fidelity with which an intervention is delivered poses a number of threats to drawing valid inferences about treatment effects (Gresham, MacMillan, Beebe-Frankenberger, & Bocian, 2000). If the components of the treatment are not well measured, no definitive conclusions can be drawn regarding the effects of the independent variables on the outcome measures. Therefore, the perceived presence of a functional relationship between the dependent (e.g., child outcomes) and independent variables (e.g., specific CTM) may be faulty (Wheeler et al., 2006). Further, in efficacy research, without a measure of fidelity, researchers cannot (a) fully account for differences between experimental and control groups (e.g., the degree to which control classrooms are using intervention components); (b) determine whether unsuccessful outcomes are due to an ineffective intervention; or (c) determine if poor outcomes are a result of a failure to implement the intervention as intended (O'Donnell, 2008). Perhaps most importantly, is the relationship between implementation and intervention outcomes. In a review of meta-analyses of community based health programs, Durlak and DuPre (2008) summarized that programs with stronger adherence resulted in mean effect sizes that were 2–3 times higher than programs with poorer implementation (up to 12 times higher in several studies). In a review of a subset of the studies (59 studies related to program implementation), Durlak and DuPre found that in 76% of the studies there was a significant positive association between the level of implementation and at least half of the program outcomes (2008). Researchers can postulate that this direct relationship between degree of treatment integrity and degree of treatment outcome is applicable to the ASD population and related intervention research.

The purpose of this study was to analyze the reliability and validity of implementation measures for two established CTMs for students with ASD, the TEACCH model and the LEAP model. The measures are currently being used in an examination of relative efficacy of the two models, and the implementation measure development is consistent with the Smith, Scahill, et al. (2007) recommendation concerning the need for establishing measurement of treatment fidelity in efficacy research. The specific research questions are: (1) Are the two implementation measures reliable as judged by the assessment of inter-rater agreement, test–retest reliability, and internal consistency? and (2) Are the implementation