An Internal and Critical Review of the PEAK Relational Training System for Children with Autism and Related Intellectual Disabilities: 2014–2017

Mark R. Dixon[1] · **Jordan Belisle**[1] · **Autumn McKeel**[2] ·
Seth Whiting[3] · **Ryan Speelman**[4] · **Jacob H. Daar**[5] ·
Kyle Rowsey[6]

Abstract The PEAK Relational Training System was designed as an assessment instrument and treatment protocol for addressing language and cognitive deficits in children with autism. PEAK contains four comprehensive training modules: *Direct Training* and *Generalization* emphasize a contingency-based framework of language development, and *Equivalence* and *Transformation* emphasize an approach to language development consistent with Relational Frame Theory. The present paper provides a comprehensive and critical review of peer-reviewed publications based on the entirety PEAK system through April, 2017. We describe both psychometric and outcome research, and indicate both positive features and limitations of this body of work. Finally, we note several research and practice questions that remain to be answered with the PEAK curriculum as well as other many other autism assessment and treatment protocols that are rooted within the framework of applied behavior analysis.

✉ Mark R. Dixon
mdixon@siu.edu

1. Behavior Analysis and Therapy Program, Southern Illinois University, Carbondale, IL 62901, USA

2. Aurora University, Aurora, IL, USA

3. Central Michigan University, Mt Pleasant, MI, USA

4. Pittsburg State University, Pittsburg, KS, USA

5. Northern Michigan University, Marquette, MI, USA

6. University of Southern Mississippi, Hattiesburg, MS, USA

If you're not your own severest critic, you are your own worst enemy.
- Jay Maisel

Applied behavior analysis (ABA) is a field dedicated to objectivity and reliable demonstrations of environmental treatment effects on socially significant behavior. Within this framework exists a systematic approach to the development of behavioral technologies and reliance on hard scientific evidence to inform treatment. Evidence of treatment effectiveness can be found in tightly controlled studies and the replication of treatment effects throughout nearly 50 years of published research in the *Journal of Applied Behavior Analysis* as well as other behavior analytic journals. Adherence to empirically validated treatment is an essential feature of ABA and more broadly to the standard of evidence based practice. The National Institute of Health (NIH) defines evidence-based practice as applying the "best available research results (evidence) when making decisions and designing programs and interventions" (National Institutes of Health, 2017). From among a vast array of available procedures, it is the ethical duty of applied service providers to choose the *best available* treatment based on several contextual factors (e.g., professional judgment, interests of the client) and, of course, the level of empirical support. Practitioners should be hesitant to adopt treatment protocols that have not been formally evaluated or that are supported by limited empirical evidence of effectiveness.

Where evidence of effectiveness is concerned, individual studies are not as persuasive as successive replications, as may be summarized in literature reviews or meta-analyses. Such systematic reviews can provide a comprehensive account of overall treatment efficacy while considering the relative strengths and limitations of individual studies. From a consumer's standpoint, a synthesis of the available literature may help clinicians identify the "best available research results" when considering a range of available treatments (NIH, 2017). The guiding principle of philosophic doubt requires scientist/practitioners to critically evaluate what constitutes the "best available" treatments (Cooper, Heron, & Heward, 2007). Only through continued evaluation and self-examination may researchers provide the best information to practitioners so they may successfully choose treatments that have the most supporting evidence.

Background: ASD Language Instruction Interventions

One area of ABA that has garnered much attention for its apparent effectiveness and evidenced-based support is in the delivery of language instruction for children with autism spectrum disorder (ASD). As a pervasive developmental disorder, ASD often presents as a life-long disability associated with repetitive behaviors and perseverative interests, deficits in social skills and social interactions, and delays in language acquisition and pragmatic language use (American Psychiatric Association, 2013). Deficits experienced by individuals with ASD vary greatly, ranging from severe (e.g., an inability to communicate even basic needs) to mild (e.g., intact communication abilities with social skill deficits and repetitive interests). ABA interventions for reducing the range of disabilities associated with ASD are currently among the most scientifically established and evidenced-based approaches available. Furthermore, ABA based interventions for ASD have been endorsed by the Surgeon General of

the United States (U.S. Public Health Service, 1999), the National Research Council (2001), the American Academy of Pediatrics (Myers & Johnson, 2007), and the National Autism Center (Howard, Ladew, & Pollack, 2009).

ABA language protocols for ASD are often delivered as a packaged intervention and referred to as Comprehensive Behavior Interventions or as Early-Intensive-Behavior Interventions (EIBI). While a variety of characteristics and procedures are associated with these approaches, key features often include an emphasis on broad programing centered around a single conceptual framework, the use of discrete trial training (DTT) among other reinforcement and naturalistic based procedures, one-on-one therapy for more than 10 hours per week, and individualized programming to address developmentally appropriate skills across a number of domains (Smith & Iadarola, 2015; Wong, Odom, Hume, Cox, Fettig, Kucharczyk, & Schultz, 2015).

Since the seminal work of Lovaas (1981, 1987, 1993; McEachin, Smith, & Lovaas, 1993), ABA procedures and EIBI programs for language learning have been the subject of hundreds of peer-reviewed studies including both single-subject demonstrations, randomized control trials, and systematic reviews. Meta-analysis and other systematic research reviews have suggested that, while treatment components may vary, EIBI approaches are generally more effective for children with autism than are non-specific or eclectic treatments (Eldevik, Hastings, Hughes, Jahr, Eikeseth, & Cross, 2009, 2010; Howlin, Magiati, & Charman, 2009; Myers & Johnson, 2007; Peters-Scheffer, Didden, Korzilius, & Sturmey, 2011; Reichow, Barton, Boyd, & Hume, 2012; Warren, McPheeters, Sathe, Foss-Feig, Glasser, & Veenstra-VanderWeele, 2011). For instance, Eldevik et al. (2010) (see also Eldevik et al., 2009) concluded that after a review of individual participant data from 34 studies, children receiving intensive behavior interventions were significantly more likely to show reliable positive changes in IQ and adaptive behaviors than those who did not receive EIBI interventions.

While reviews have generally been positive, some have been critical of the evidence presented in studies evaluating EIBI procedures. Warren et al. (2011) noted that gains in cognitive performance, language skills, and adaptive behavior were associated with EIBI treatments, but the overall evidence for EIBI was limited due to the lack of high-quality RCT studies and no studies that have directly compared manualized treatment approaches. Partially illustrating this point, Spreckley and Boyd (2009) conducted a literature review which included 13 RCT or quasi-RCT studies that included reporting on DTT procedures. The results of this review found that only 6 studies could be rated as having "adequate internal validity for quantitative meta-analysis" according to the Physiotherapy Evidence Database (PEDro) Scale, and only 4 of these studies included the information needed to be included in a meta-analysis. That analysis showed little evidence that ABA improved cognitive, language, or adaptive skills better than standard care. Even when EIBI appears to be effective at the group level, Howlin et al. (2009) noted that gains are quite variable at the individual level.

Operationalization and manualization of models plays an important role in treatment evaluation and therefore in the determination of evidence-based best practice. Manualized protocols allow researchers to evaluate utility across a broad set of prescribed learning objectives; a factor that is often limited when efficacy is evaluated in non-standardized interventions (Odom, Boyd, Hall, & Hume, 2010a; Odom, Collet-Klinenberg, Rogers, & Hatton, 2010b). Additionally, manualized protocols allow for the standardization of implementation tools (e.g. Gould, Dixon, Najdowski, Smith, & Tarbox, 2011) and

assessment of procedural fidelity (Bellg et al., 2004). In a review of comprehensive treatment models for ASD, Odom et al. (2010a) observed that out of 30 models (20 of which were ABA-based), few included tools to collect implementation data and only one included psychometric assessment of the implementation tools. Furthermore, efficacy evidence in peer-reviewed journals was available for only 16 out of the 30 models. Lack of RCT investigations was noted as a major limitation in determining efficacy. Indeed, most systematic reviews of EIBI procedures expressly excluded the single subject research designs (e.g., Eldevik et al., 2009, 2010; Howlin et al., 2009; Warren et al., 2011) which many behavior analysts focus research efforts on.
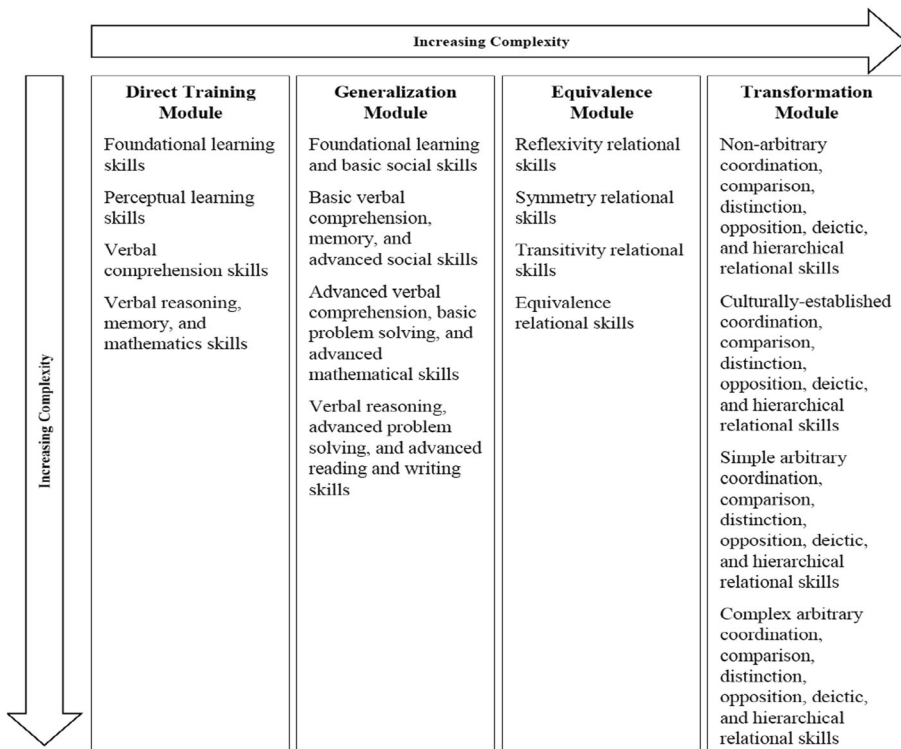
The PEAK System

Despite the availability of operationalized models and manualized curricula, many ABA practitioners continue to design language learning programs by borrowing procedures from more than one published curriculum (Love, Carr, Almason, & Petursdottir, 2009). Per Love et al. (2009), the two most commonly referenced curricula models were those published by Sundberg and Partington (1998) and Lovaas (1981). Assessment protocols focusing on Skinner's verbal operants, such as the *Verbal Behavior Milestones Assessment and Placement Program* (VB-MAPP; 2008) and the *Assessment of Basic Language and Learning Skills—Revised* (ABLLS-R; 2008), have become common components of EIBI programs (e.g. Fisher & Zangrillo, 2015) despite the fact that few published studies have assessed the psychometric properties of these assessment protocols or the efficacy treatments based on them. An additional limitation is that, by focusing exclusively on Skinner's verbal operants, these protocols ignore hundreds of research studies on how language develops through derived relational responding and corresponding transformations of stimulus function (Hayes, Barnes-Holmes, & Roche, 2001). Although the bulk of the behavioral literature focuses on implementing verbal operant procedures with children with autism, also critical to consider are advances in our understanding of language and cognitive development resultant from stimulus equivalence theory (Sidman, 1971; Sidman & Tailby, 1982), naming theory (Horne & Lowe, 1996), and Relational Frame Theory (Hayes et al., 2001). At the theoretical level, attempts have been made to integrate approaches based on verbal operants and derived relations (Barnes-Holmes, Barnes-Holmes, & Cullinan, 2000), but ultimately what defines the best approach to developing language and cognition (i.e., verbal operant training exclusively, derived relational training exclusively, or some combination therein) will be published data, including both demonstrations of assessment reliability and validity, and experimental (both single-case and between-group designs) evaluations of treatment outcomes and ancillary changes in other domains of life functioning and quality.

The Promoting the Emergence of Advanced Knowledge (PEAK: Dixon, 2014a) system is an assessment and curriculum guide composed of four unique modules: Direct Training, Generalization, Equivalence, and Transformation. Each module offers 184 individual programs designed to encapsulate a distinct learning modality, and is consistent with a merger of traditional Skinnerian verbal operant training with post-Skinnerian procedures to produce derived relational responding. Within each module practitioners are provided detailed instructions on how to conduct initial assessments

and place clients into an appropriate skill range, based upon current abilities. Each program outlines goals, materials, and typical stimuli, as well as instructions on implementation and data collection. Skill progression through each module is summarized in Fig. 1. Although the modules generally increase in complexity from the first to the fourth, programs across modules are intended to be run synchronously to exercise each of the four modalities simultaneously; however, the best approach to progressing through the modules has not yet been empirically evaluated.

The first two modules were designed based on a contingency-based verbal operant account of language development. These two modules are similar to the VB-MAPP and ABLLS-R in terms of their theoretical construction (i.e., Skinner's verbal behavior theory), but may provide more complex topographies of verbal operant targets relative to the above assessments. For example, the VB-MAPP provides targets up to the skills expected of a typically developing 4-year old, whereas PEAK Direct Training provides targets up to age 8, and PEAK Generalization provides targets up-to age 11. PEAK Direct Training (PEAK-DT; Dixon, 2014a) teaches foundational language skills (e.g., eye contact, object permanence, echoics, mands) like Skinner's (1957) verbal operants using a discrete trial methodology in which specific discriminative stimuli come to occasion desired responses through prompt fading and delivery of reinforcers. As individuals gain proficiency the complexity of programming accelerates to include much more advanced

Fig. 1 Graphical display of the progression of skills targeted in and across each of the PEAK Relational Training System Modules. PEAK-DT and PEAK-G skill groups were taken from the component analysis for reach module. PEAK-E and PEAK-T skill groups were taken directly from the complexity conceptualizations found in each of the modules

topographies of these operants (e.g., metonymical tacts, autoclitic mands), along with social components of understanding the role of the audience in modifying verbal responses, discriminating private events to publically accompanying stimuli, guessing about events that have no literal correct answer, and using working memory. Early programs in PEAK–DT focus on the elementary forms of verbal operant behaviors that have been extensively empirically demonstrated with disabled populations (Dixon, Small, & Rosales, 2007; Dymond, O Hora, Whelan, & O Donovan, 2006), such as simple tacts (Arntzen & Almås, 2002) and mands (Hall & Sundberg, 1987), as well as rote intraverbal response sequences (Ingvarsson & Hollobaugh, 2011). Later programs in PEAK–DT are more complex forms of verbal operants, such as those proposed by Skinner that extend upon these basic verbal operants, upon which substantially less research has been conducted. PEAK Generalization (PEAK-G; Dixon, 2014b) moves beyond traditional discrete trial training using a train-test methodology in which novel untrained stimuli are presented within embedded blocks of directly trained stimuli with hopes that these never reinforced targets will come to occasion correct responses from the learner due to formal similarity with trained stimuli. This module is designed to promote stimulus and response generalization as an active process to establish and maintain skills in new and novel contexts. As with PEAK–DT, empirically supported procedures for promoting generalization of verbal operant responses are well established in the behavior analytic literature, and are the basis of the PEAK–G module. Early programs in PEAK–G target generalized response repertoires such as imitation and generalized tacts that are empirically established, and latter programs include generalized response topographies in need of further empirical evaluation (e.g., tacting actions metaphorically, superstitious manding).

The final two modules, *PEAK Equivalence* (PEAK-E; Dixon, 2015) and *PEAK Transformation* (PEAK–T; Dixon, 2016) offer a conceptually systematic approach that capitalizes on behavioral technologies derived from stimulus equivalence (Sidman, 1971) and Relational Frame Theory (Hayes et al., 2001). Currently these are the only comprehensive manualized protocols emphasizing derived relational responding in children with autism that are supported by peer-reviewed investigations of treatment outcomes. Unlike any previous autism curriculum, the methods efficiently teach verbal skills in which meaning transfers to new stimuli without direct training. This is done by providing exemplars that set the occasion for derived responding to emerge as a functional operant. Advances in verbal behavior approaches have shown that early learning skills can be taught under various sources of stimulus control and in natural environment settings (Johnson, Kohler, & Ross, 2017), but once these pre-requisite learning skills have been established, instructional programming can advance to more advanced, relational targets. PEAK-E begins by testing for the emergence of basic reflexive (e.g., identity matching), symmetrical and transitive responding across sense modalities (e.g., identifying the name of gustatory sensations), complex stimulus class formations (e.g., animal names and habitats), and transfer of function between stimuli (e.g., modification of performance during common games). PEAK-T evaluates and treats deficits in relational responding including equivalence and also the arrangements of opposite, difference, comparison, hierarchy, and perspective taking. Building from basic non-arbitrary relational tasks (sorting by size or color), through cultural conventions of relations (the word *pig* means a bigger animal than the word *ant*), to the most complex transformations involving arbitrary stimuli relations (if cux is like a sandwich and veb is better than a brick, which one, cux or veb, is better to hold your tent in place

during a wind storm?). Overall, the PEAK protocol claims to teach those with language deficits to "speak with meaning, and listen with understanding" (Hayes et al., 2001) through the implementation of the four component modules of the system.

Beginning in 2014, elements of the PEAK assessments and curricula have been the subject of a number of published research studies investigating the psychometric properties of the assessment tools themselves and the effectiveness of curriculum goals and procedures in producing skill acquisition. Because the relevant studies are dispersed across a number of journals, and address a variety of research questions with variety of designs and methodologies, it may be difficult for practitioners to track the overall patterns of the evidence base that would support a critical appraisal of the PEAK system. To help those who are unfamiliar identify the strengths and weakness of the PEAK program, the current study provides a systematic overview of the published literature concerning the validity of the PEAK assessment protocols and the treatment efficacy of procedures and curricula described within the PEAK model. The hope is that this survey will stimulate others to explore the potential and pitfalls of PEAK as one means of working toward the best possible language interventions for children with autism.

Methods

A review of the literature was conducted to identify peer-reviewed empirical evaluations of the PEAK Relational Training System that bear upon the efficacy of the system and therefore provide a basis for recommendations about future research on this applied technology. The reviewers (authors of the present investigation) identified all articles that have reported data on PEAK from January 2014 to April 2017, which was exhaustive at the time of the investigation. As described below, reviewers analyzed the social validity, subject demographics, psychometric evaluations, between-group and single-case evaluations of training efficacy, and critical limitations of each study.

Identifying Peer Reviewed Research Articles

Search Procedures Searches for peer-reviewed articles were conducted in two phases. First, we used the Google Scholar and PsycINFO journal databases to identify published articles. For all searches, peer-reviewed articles were isolated and the date perimeter was set from January 2014 to present. All searches were conducted on February 15, 2017. Search terms included "PEAK Relational Training System," "PEAK Direct Training," "PEAK Generalization," "PEAK Equivalence," "PEAK Transformation," and "Promoting the Emergence of Advanced Knowledge." Where more than 100 results were reported in either search engine, the first 100 results were considered sufficient for inclusion in the current study.[1] Second, a supplemental search was conducted by obtaining the curriculum vitas of each author of a published paper identified using the above method to capture any additional work that was in press from authors of prior PEAK publications. No archival source was available for systematically identifying unpublished articles from other authors.

---

[1] We knew that our group had conducted about two dozen studies, and thought it unlikely that other research teams had published more than four times that many without our knowledge.

Inclusion and Exclusion Criteria For an article to be included, it had to (a) explicitly utilize PEAK Relational Training System's assessment and/or curriculum, and (b) be published or in press in a peer reviewed journal. Articles were excluded if they were non-empirical (e.g., Reed & Luiselli, 2016) or if the training technology deviated substantially from what is described in the PEAK curriculum (e.g., Dixon, Belisle, Munoz, Stanley, & Rowsey, In press). In addition, our plan was to exclude any psychometric articles in which statistically significant outcomes were not presented or were of such poor quality that the outcomes were likely to have occurred by chance. We also would have excluded any outcome research that did not have a control comparison consistent within the logic of either single-case or between-group experimental design (e.g., pre-post test only with no-control group; case study); however, no articles were identified that had to be excluded based on the criteria.

Data Extraction Table 1 shows the types of data that were extracted from the 21 articles that were identified in our search. Information recorded for each study included the type of study (single case experimental, between-group experimental, or psychometric non-experimental), the module from PEAK Relational Training System (Direct, Generalization, Equivalence, or Transformation), the authors, year of publication, the title, the participants, design, results, and the first three limitations described in each study by the authors of the study.

Interrater Agreement One reviewer extracted the information from all 21 articles, and an independent reviewer evaluated 7 of the articles to determine agreement between independent reviewers for the categories shown in Table 1. If, for a given category for a given article, the same information was extracted by both reviewers, this was counted as an agreement. If the second reviewer extracted the same information as the first, but provided additional information, this was counted as an agreement. If the second reviewer omitted information that the first reviewer included, this was counted as a disagreement. To determine the interrater agreement between the reviewers, the number of agreements in each category from Table 1 (agreements) was divided by the number of agreements + disagreements and multiplied by 100 ([agreements / agreements + disagreements] x 100). IOA of 95.9% was obtained between two reviewers.

Results

Table 1 summarizes the information that was extracted from the 21 articles. Ten of these were single-case evaluations of programs contained in the PEAK modules, 1 was a randomized control trial evaluation of the PEAK curriculum, and 10 were psychometric evaluations of the PEAK assessments. Collectively, the studies examined aspects of all four PEAK modules, including PEAK-DT (2 single subject evaluation of treatment outcome, 1 RCT evaluation of treatment outcomes, and 8 psychometric evaluations of assessment), PEAK-G (1 single-subject, 2 psychometric), PEAK-E (6 single subject), and PEAK-T (1 single subject).

Table 1 Summary of research conducted between 2014–2017 on the PEAK Relational Training System

Study Type	Module	Authors	Year	Title	Participants/ Design	Results Summary	Limitations Noted
Single-Case Experimental	Direct Training	Belisle, Rowsey, Dixon	2016	The use of in-situ behavioral skills training to improve staff implementation of the PEAK relational training system	3 direct-care staff, 3 children with autism / Multiple-baseline across staff-student dyads	- Mastery of staff implementation fidelity for all staff - Improvement in PEAK performance for 2 out of 3 children	- In-situ BST was not compared to other staff training models - Short maintenance phase - IOA was not conducted on child performance
	Direct Training	McKeel, Rowsey, Belisle, Dixon, Szekely	2015	Teaching complex verbal operants with the PEAK Relational Training System	3 children with autism / Multiple-baseline across subjects across 3 PEAK programs	- Mastery of directly reinforced basic autoclitics, receptive metonymical tacts, expressive metonymical tacts, expressive planet identification for 2 participants - Mastery of directly reinforced guessing, receptive metonymical tacts, expressive metonymical tacts for 1 participant	- Similar age and diagnoses of participants - Limited number of stimuli used in each program - Limited trial blocks to mastery criterion
	Generalization	Dixon, Peach, Daar, Penrod	2017	Teaching complex verbal operants to children with autism and establishing generalization using the peak curriculum	3 children with autism / Multiple- baseline across skills, replicated across 3 participants	- Mastery of directly reinforced creative paths, autoclitic mands, distorted tacts - Generalized mastery of untrained exemplars of creative paths, autoclitic mands, distorted tacts	- Minor differences in stimulus topography across train and test stimuli - Maintenance of skills was not examined
	Equivalence	Stanley, Belisle, Dixon	In Press	Equivalence-based instruction of academic skills: Application to adolescents with autism	3 adolescents with autism/ Multiple-baseline across participants, replicated across 3 skills	- Mastery of directly trained chemistry technical name (A) – symbol (B) relation and symbol (B) – common name (C) relation - Derived mastery of untrained equivalent technical name (A) – common name (C) relation and common name (C) - technical name (A) relation - Mastery of directly trained measurement inches (B) – yards (A) relation and measurement feet (C) – yards (A) relation - Derived mastery of untrained equivalent inches (B) – feet (C) relation and feet (C) - inches (B) relation - Mastery of directly trained history name (A) – picture (B) relation, name (A) to fact (C) relation, and name (A) to region (D) relation - Derived mastery of untrained equivalent picture (B) – fact	- Cannot assume that procedures can be replicated with individuals with fewer verbal skills - Experimenters were not typical classroom staff - Symmetry tests were not conducted for all relations

Table 1 (continued)

Study Type	Module	Year	Title	Authors	Participants/ Design	Results Summary	Limitations Noted
	Equivalence	2017	Derived coordinated symmetrical and transitive gustatory-visual-auditory relations in children with disabilities	Dixon, Belisle, Stanley, Munoz, Speelman	2 children with autism, 1 child with cognitive delay / Multiple-baseline across skills, replicated across 3 participants	(C) relation, region (D) – picture (B) relation, and fact (C) – region (D) relation - Mastery of the directly reinforced gustatory (A) – visual (B) relation - Mastery of directly reinforced visual (B) – auditory (C) relation - Derived mastery of untrained symmetrical auditory (C) – visual (B) relation - Derived mastery of untrained transitive gustatory (A) – auditory (C) relation for 2 of 3 participants - Emergence of mands for 2 of 3 participants	- High levels of baseline performance for 1 participant - Mixed training was required for 1 participant - Mastery of the transitive relation was not demonstrated by 1 participant
	Equivalence	2016	Evaluating emergent naming relations through representational drawing in individuals with developmental disabilities using the PEAK-E curriculum	Dixon, Belisle, Rowsey, Speelman, Stanley, Kime	1 child with autism, 1 child with autism and ADHD, 1 child with an emotional and cognitive delay / Multiple-baseline across subjects	- Mastery of directly reinforced animal blend picture (A) – vocal animal name (B) relation - Derived mastery of untrained symmetrical vocal animal name (B) – animal blend picture (A) relation	- Replication may be unlikely with participants who cannot already name and draw - Several aspects of naming were not tested
	Equivalence	2016	Establishing derived categorical responding in children with disabilities using the PEAK-E curriculum	Dixon, Belisle, Stanley, Speelman, Rowsey, Daar	1 child with an intellectual disability and speech impairment, 1 child with a cognitive delay, 1 child with autism / Multiple-baseline across subjects	- Mastery of directly reinforced category item (A) – category item (B) relation - Mastery of directly reinforced category item (B) – category item (C) relation - Mastery of directly reinforced category name (D) – category item (A) relation - Derived mastery of untrained transitive category name (A) – category items (A/B/C) relations under a lag schedule - Emergence of untrained intraverbal naming of category items under a lag schedule for 2 of 3 participants	- Emergence of intraverbal naming of category items was not observed for 1 participant - PEAK scores were used instead of percentage correct
	Equivalence	2016	Derived equivalence relations of geometry skills in students with autism: An	Dixon, Belisle, Stanley, Williams	2 adolescents with autism / Multiple-baseline across subjects	- Mastery of directly reinforced vocal shape name (A) – vocal shape property (B) relation	- Participants had high relational abilities at the on-set of the study

Table 1 (continued)

Study Type	Module	Authors	Year	Title	Participants/ Design	Results Summary	Limitations Noted
				application of the PEAK-E curriculum		- Derived mastery of untrained symmetrical vocal shape property (B) – vocal shape name (A) relation - Derived mastery of untrained transitive vocal shape name (A) – shape picture (C) relation	- Participant's abilities to identify "sides" of shapes were not tested at the on-set of the study - no B-A or A-C test probes were conducted during the training phase
	Equivalence	Dixon, Speelman, Rowsey, Belisle	2016	Derived rule-following and transformation of stimulus functions in a children's game: An application of PEAK-E with children with developmental disabilities	1 child with an intellectual disability, 1 child with an emotional and learning disability, 1 child with autism / Multiple-baseline across subjects across 3 skills	- Mastery of directly reinforced known body part name (A) – unknown body part name (B) relation - Mastery of directly reinforced unknown body part name (B) – unknown body part name (C) relation - Emergence of untrained performance in a game of Twister using unknown words for head, hand, and foot	- Only 3 stimulus classes were established for each participant - Combinatorial entailment (transitivity and equivalence) was not directly tested - Sequence effects may have been introduced by only training a single relational class in each phase
	Transformation	Belisle, Dixon, Stanley, Munoz, Daar	2016	Teaching foundational perspective-taking skills to children with autism using the PEAK-T curriculum: Single-reversal "I-You" deictic frames	3 adolescents with autism / Multiple-baseline across subjects	- Mastery of directly reinforced single-reversal 'YOU' relation - Derived mastery of untrained single-reversal 'I' relation for 2 of 3 participants - Mastery of directly reinforced single-reversal 'I' relation for 1 participant - Transfer of untrained single-reversal 'YOU' and 'I' relations to untrained stimuli	- Only simple visual perspective taking was evaluated - 1 participant showed high levels of correct responding to the single-reversal 'I' relation in baseline - Mixed 'YOU' and 'I' training was required for one of the participants
Between-Group Experimental	Direct Training	McKeel, Dixon, Daar, Rowsey, Szekely	2015	The efficacy of the PEAK relational training system using a randomized controlled trial of children with autism	40 children with autism (27 experimental, 13 control) / Randomized control-trial evaluation	- Statistically significant differences between the groups were observed from time 1 to time 2 on PEAK assessment scores	- Small sample size increases the chance of a sampling error - Skills were reassessed after only 1-month

Table 1 (continued)

Study Type	Module	Authors	Year	Title	Participants/ Design	Results Summary	Limitations Noted
						- Statistically significant differences were observed in the experimental group from time 1 to time 2 PEAK assessment scores - Statistically significant differences were not observed in the control group from time 1 to time 2 PEAK assessment scores	- Treatment fidelity was not assessed
Psychometric Non-Experimental	Direct Training	Dixon, Stanley, Belisle, Rowsey	2016	The test-retest and interrater reliability of the Promoting the Emergence of Advanced Knowledge-Direct Training assessment for use with individuals with autism and related disabilities	39 individuals with autism / correlation between assessments by independent observers and at two time occasions	- Strong, positive correlation between scores by independent observers - High interobserver agreement across independent observers - Strong, positive correlation between raw scores at both measurement periods, across PEAK Factors - Strong, positive correlation between age-normed scores at both measurement periods, across PEAK Factors	- Small sample size - Reliability of only one PEAK module was tested - No comparison between items that were indirectly tested with items that were directly tested
	Direct Training	Malkin, Dixon, Speelman, Luke	2016	Evaluating the relationship between the PEAK Relational Training System – Direct Training Module, Assessment of Basic Language and Learning Skills – Revised, and the Vineland Adaptive Behavior Scales – II	21 children with autism / correlation between three assessments	- A strong, positive correlation was observed between PEAK-DTA and ABLLS-R - A moderate, positive correlation was observed between PEAK-DTA and VABS-II - A strong, positive correlation was observed between ABLLS-R and VABS-II	- Small sample size - All assessments were conducted non-concurrently - IOA was not conducted
	Direct Training	McKeel, Rowsey, Dixon, Daar	2015	Correlation between PEAK relational training system and one-word picture vocabulary tests	40 individuals with autism / correlation between two assessments	- A strong, positive correlation was observed between PEAK-DTA and ROWPVT-4 - A strong, positive correlation was observed between PEAK-DTA and EOWPVT-4	- IOA was not conducted - Assessments were not conducted at the same time each day - Small sample size with large age range (5-21)

Table 1 (continued)

Study Type	Module	Year	Title	Authors	Participants/ Design	Results Summary	Limitations Noted
	Direct Training	2014	Normative sample of the PEAK Relational Training System: Direct Training Module and subsequent comparisons to individuals with autism	Dixon, Belisle, Whiting, Rowsey	206 individuals without a diagnosed disability, 70 individuals with autism	- A strong, positive correlation was observed between ROWPVT-4 and EOWPVT-4 - A strong, positive correlation was observed between PEAK-DTA and age for the normative sample - A cubic regression provided the strongest fit for the normative data - No correlation was observed between PEAK-DTA and age for the autism sample - Individuals with autism scores significantly lower than the normative sample on the PEAK-DTA	- Small sample size - Sample was homogenous
	Direct Training	2014	Assessing the relationship between intelligence and the PEAK Relational Training System	Dixon, Whiting, Rowsey, Belisle	43 individuals with autism, 4 individuals with an intellectual disability, 1 individual with ADHD, 1 individual with Fragile X, 1 individual with OHI / Correlation between PEAK-DTA and IQ	- A strong, positive correlation was observed between PEAK-DT-A and measures of IQ - no correlation was observed between PEAK-DTA and age	- Several different IQ tests were used in the study - IQ tests taken from student records varied in the time since they were administered
	Direct Training	2014	Principal component analysis of the PEAK relational training system	Rowsey, Belisle, Dixon	98 individuals with autism or a related intellectual or developmental disability / Principle component analysis of the PEAK-DTA	- The results of the principal component analysis yielded four factors - A high percentage of variance is explained in the four-factor model - Cronbach's alphas assessing the internal consistency of each factor as well as the PEAK-DTA indicated strong internal consistency for each	- All participants had autism or a related disability - Small sample size
	Direct Training	2014	PEAK relational training system for children with autism and developmental disabilities: Correlations with Peabody picture vocabulary test and assessment reliability	Dixon, Carman, Tyler, Whiting, Enoch, Daar	13 children with autism or a related intellectual or developmental disability / Correlation between PEAK-DTA and two assessments	- A strong, positive correlation was observed between PEAK-DT-A and PPVT - A strong, positive correlation was observed between PEAK-DT-A and Illinois Early Learning Standard Scores	- Small sample size - Participant demographics were limited (ages and diagnoses)

Table 1 (continued)

Study Type	Module	Authors	Year	Title	Participants/ Design	Results Summary	Limitations Noted
	Direct-Training/ Generalization	Dixon, Belisle, Stanley, Rowsey, Daar, Szekely	2015	Toward a behavior analysis of complex language for children with autism: Evaluating the relationship between PEAK and the VB-MAPP	40 individuals with autism / Correlation between three assessments	- Strong, significant correlation between PEAK-DTA + PEAK-GA and each level of the VB-MAPP and VB-MAPP Total Score - Strong, significant correlation between PEAK-DTA and each level of the VB-MAPP and VB-MAPP Total Score - Strong, significant correlation between PEAK-GA and each level of the VB-MAPP and VB-MAPP Total Score	- Only two PEAK modules were included in the analysis - Small sample size
	Generalization	Rowsey, Belisle, Stanley, Daar, Dixon	2017	Principal component analysis of the PEAK Generalization module	84 individuals diagnosed with an autism spectrum disorder/Principal Component analysis of the PEAK-GA	- The results of the principal component analysis yielded four factors - A high percentage of variance is explained within the model - Cronbach's alphas assessing the internal consistency of each factor as well as the PEAK-DTA indicated strong internal consistency for each	- A relatively homogenous sample - Small sample size - All participants had a diagnosis of autism
	Generalization	Dixon, Rowsey, Gunnarsson, Belisle, Stanley, Daar	2016	Normative sample of the PEAK Relational Training System: Generalization Module with comparison to individuals with autism	183 individuals without a diagnosed disability, 84 individuals with autism	- A strong, positive correlation was observed between PEAK-DTA and age for the normative sample - No correlation was observed between PEAK-DTA and age for the autism sample - Individuals with autism scores significantly lower than the normative sample on the PEAK-DTA	- Small sample sizes for both groups - Autism sample was from the same state where as the normative sample was from several states - Demographic data other than age, gender, and state of residence are not reported for the normative sample

Psychometric Properties of the PEAK System

One of the standard components of psychometric evaluations is reliability, which implies that a measurement tool measures in the same way each time it is employed. Table 2 provides summary information which indicates good reliability for the first two PEAK modules in the form of test-retest reliability (Dixon, Stanley, Belisle, & Rowsey, 2016f), assessment administration fidelity (McKeel, Rowsey, Dixon, & Daar, 2015c), and inter-observer reliability (IOR; Dixon, Carman, Tyler, Whiting, Enoch, & Daar, 2014b; Dixon et al., 2016f; Dixon, Whiting, Rowsey, & Belisle, 2014c; McKeel et al., 2015c). Reliability has not been evaluated for the other two PEAK modules. Four studies that assessed IOR and found that it ranged from 85% to 99.1%; one study reported a Cohen's kappa reliability coefficient of .981.

A second component of psychometric evaluations is validity, which implies that a measurement tool measures what it is supposed to. Convergent validity is assessed when outcomes of a target measurement tool are correlated with outcomes from a different, but, thematically-related tool. Table 2 shows that, to date, PEAK assessment outcomes have been found to correlate strongly with the following other assessments of language, cognitive, or adaptive functioning (median r = +.908): IQ (as measured by multiple tests based on participant age and/or agency preference), Assessment of Basic Learning and Language Skills – Revised (ABLLS-R; Partington, 2006), Vineland Adaptive Behavior Scales (VABS-II; Furniss, 2009), Peabody Picture Vocabulary Test (PPVT; Dunn & Dunn, 2007), Illinois Early Learning Standards (Illinois State Board of Education, 2013), and One-Word Picture Vocabulary Tests, both Expressive (EOWPVT; Martin & Brownell, 2011a) and Receptive (ROWPVT; Martin & Brownell, 2011b). Finally, PEAK scores appear not to be correlated with chronological

Table 2 Summary of correlation coefficients and tests of statistical significance across convergent validity evaluations of PEAK

PEAK Assessment	Correlated Assessment	Correlation	$p < 0.05$?
Test-Retest Reliability			
PEAK-Direct Training	PEAK – DTA after 1-month	PEAK-DTA Raw Scores: ICC: 0.987 PEAK-DTA Normative Scores: ICC: 0.973	Yes
Convergent Validity			
PEAK-Direct Training	VB-MAPP	Pearson's r: 0.932	Yes
	ABLLS-R	Pearson's r: 0.951	Yes
	VABS-II	Pearson's r: 0.453	Yes
	ROWPVT-4	Pearson's r: 0.908	Yes
	EOWPVT-4	Pearson's r: 0.901	Yes
	Illinois Early Learning Standard	Pearson's r: 0.916	Yes
	IQ	Pearson's r: 0.759	Yes
	Age	Pearson's r: -0.086	No
	PPVT	Pearson's r: 0.908	Yes
PEAK-Generalization	VB-MAPP	Pearson's r: 0.577	Yes

age of persons with ASD, but are for neurotypical children who were administered the assessment. This suggests that PEAK scores might be useful in identifying levels of developmental deviance in persons with ASD.

Content validity also has been assessed for the first two PEAK modules. Normative samples and samples of individuals with autism spectrum disorders were assessed across a range of ages in both the PEAK-DT (Dixon, Belisle, Whiting, & Rowsey, 2014a) and the PEAK-G (Dixon, Belisle, Stanley, Munoz, & Speelman, 2017). Both studies found that total scores on both the PEAK-DTA and PEAK-GA were strongly correlated with age in a normative sample, however not correlated with age in a sample of individuals with autism. This finding is consistent with the observation that individuals with autism tend to deviate from typically developing peers in terms of language development. It also suggests that items on both assessments may be applicable to individuals with autism across age groups, depending on their idiosyncratic level of language impairment.

A different approach to content validity was taken by examining the PEAK-DTA [2] and PEAK-GA through Principal Component Analyses (PCA; Rowsey, Belisle, & Dixon, 2014; Dixon, Rowsey, Gunnarsson, Belisle, Stanley, & Daar, 2017). Findings indicate that the skills in both the PEAK DT and PEAK G can be grouped into four distinct components. The authors entitled these components, or factors, very broadly and as such they articulate a range of skills that statistically cluster together in terms of development within a child's repertoire. For example, in the PEAK DTA, the four factors were "Foundational Learning Skills" (eye contact; keeping hands still; basic echoics; following 1 step directions), "Perceptual Learning Skills" (tact and listener responding), Verbal Comprehension Skills" (basic social skills; delayed picture matching or naming), and Verbal Reasoning, Memory & Math Skills (speaking with and understanding metaphors; the role of the audience as a listener; private events)."

Overall, psychometric studies suggest that PEAK provides a sound measurement system that could have utility separate from PEAK-based interventions. That is, PEAK assessment scores might serve as a standardized dependent variable for evaluating the effectiveness of any intervention. Possibilities include well known behavior-analytic approaches to enhancing language and cognition, such as VB-MAPP, functional communication training, social skills programs, and augmentative communication systems. PEAK also may be suitable for evaluating the effectiveness of "treatment as usual", such as the instruction provided in a special education classroom, which could be evaluated in terms of student growth as measured by PEAK scores at the beginning and end of the academic year. The relatively rapid assessment that is possible with PEAK might therefore provide a cost-effective support to meeting the increasing demands for accountability within special education (e.g., Cusumano, 2007; Snow, Burns, & Griffin, 1998). Finally, the standardized assessment protocols of PEAK may prove useful in evaluating fad treatments for children with autism, such as special diets (see Jacobson, Foxx, & Mulick, 2005).

---

[2] When referring specifically to the assessment portion of a module we will append "A" (for assessment) to the relevant acronym. Thus, PEAK-DTA is the assessment from the PEAK-DT module. PEAK-GA is the assessment from the PEAK-G module, and so forth.

Effectiveness: Group Design Research

To date, only one study has employed a group-design (RCT) to evaluate the effectiveness of the PEAK system with children with autism (McKeel, Dixon, Daar, Rowsey, & Szekely, 2015a). This study compared PEAK-DTA scores for a control group of individuals with autism receiving treatment as usual (i.e., instruction using their school's typical curriculum) and a treatment group of individuals with autism receiving two training sessions per week from the PEAK-DT Curriculum both before and after intervention. The study found that there were statistically significant differences between pre-test and post-test scores within the experimental (PEAK-DT) group, but not within the control group. Additionally, statistically significant differences were found between the experimental and control groups' post-test scores on the PEAK-DTA. We used raw pre-test and post-test PEAK-DTA scores to determine the mean, standard deviation, and correlations of within-subject assessment results to calculate effect size (Cohen's d) for both groups. For the experimental group, $d = .99$, which is generally considered a large effect size (Cohen, 1988). For the Control Group $d = 0.26$, which is generally considered a small effect size. Overall, these findings indicate that the PEAK DT curriculum was more effective than treatment as usual in increasing participants' PEAK DTA scores.

Effectiveness: Single Subject Design Research

Table 3 shows that single-subject research on PEAK has included a wide array of advanced verbal and academic targets, including response variability as an operant (Dixon, Peach, Daar, & Penrod, 2017), tact extensions (McKeel, Rowsey, Belisle, Dixon, & Szekely, 2015b), autoclitics (Dixon, Peach, Daar, & Penrod, 2017), and stimulus equivalence outcomes such as derived listener responding (Dixon, Belisle, Rowsey, Speelman, Stanley, & Kime, 2016a) and gross motor equivalence (Dixon, Speelman, Rowsey, & Belisle, 2016e).

All published studies have employed the multiple baseline design (across individuals or across behavioral targets), with some incorporating additional components such as successive phases with treatment provided for different targets (e.g., Dixon et al., 2016e) for further replications, successive phases showing changes in untrained targets (e.g., Dixon, Belisle, Stanley, Daar, & Williams, 2016c), Dixon et al., 2016b or maintenance phases (e.g., Belisle, Dixon, Stanley, Munoz, & Daar, 2016a; Belisle, Rowsey, & Dixon, 2016b; Dixon et al., 2016e). Of 10 treatment efficacy evaluations, seven involved child participants, and three involved adolescent participants. Therefore, the supporting data have evaluated the efficacy of PEAK with children and adolescents, of the children and adolescents, all ten studies contained individuals with an autism diagnosis, four of these studies also included children intellectual or cognitive delays, and in two contained children with speech impairment, emotional disabilities, and/or attention deficit hyperactivity disorders. In terms of effects, across all studies the vast majority of participants were noted as successes by authors of each study, thus meeting mastery criteria as defined in each individual case. The existing single-subject research has targeted a total of 44 unique skills from the PEAK training modules (these skills were either directly trained or tested for untrained, but expected,

Table 3 Summary of skills taught and effect size analyses across single-subject experimental evaluations of PEAK

| Single-Case Experimental Design | | |
PEAK Module	Target Skill (Train/Tested)	Percentage non-overlapping data (PND)
PEAK-Direct Training	Basic autoclitics (Tr)†	100%***
	Receptive metonymical tacts (Tr)†	100%***
	Expressive planet identification (Tr)†	100%***
	Guessing (Tr)†	100%***
	Expressive metonymical tacts (Tr)†	100%***
PEAK-Generalization	Creative paths (Tr) †	100%***
	Creative Paths (Te) †	98.1%***
	Autoclitic mands (Tr) †	100%***
	Autoclitic mands (Te) †	27.5%
	Distorted tacts (Tr) †	100%***
	Distorted tacts (Te) †	73.9%**
PEAK- Equivalence	Chemistry technical name – symbol (Tr)	100%***
	Chemistry symbol – common name (Tr)	95.5%***
	Chemistry technical name – common name (Te)	100%***
	Chemistry common name – technical name (Te)	100%***
	Measurement inches – yards (Tr)	100%***
	Measurement feet – yards (Tr)	100%***
	Measurement inches – feet (Te)	100%***
	Measurement feet – inches (Te)	100%***
	History name – picture (Tr)	100%***
	History name – fact (Tr)	100%***
	History name – region (Tr)	100%***
	History picture – fact (Te)	100%***
	History region – picture (Te)	100%***
	History fact – region (Te)	100%***
	Gustatory – picture (Tr) §	62.9%*
	Picture – auditory name (Tr) §	93.5%***
	Auditory name – picture (Te)	100%***
	Gustatory – picture (Te)	61.9%*
	Animal blend – animal name (Tr)	100%***
	Animal name – animal blend (Te)	93.3%***
	Category item 1 – category item 2 (Tr) †	88.9%**
	Category item 2 – category item 3 (Tr) †	90.5%***
	Category name – category item 1 (Tr) †§	97.8%***
	Category name – category item 1, 2, 3 (Te) †	95.7%***
	Categorical vocal intraverbal	83.3%**
	Shape name – shape property (Tr)	100%***
	Shape property – shape name (Te)	100%***
	Shape name – shape picture (Te)	100%***

Table 3 (continued)

| Single-Case Experimental Design | | |
PEAK Module	Target Skill (Train/Tested)	Percentage non-overlapping data (PND)
	Twister with unknown body part term (Te)	89.5%**
PEAK- Transformation	Single-reversal YOU set 1 (Tr) §	92.1%***
	Single-reversal YOU set 2 (Te) §	100%***
	Single-reversal I set 1 (Te) §	77.8%**
	Single-reversal I set 2 (Te) §	81.8%**

* = PND minimally effective (50–70), ** = PND moderately effective (70–90), *** = PND highly effective (< 90); † = PEAK Score was used as the dependent variable rather than percentage correct, which could artificially inflate PND values; § = Mixed training trial blocks and training with a condensed stimulus array were not included in calculation of PND

emergence). In effort to both summarize these skills for purposes of this review, and also to add statistical support for the traditional visual inspection of single-subject data, the percentage of non-overlapping data (PND; Scruggs, Mastropieri, & Casto, 1987) was calculated as an indicator of effect size. Higher percentages imply larger treatment effects. Table 3 shows that PND exceeded 70% in 41 of 44 instances, with a median of 100%.

Discussion

Psychometric Evaluations

Although historically behavior analysts have not engaged in much psychometric research, the concepts of reliability and validity are integral to all forms of measurement, and it makes sense to formally evaluate these dimensions of behavioral assessment whenever possible. Doing so does not simply document the operating characteristics of behavioral assessment tools for those who are already inclined to use them; it also facilitates the use of such tools by a broad range of human services professions who may be interested in conditions like autism but who trust only forms of measurement that have been psychometrically vetted. They do so for two good reasons. First, an intervention can be called "evidence based" only if socially-valued changes have been properly documented, and reliability and validity evaluations are required to verify that what a measurement tool yields is not capricious or arbitrary. Second, many insurance companies will pay only for assessments that are empirically supported (i.e., psychometrically sound).

The present review shows that PEAK modules that have been psychometrically evaluated generally perform well with respect to reliability and validity. Among assessment tools and curricula that behavior analysts have developed for addressing the challenges of autism, the PEAK system is unique in its level of psychometric support. Indeed, we are aware of no published psychometric evaluation of any other system, including the VB-MAPP (Sundberg, 2008), the ABBLS-R (Partington, 2006), *A Work*

in Progress (Leaf, McEachin, & Harsh, 1999) or *STAR* (STARautismsupport.com, 2016).

Nevertheless, the studies reviewed constitute preliminary evidence and three main limitations can be noted. First, existing psychometric evaluations of PEAK are based on smaller and more heterogeneous samples than seen in many mainstream efforts (e.g., Nordin & Nordin, 2013; Whisman & Richardson, 2015). Although statistically significant results obtained with small sample sizes may indicate strong relations, small samples also can constrain generality if they do not mirror population characteristics. Note, for instance, that all but one of the PEAK psychometric studies (cf. Malkin, Dixon, Speelman, & Luke, 2016) incorporated a sample of individuals who were recruited from the Midwest region of the United States, raising question of how well the findings might apply to populations such as that of the United States as a whole, or of the world. A second limitation is that psychometric data are available only for the first two PEAK modules, which the preponderance of evidence addressing the PEAK-DTA. Only two studies have the PEAK-GA, and no studies have assessed the psychometric properties of the PEAK-E and PEAK-T modules, and it is important to acknowledge that the findings from one PEAK module do not necessarily generalize to the other PEAK modules. A third limitation of current research on the PEAK modules is the lack of integrity data and IOR. While several studies reported IOR data, these assessments were conducted by trained graduate students, so it is not known how reliably PEAK assessments might be conducted in the field by more typically trained clinicians. Field psychometric studies therefore are needed.

Investigation into normative performance on the PEAK modules marks a first attempt at determining a reference point for how the current functional level of individuals compares to that of typically developing peers. These findings constitute an early step towards understanding how various skills within the first two PEAK modules emerge in a typically developing population, and therefore allow for a preliminary assessment of which skills individuals with autism and other developmental disabilities should be expected to exhibit based on their typically developing peers of the same age. In addition, these studies may suggest the ages at which typically developing individuals would likely benefit from instruction with PEAK modules. These findings represent an early step toward understanding the relationships between the various skills targeted in the modules and may help to identify the order in which it is appropriate to train these skills. In addition, the grouping of these skills into underlying components may aid in promoting understanding in non-behavior analytic implementers of how to identify and target related skills. Parallel studies are needed for the PEAK-E and PEAK-T modules.

Treatment Outcome Research

Given that the PEAK modules focus on behavioral technologies to increase academic and verbal skills, it is little surprise that a large proportion of research spawned from the PEAK curricula utilizes single-subject experimental designs. Single-subject experiments have documented the effectiveness of PEAK in several ways. First and foremost, PEAK has been shown to establish a variety of advanced verbal operants, mostly found within the PEAK-DT module, and more advanced verbal skills, mostly found within the PEAK-E module. This finding shows that, at least where it has been evaluated, PEAK is effective in creating the effects it was designed to create. Second, effectiveness has been documented mostly in persons with autism, who tend to show particular

deficits in the relevant domains of functioning. This finding supports the proposition that PEAK is worth considering as an alternative to autism intervention packages that have been available for longer and thus may be better known. Some of these packages have passionate adherents but little in the way of empirical support. Third, some studies have also included individuals with problems other than autism. This finding suggests that PEAK's utility may not be limited to one clinical population. Indeed, a strength of the growing body of PEAK research literature is the establishment of age-appropriate norms that can be used to identify specific skill deficits, even in individuals who lack any clinical diagnosis. In principle, PEAK-based interventions are suitable for any individual for whom the targeted skills are appropriate, although much more research is needed to evaluate effectiveness in a variety of populations. One obvious opportunity is to target adults with disabilities. Although the PEAK system was designed for children with autism, learning can occur at any age (Wong et al., 2015), and hundreds of thousands of individuals with autism and other intellectual disabilities reached adulthood before modern autism interventions were introduced. Pilot data from our research team suggests that adults may benefit from PEAK-based instruction, but systematic studies are needed to explore the magnitude of possible change.

Our survey of the effectiveness research revealed several limitations that must be addressed in future studies. First and foremost, the preponderance of the research we reviewed is from the same group of authors, some of whom had a role in developing PEAK. This naturally raises questions about author objectivity. As a frame of reference, for instance, consider the widespread belief that research on commercial products that is funded by industry sources is more likely to find benefits than independently-funded research (e.g., Nestle & Pollan, 2013). Systematic literature reviews show, however, that the expected bias is not always present (e.g., Chartres, Fabbri, & Bero, 2016). Although readers of PEAK research that our group has authored may draw their own judgments about the quality of the methods we employed, an even better solution to this concern is for different teams to replicate and extend our work. This would increase confidence in the observed effects.

Second, the multiple baseline designs employed in single subject PEAK research were not always ideally implemented. Multiple baselines logic suggests implementation of the intervention across legs in a staggered fashion, so as to demonstrate that effects occur uniquely upon introduction of an independent variable. However, in several of the PEAK studies implementation was staggered across legs by one or a very few sessions (e.g., Dixon et al., 2016e). Although such baseline to treatment phase shifts always yielded positive changes of the dependent variable, because implementation on subsequent legs of the design occurred before the initial change was clearly stabilized, causal interpretations of resulting data may be weakened. While these issues are the exception rather than the rule in current PEAK research, they remain a limitation nonetheless.

Third, the skills for which effectiveness evidence is available constitute only a subset of the range of skills that the PEAK system addresses. Specifically, the system incorporates 736 unique programs, only a minority of which has been targeted in effectiveness research. Although the behavioral targets and basic procedures are "evidence inspired" (i.e., derived from prior empirical research), many of the program topographies have rarely been studied, particularly in practice settings. For example, whereas the training of basic tact and mand repertoires has been addressed in research unrelated to PEAK, other targets like superstitious mands, self-editing responses, and

metonymical tacts have not been systematically studied. Similarly, PEAK targets many different types of relational frames, such as frames of distinction (Ming & Stewart, 2017), which have not received much attention in autism research. Effectiveness research on these components of PEAK is especially needed, and would serve not only to evaluate PEAK but also to move autism intervention research in potentially important new directions.

Fourth, as Smith (2013) has noted, a single procedure is often not enough to solve most client problems; rather, solutions may require a package of interventions. Although PEAK represents such a package, research has evaluated effectiveness of only selected components of it in isolation. Currently unknown is how the entire PEAK package might affect overall development. Before research addressing this question can be undertaken, it is also important to determine the relative contributions of different components of the PEAK package. For instance, are all four modules of the system are necessary to produce meaningful and measurable repertoire changes at the global level? Because different portions of PEAK are inspired by the Skinnerian verbal operant approach (Skinner, 1957), Relational Frame Theory (Hayes et al., 2001) and stimulus equivalence (Sidman, 1971; Sidman & Tailby, 1982), research on the relative contributions of the various PEAK modules may well also have bearing on theoretical discussions about the relative merits and possible syntheses of those foundational principles (Barnes-Holmes et al., 2000; Murphy, Barnes-Holmes, & Barnes-Holmes, 2005; Rehfeldt & Barnes-Holmes, 2009).

Also, how do the various components of PEAK interact? And are all learning programs in PEAK really necessary? Currently little is known about how success on an individual program will affect responding on other programs. Research is needed to explore how certain skills may serve as the foundation for splintering novel skills (i.e., *behavioral cusps*; Rosales-Ruiz & Baer, 1997; Hixson, 2004). With clinical time always limited, motivation exists to determine whether it is productive to skip or combine steps, thereby potentially eliminating the painstaking task of addressing 736 separate language and cognitive goals for each child exposed to the curriculum.

Finally, as noted in our Introduction, mainstream audiences and policy makers are most likely to resonate to effectiveness evidence in the form of group-based experiments like RCTs (Guyatt, Sackett, Taylor, Ghong, Roberts, & Pugsley, 1986; United States Preventive Services Task Force, 2017), such as the replication of Lovaas (1987) by Cohen, Amerine-Dickens, and Smith (2006). To date, however, PEAK has been examined in only one such study (McKeel et al., 2015a) that had a limited and homogeneous sample and addressed very limited aspects of the PEAK system. More studies of this type, incorporating larger and more heterogeneous samples, and evaluating all components of PEAK, will be needed before the overall PEAK system can be considered fully "evidence based". Unfortunately, behavior analysts tend to shy away from group-comparison experimentation (e.g., Johnston, 1988; for a review see Smith, 2013), which can serve to their detriment when seeking to inform and persuade audiences who are not behavior analysts. There is a need for PEAK RCTs with larger samples, addressing more components of the system, and conducted for greater amounts of time than were the case for the McKeel et al. (2015a, b, c) study. Group-comparison studies may be especially useful in evaluating the global impact on development of a complex treatment package like PEAK, but many other investigatory possibilities can be identified, including studies to map dose-response relationships

(e.g., groups differentiated by number of hours per week of treatment, as per Lovaas, 1987), and active-control designs pitting PEAK against alternative programs like STARS (STARautismsupport.com, 2016) or Skills (SkillsforAutism.com, 2016). It will be important to supplement PEAK-specific assessments of change with global outcome measures such as IQ, behavior problem frequency reductions, vocabulary tests, and/or achievement testing.

Three Additional Directions for Future Research

Dissemination issues PEAK's systematic, manualized format was initially chosen with front-line service-delivery staff in mind. The assumption was that even paraprofessionals might be able to implement PEAK assessments and interventions. However, most research on PEAK has been conducted with unusually well trained implementers such as students in behavior analysis graduate programs, and in only one study were less-skilled front-line staff employed as implementers. In this study, the manual alone did not promote ideal implementation, and only after behavioral skills training were staff able to implement PEAK with integrity (Belisle et al., 2016b). Research is needed to determine what supports or modifications of the PEAK manuals may be needed to allow the program to be properly implemented in the field by relatively unskilled workers.

PEAK applications with neurotypical children Preliminary psychometric work suggests that PEAK targets not only basic and remedial skills but also more sophisticated language and cognitive skills that are appropriate to individuals without autism. It is common for neurotypical children to have gaps in these skills, which raises the possibility of PEAK assessments being used to identify them, and PEAK interventions being used to address them. Encouragement may be taken from recent research within the area of derived relational responding which showed significant improvements in intelligence test scores following brief exposure to a series of computerized discrete trials of logic deduction with mostly arbitrary stimuli (Cassidy, Roche, & Hayes, 2011). Such effects make logical sense when the derived relations that have interested behavior analysts are compared to the types of skills evaluated in intellectual assessments and promoted in commercial products such as RaiseyourIQ (Thirus, 2016). We therefore see opportunities to employ PEAK technology with neurotypical individuals for purposes of both remediation, as in after school tutoring programs such as FIT Learning (FitLearners.com, 2016) or Sylvan (Sylvianlearning.com, 2016), and for prevention, as in school-based Response to Intervention programs.

Neuroimaging as a measure of PEAK interventions effects In this review we have focused on practical goals of intervention but there is also value in considering theoretical questions that arise in conjunction with intensive intervention programs like PEAK. For example, there is no greater underexplored area of cross-disciplinary collaboration for behavior analysis than neuroscience (Ortu & Vaidya, 2016; Zilio, 2016). Although a few studies have been done examining the neurological correlates of various language training tasks such as stimulus equivalence (Dickins, Singh, Roberts, Burns, Downes, Jimmieson, & Bentall, 2001; Schlund, Cataldo, & Hoehn-Saric, 2008), to our knowledge none have done so using children with autism. Given mainstream interest in the neurological deficits of these children, and the brain's vast potential for

neural plasticity, the time seems ripe to evaluate how behavioral interventions may alter and interact with neurological markers. An advantage of neuroscience-focused research is that it places behavioral intervention in a context that mainstream audiences, including policy makers and funding agencies, are predisposed to appreciate, and thereby promotes interdisciplinary discussion. We believe that such neurological changes will be shown following PEAK or other ABA approaches at language training, and when documented will open a multitude of opportunities for research funding, mainstream media validation, and the suppression of unsupported treatments that often are touted as changers of such physiology.

Conclusion

PEAK is the first program to integrate Skinnerian and post-Skinnerian (RFT) formulations of verbal behavior into its assessment and treatment approach for individuals with autism. As we hope is illustrated by this internal research review, however, PEAK was always intended to be considered within the values and procedures of the evidence-based practice movement. The evidence summarized here suggests that valuable steps have been taken toward developing a rigorous technology for delivering language interventions based on applied behavior analysis to persons with autism and related conditions. Ultimately, however, the task of determining what works (and does not work) for individuals with autism is too large to be met by any one research team. Our goal in preparing the present review was twofold: to recruit additional investigators to examine the PEAK system, and to share a framework (psychometric evaluation plus effectiveness assessments of various types) that can be employed in examining any intervention system. Regardless of what future research reveals about PEAK (or any other intervention package), the process of investigation serves the general goal of identifying the best possible interventions for people who have a pressing need for them.

References

Studies that were included in the review of the PEAK system summarized in Tables 1, 2, and 3 are indicated with asterisks****

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: American Psychiatric Publishing.

Arntzen, E., & Almås, I. K. (2002). Effects of mand-tact versus tact-only training on the acquisition of tacts. *Journal of Applied Behavior Analysis, 35*, 419–422. doi:10.1901/jaba.2002.35-419.

Barnes-Holmes, D., Barnes-Holmes, Y., & Cullinan, V. (2000). Relational frame theory and skinner's verbal behavior: a possible synthesis. *The Behavior Analyst, 23*, 69–84.

** Belisle, J., Dixon, M. R., Stanley, C., Munoz, B., & Daar, J. H. (2016a). Teaching foundational perspective-taking skills to children with autism using the PEAK-T curriculum: single-reversal "I-You" deictic frames. *Journal of Applied Behavior Analysis, 49*, 965–969. doi:10.1002/jaba.324.

** Belisle, J., Rowsey, K. E., & Dixon, M. R. (2016b). The use of in-situ behavioral skills training to improve staff implementation of the PEAK relational training system. *Journal of Organizational Behavior Management, 36*, 71–79. doi:10.1080/01608061.2016.1152210.

Bellg, A. J., Borrelli, B., Resnick, B., Hecht, J., Minicucci, D. S., Ory, M., & Czajkowski, S. (2004). Enhancing treatment fidelity in health behavior change studies: best practices and recommendations from the NIH Behavior Change Consortium. *Health Psychology, 23*(5), 443.

Cassidy, S., Roche, B., & Hayes, S. C. (2011). A relational frame training intervention to raise intelligence quotients: a pilot study. *The Psychological Record, 61*, 173–198.

Chartres, N., Fabbri, A., & Bero, L. A. (2016). Association of industry sponsorship with outcomes of nutrition studies: a systematic review and meta-analysis. *JAMA Internal Medicine, 167*, 1769–1777. doi:10.1001/jamainternmed.2016.6721.

Cohen, H., Amerine-Dickens, M., & Smith, T. (2006). Early intensive behavioral treatment: replication of the UCLA model in a community setting. *Journal of Developmental & Behavioral Pediatrics, 27*, S145–S155.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Erlbaum.

Cooper, J. O., Heron, T. E., & Heward, W. L. (2007). *Applied behavior analysis* (2nd ed.). Upper Saddle River, NJ: Pearson Education.

Cusumano, D. L. (2007). Is it working?: an overview of curriculum based measurement and its uses for assessing instructional, intervention, or program effectiveness. *The Behavior Analyst Today, 8*, 24–34.

Dickins, D. W., Singh, K. D., Roberts, N., Burns, P., Downes, J. J., Jimmieson, P., & Bentall, R. P. (2001). An fMRI study of stimulus equivalence. *Neuroreport, 12*, 405–411.

Dixon, M. R. (2014a). *The PEAK relational training system: direct training module*. Carbondale: Shawnee Scientific Press.

Dixon, M. R. (2014b). *The PEAK relational training system: generalization module*. Carbondale: Shawnee Scientific Press.

Dixon, M. R. (2015). *The PEAK relational training system: equivalence module*. Carbondale: Shawnee Scientific Press.

Dixon, M. R. (2016). *The PEAK relational training system: transformation module*. Carbondale: Shawnee Scientific Press.

Dixon, M. R., Belisle, J., Munoz, B. E., Stanley, C. R., & Rowsey, K. E. (In press). Teaching metaphorical extensions of private events through rival-model observation to children with autism. *Journal of Applied Behavior Analysis.*

** Dixon, M. R., Belisle, J., Rowsey, K. E., Speelman, R., Stanley, C., Kime, D. (2016a). Evaluating emergent naming relations through representational drawing in individuals with developmental disabilities using the PEAK-E curriculum. *Behavior Analysis: Research and Practice*, 1–6. doi:10.1037/bar0000055.

** Dixon, M. R., Belisle, J., Stanley, C., Munoz, B. E., & Speelman, R. (2017a). Establishing derived coordinated symmetrical and transitive gustatory-visual-auditory relations in children with disabilities. *Journal of Contextual Behavioral Science*. doi:10.1016/j.jcbs.2016.11.001.

** Dixon, M. R., Belisle, J., Stanley, C., Rowsey, K. E., & Daar, J. (2015). Toward a behavior analysis of complex language for children with autism: evaluating the relationship between PEAK and the VB-MAPP. *Journal of Developmental and Physical Disabilities, 27*, 223–233. doi:10.1007/s10882-014-9410-4.

** Dixon, M. R., Belisle, J., Stanley, C., Speelman, R., Rowsey, K. E., & Daar, J. H. (2016b). Establishing derived categorical responding in children with disabilities using the PEAK-E curriculum. *Journal of Applied Behavior Analysis, 50*, 134–145. doi:10.1002/jaba.355.

** Dixon, M. R., Belisle, J., Stanley, C., Daar, J. H., & Williams, L. A. (2016c). Derived equivalence relations of geometry skills in students with autism: an application of the PEAK-E curriculum. *Analysis of Verbal Behavior, 32*, 38–45. doi:10.1007/s40616-016-0051-9.

** Dixon, M. R., Belisle, J., Whiting, S. W., & Rowsey, K. E. (2014a). Normative sample of the PEAK Relational Training System: direct training module and subsequent comparisons to individuals with autism. *Research in Autism Spectrum Disorders, 8*, 1597–1606. doi:10.1016/j.rasd.2014.07.020.

** Dixon, M. R., Carman, J., Tyler, P. A., Whiting, S. W., Enoch, M. R., & Daar, J. H. (2014b). PEAK relational training system for children with autism and developmental disabilities: correlations with

Peabody picture vocabulary test and assessment reliability. *Journal of Developmental and Physical Disabilities, 26*, 603–614. doi:10.1007/s10882-014-9384-2.

** Dixon, M. R., Peach, J., Daar, J. H., & Penrod, C. (2017). Teaching complex verbal operants to children with autism and establishing generalization using the peak curriculum. *Journal of Applied Behavior Analysis, 50*(2), 317–331.

*Dixon, M. R., Rowsey, K. E., Gunnarsson, K. F., Belisle, J., Stanley, C. R., & Daar, J. H. (2017). Normative sample of the peak relational training system: generalization module with comparison to individuals with autism. *Journal of Behavioral Education, 26*(1), 101–122.

Dixon, M. R., Small, S. L., & Rosales, R. (2007). Extended analysis of empirical citations with skinner's verbal behavior: 1984–2004. *Behavior Analyst, 30*, 197.

** Dixon, M. R., Speelman, R., Rowsey, K. E., & Belisle, J. (2016e). Derived rule-following and transformation of stimulus functions in a children's game: an application of PEAK-E with children with developmental disabilities. *Journal of Contextual Behavioral Science, 5*, 186–192. doi:10.1016/j.jcbs.2016.05.002.

** Dixon, M. R., Stanley, C., Belisle, J., & Rowsey, K. E. (2016f). The test-retest and interrater reliability of the promoting the emergence of advanced knowledge-direct training assessment for use with individuals with autism and related disabilities. *Behavior Analysis: Research and Practice, 16*, 34–40. doi:10.1037/bar0000027.

** Dixon, M. R., Whiting, S., Rowsey, K. E., & Belisle, J. (2014c). Assessing the relationship between intelligence and the PEAK relational training system. *Research in Autism Spectrum Disorders, 8*, 1208–1213. doi:10.1016/j.rasd.2014.05.005.

Dunn, D. M., & Dunn, L. M. (2007). *Peabody picture vocabulary test*: Manual. Pearson.

Dymond, S., O Hora, D., Whelan, R., & O Donovan, A. (2006). Citation analysis of skinner's verbal behavior: 1984–2004. *Behavior Analyst, 29*, 75.

Eldevik, S., Hastings, R. P., Hughes, J. C., Jahr, E., Eikeseth, S., & Cross, S. (2009). Meta-analysis of early intensive behavioral intervention for children with autism. *Journal of Clinical Child & Adolescent Psychology, 38*, 439–450. doi:10.1080/15374410902851739.

Eldevik, S., Hastings, R. P., Hughes, J. C., Jahr, E., Eikeseth, S., & Cross, S. (2010). Using participant data to extend the evidence base for intensive behavioral intervention for children with autism. *American Journal on Intellectual and Developmental Disabilities, 115*, 381–405. doi:10.1352/1944-7558-115.5.381.

Fisher, W. W., & Zangrillo, A. N. (2015). Applied behavior analytic assessment and treatment of autism spectrum disorder. In H. S. Roane, J. E. Ringdahl, T. S. Falcomata, H. S. Roane, J. E. Ringdahl, & T. S. Falcomata (Eds.), *Clinical and organizational applications of applied behavior analysis* (pp. 19–45). San Diego: Elsevier Academic Press. doi:10.1016/B978-0-12-420249-8.00002-2.

FitLearners.com (2016). *The fit learning model: informed by science, loved by learners*. Retrieved from http://fitlearners.com/.

Furniss, F. (2009). Assessment methods. In J. L. Matson (Ed.), Applied behavior analysis for children with autism spectrum disorders (pp. 33–66). New York: Springer.

Gould, E., Dixon, D. R., Najdowski, A. C., Smith, M. N., & Tarbox, J. (2011). A review of assessments for determining the content of early intensive behavioral intervention programs for autism spectrum disorders. *Research in Autism Spectrum Disorders, 5*(3), 990–1002.

Guyatt, G., Sackett, D., Taylor, D. W., Ghong, J., Roberts, R., & Pugsley, S. (1986). Determining optimal therapy—randomized trials in individual patients. *New England Journal of Medicine, 314*, 889–892. doi:10.1056/NEJM198604033141406.

Hall, G., & Sundberg, M. L. (1987). Teaching mands by manipulating conditioned establishing operations. *The Analysis of Verbal Behavior, 5*, 41–53.

Hayes, S. C., Barnes-Holmes, D., & Roche, B. (2001). *Relational frame theory a post-skinnerian account of human language and cognition*. New York: Kluwer Academic/Plenum Publishers.

Hixson, M. D. (2004). Behavioral cusps, basic behavioral repertoires, and cumulative-hierarchical learning. *The Psychological Record, 54*, 387–403.

Horne, P. J., & Lowe, C. F. (1996). On the origins of naming and other symbolic behavior. *Journal of the Experimental Analysis of Behavior, 65*, 185–241. doi:10.1901/jeab.1996.65-185.

Howard, H. A., Ladew, P., & Pollack, E. G. (2009). *The national autism center's national standards project "findings and conclusions"*. Randolph, MA: National Autism Center.

Howlin, P., Magiati, I., & Charman, T. (2009). Systematic review of early intensive behavioral interventions for children with autism. *American Journal on Intellectual and Developmental Disabilities, 114*, 23–41. doi:10.1352/2009.114:23-41.

Ingvarsson, E. T., & Hollobaugh, T. (2011). A comparison of prompting tactics to establish intraverbals in children with autism. *Journal of Applied Behavior Analysis, 44*, 659–664. doi:10.1901/jaba.2011.44-659.

Jacobson, J. W., Foxx, R. M., & Mulick, J. A. (Eds.). (2005). *Controversial therapies for developmental disabilities: fad, fashion, and science in professional practice*. Mahwah, NJ: Erlbaum.

Johnson, G., Kohler, K., & Ross, D. (2017). Contributions of skinner's theory of verbal behaviour to language interventions for children with autism spectrum disorders. *Early Child Development and Care, 187*(3), 436–446. doi:10.1080/03004430.2016.1236255.

Johnston, J. M. (1988). Strategic and tactical limits of comparison studies. *The Behavior Analyst, 11*, 1–9.

Leaf, R., McEachin, J., & Harsh, J. D. (1999). *A work in progress: behavior management strategies and a curriculum for intensive behavioral treatment of autism*. New York: Autism Partnership.

Lovaas, O. I. (1981). *The ME book: teaching developmentally disabled children*. Austin: Pro-Ed Inc..

Lovaas, O. I. (1987). Behavioral treatment and normal educational and intellectual functioning in young autistic children. *Journal of Consulting and Clinical Psychology, 55*, 3–9. doi:10.1037/0022-006X.55.1.3.

Lovaas, O. I. (1993). The development of a treatment-research project for developmentally disabled and autistic children. *Journal of Applied Behavior Analysis, 26*, 617–630. doi:10.1901/jaba.1993.26-617.

Love, J. R., Carr, J. E., Almason, S. M., & Petursdottir, A. I. (2009). Early and intensive behavioral intervention for autism: a survey of clinical practices. *Research in Autism Spectrum Disorders, 3*, 421–428. doi:10.1016/j.rasd.2008.08.008.

** Malkin, A., Dixon, M. R., Speelman, R., & Luke, N. (2016). Evaluating the relationship between the PEAK relational training system—direct training module, assessment of basic language and learning skills—revised, and the vineland adaptive behavior scales—II. *Journal of Developmental and Physical Disabilities*. doi:10.1007/s10882-016-9527-8.

Martin, N. A., & Brownell, R. (2011a). *Expressive one-word picture vocabulary test 4*. Academic Therapy Publications.

Martin, N., & Brownell, R. (2011b). *Receptive one-word picture vocabulary test*. ATP Assessmenta, a division of Academic Therapy Publications.

McEachin, J. J., Smith, T., & Lovaas, O. I. (1993). Long-term outcome for children with autism who received early intensive behavioral treatment. *American Journal of Mental Retardation, 97*, 359–372.

** McKeel, A. N., Dixon, M. R., Daar, J. H., Rowsey, K. E., & Szekely, S. (2015a). The efficacy of the PEAK relational training system using a randomized controlled trial of children with autism. *Journal of Behavioral Education, 24*, 230–241. doi:10.1007/s10864-015-9219-y.

** McKeel, A. N., Rowsey, K. E., Belisle, J., Dixon, M. R., & Szekely, S. (2015b). Teaching complex verbal operants with the PEAK relational training system. *Behavior Analysis in Practice, 8*, 241–244. doi:10.1007/s40617-015-0067-y.

** McKeel, A. N., Rowsey, K. E., Dixon, M. R., & Daar, J. (2015c). Correlation between PEAK relational training system and one-word picture vocabulary tests. *Research in Autism Spectrum Disorders, 12*, 34–39. doi:10.1016/j.rasd.2014.12.007.

Ming, S., & Stewart, I. (2017). When things are not the same: a review of research into relations of difference. *Journal of Applied Behavior Analysis*.

Murphy, C., Barnes-Holmes, D., & Barnes-Holmes, Y. (2005). Derived manding in children with autism: synthesizing skinner's verbal behavior with relational frame theory. *Journal of Applied Behavior Analysis, 38*, 445–462. doi:10.1901/jaba.2005.97-04.

Myers, S. M., & Johnson, C. P. (2007). Management of children with autism spectrum disorders. *Pediatrics, 120*, 1162–1182. doi:10.1542/peds.2007-2362.

National Institutes of Health. (2017). *What does it mean for something to be evidence-based?* Retrieved from https://prevention.nih.gov/resources-for-researchers/dissemination-and-implementation-resources/evidence-based-programs-practices.

National Research Council. (2001). *Educating children with autism*. Washington: National Academy Press.

Nestle, M., & Pollan, M. (2013). *Food politics: how the food industry influences nutrition and health*. Berkeley: University of California Press.

Nordin, M., & Nordin, S. (2013). Psychometric evaluation and normative data of the Swedish version of the 10-item perceived stress scale. *Scandinavian Journal of Psychology, 54*, 502–507. doi:10.1111/sjop.12071.

Odom, S. L., Boyd, B. A., Hall, L. J., & Hume, K. (2010a). Evaluation of comprehensive treatment models for individuals with autism spectrum disorders. *Journal of Autism and Developmental Disorders, 40*, 425–436. doi:10.1007/s10803-009-0825-1.

Odom, S. L., Collet-Klinenberg, L., Rogers, S. J., & Hatton, D. D. (2010b). Evidence-based practices in interventions for children and youth with autism spectrum disorders. *Preventing School Failure, 54*, 275–282. doi:10.1080/10459881003785506.

Ortu, D., & Vaidya, M. (2016). The challenges of integrating behavioral and neural data: bridging and breaking boundaries across levels of analysis. *The Behavior Analyst. Advance online publication. doi.* doi:10.1007/s40614-016-0074-5.

Partington, J. (2006). *Assessment of basic language and learning skills-revised (The ABLLS-R)*. Pleasant Hill: Behavior Analysts.

Peters-Scheffer, N., Didden, R., Korzilius, H., & Sturmey, P. (2011). A meta-analytic study on the effectiveness of comprehensive ABA-based early intervention programs for children with autism spectrum disorders. *Research in Autism Spectrum Disorders, 5*, 60–69. doi:10.1016/j.rasd.2010.03.011.

Reed, D. D., & Luiselli, J. K. (2016). Promoting the emergence of advanced knowledge: a review of peak relational training system: direct training module by Mark R Dixon. *Journal of Applied Behavior Analysis, 9*, 205–211. doi:10.1002/jaba.281.

Rehfeldt, R. A., & Barnes-Holmes, Y. (Eds.). (2009). *Derived relational responding: applications for learners with autism and other developmental disabilities: a progressive guide to change*. Oakland: New Harbinger Publications.

Reichow, B., Barton, E. E., Boyd, B. A., & Hume, K. (2012). Early intensive behavioral intervention (EIBI) for young children with autism spectrum disorders (ASD). *Cochrane Database of Systematic Reviews*. CD009260. doi:10.1002/14651858.CD009260.pub2.

Rosales-Ruiz, J., & Baer, D. M. (1997). Behavioral cusps: a developmental and pragmatic concept for behavior analysis. *Journal of Applied Behavior Analysis, 30*, 533–544. doi:10.1901/jaba.1997.30-533.

**Rowsey, K. E., Belisle, J., Stanley, C. R., Daar, J. H., & Dixon, M. R. (2017). Principal component analysis of the peak generalization module. *Journal of Developmental and Physical Disabilities, 29*(3), 489–501.

** Rowsey, K. E., Belisle, J., & Dixon, M. R. (2014). Principal component analysis of the PEAK relational training system. *Journal of Developmental and Physical Disabilities, 27*, 15–23. doi:10.1007/s10882-014-9398-9.

Schlund, M. W., Cataldo, M. F., & Hoehn-Saric, R. (2008). Neural correlates of derived relational responding on tests of stimulus equivalence. *Behavioral and Brain Functions, 4*, 6.

Scruggs, T. E., Mastropieri, M. A., & Casto, G. (1987). The quantitative synthesis of single-subject research: methodology and validation. *Remedial and Special Education, 8*, 24–33.

Sidman, M. (1971). Reading and auditory-visual equivalences. *Journal of Speech and Hearing Research, 14*, 5–13.

Sidman, M., & Tailby, W. (1982). Conditional discrimination vs. matching to sample: an expansion of the testing paradigm. *Journal of the Experimental Analysis of Behavior, 37*, 5–22. doi:10.1901/jeab.1982.37-5.

SkillsforAutism.com (2016). *Research of skills and ABA*. Retrieved from https://www.skillsforautism.com/Research.

Skinner, B. F. (1957). *Verbal behavior*. BF Skinner Foundation.

Smith, T. (2013). What is evidence-based behavior analysis? *The Behavior Analyst, 36*, 7–33.

Smith, T., & Iadarola, S. (2015). Evidence base update for autism spectrum disorder. *Journal of Clinical Child & Adolescent Psychology, 44*, 897–922.

Snow, C. E., Burns, M. S., & Griffin, P. (1998). *Preventing reading difficulties in young children*. Washington: National Academy Press.

Spreckley, M., & Boyd, R. (2009). Efficacy of applied behavioral intervention in preschool children with autism for improving cognitive, language, and adaptive behavior: a systematic review and meta-analysis. *The Journal of Pediatrics, 154*, 338–344. doi:10.1016/j.jpeds.2008.09.012.

** Stanley, C., Belisle, J., Dixon M. R. (In Press). Equivalence-based instruction of academic skills: application to adolescents with autism. *Journal of Applied Behavior Analysis.*

STARautismsupport.com (2016). *STAR program.* Retrieved from http://starautismsupport.com/curriculum/star-program.

Sundberg, M. L. (2008). *VB-MAPP: verbal behavior milestones assessment and placement program*. Concord: AVB Press.

Sundberg, M. L., & Partington, J. W. (1998). *Teaching language to children with autism or other developmental disabilities*. Pleasant Hill: Behavior Analysts, Inc..

Sylvianlearning.com (2016). *Sylvian learning*. Retrieved from https://www.sylvanlearning.com/.

Thirus, J. (2016). Relational frame theory, mathematical and logical skills: a multiple exemplar training intervention to enhance intellectual performance. *International Journal of Psychology and Psychotherapy.*

United States Preventive Services Task Force. (2017). *Grade definitions. 2007.* Retrieved from http://www.uspreventiveservicestaskforce.org/uspstf/grades.htm.

U.S. Public Health Service. (1999). *Mental health: a report of the Surgeon General (chap. 3, section on autism).* Retrieved from http://www.surgeongeneral.gov/library/mentalhealth/chapter3/sec6.html#autism.

Warren, Z., McPheeters, M. L., Sathe, N., Foss-Feig, J. H., Glasser, A., & Veenstra-VanderWeele, J. (2011). A systematic review of early intensive intervention for autism spectrum disorders. *Pediatrics, 127*, e1303–e1311. doi:10.1542/peds.2011-0426.

Whisman, M. A., & Richardson, E. D. (2015). Normative data on the beck depression inventory–second edition (BDI-II) in college students. *Journal of Clinical Psychology, 71*, 898–907. doi:10.1002/jclp.22188.

Wong, C., Odom, S. L., Hume, K. A., Cox, A. W., Fettig, A., Kucharczyk, S., et al. (2015). Evidence-based practices for children, youth, and young adults with autism spectrum disorder: a comprehensive review. *Journal of Autism and Developmental Disorders, 45*, 1951–1966. doi:10.1007/s10803-014-2351-z.

Zilio, D. (2016). Who, what, and when: Skinner's critiques of neuroscience and his main targets. *The Behavior Analyst, 39*, 197–218. doi:10.1007/s40614-016-0053-x.