Missouri State
UNIVERSITY

BearWorks

Spring 2020

Evaluating the Construct Validity of the Peak Comprehensive Assessment: Measuring Language and Cognition

Nicole Rae McDonald
Missouri State University, Choate88@live.missouristate.edu

As with any intellectual project, the content and views expressed in this thesis may be considered objectionable by some readers. However, this student-scholar's work has been judged to have academic value by the student's thesis committee members trained in the discipline. The content and views expressed in this thesis are those of the student-scholar and are not endorsed by Missouri State University, its Graduate College, or its employees.

EVALUATING THE CONSTRUCT VALIDITY OF THE PEAK COMPREHENSIVE

ASSESSMENT: MEASURING LANGUAGE AND COGNITION

A Master's Thesis

Presented to

The Graduate College of

Missouri State University

In Partial Fulfillment

Of the Requirements for the Degree

Master of Science, Applied Behavior Analysis

By

Nicole McDonald

May 2020

EVALUATING THE CONSTRUCT VALIDITY OF THE PEAK COMPREHENSIVE

ASSESSMENT: MEASURING LANGUAGE AND COGNITION

Psychology

Missouri State University, May 2020

Master of Science, Applied Behavior Analysis

Nicole McDonald

ABSTRACT

The purpose of the current study was to conduct a pilot investigation of the internal construct validity of the four modules of the PEAK Comprehensive Assessment (PCA). The PCA has been developed through robust research over the past five years (Dixon et al. 2017) and is designed to evaluate language and cognitive skills of individuals with developmental disabilities, including neurodevelopmental disorders such as Autism Spectrum Disorder. Although the PCA contains four modules exemplifying four distinct learning processes (Direct Training, Generalization, Equivalence, and Relational Learning), these four processes may represent one singular learning construct, described loosely as "executive functioning" or "cognitive ability" in domains outside of applied behavior analysis. Within applied behavior analytic models, the common feature among these modules is that all are operant learning accounts. I evaluated the construct validity of the PCA using a principle component analysis in a sample of 55 participants with disabilities collected from multiple clinical sites throughout the United States. Results supported a one-factor model, suggesting that although scores in each module may differentially direct programming decisions, they are representative of a single underlying construct. Implications of these results are discussed.

EVALUATING THE CONSTRUCT VALIDITY OF THE PEAK COMPREHENSIVE

ASSESSMENT: MEASURING LANGUAGE AND COGNITION

By

Nicole McDonald

A Master's Thesis
Submitted to the Graduate College
Of Missouri State University
In Partial Fulfillment of the Requirements
For the Degree of Master of Science, Applied Behavior Analysis

May 2020

Approved:

Jordan Belisle, Ph.D., Thesis Committee Chair

Dana Paliliunas, Ph.D., Committee Member

Michael Clayton, Ph.D., Committee Member

Julie Masterson, Ph.D., Dean of the Graduate College

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

Construct validity is an internal consistency measure which has long been examined by individuals within the behavioral science community. Throughout the late 1930's and early 1940's, researchers such as Peak and Cronbach (Cronbach and Meehl 1955) began to investigate the different components of constructs and validity which lead to a more comprehensive development of what we now see as construct validity theory. A construct is an idea or theory typically used by scientists to describe an unknown cluster of events or phenomenon. In the realm of autism spectrum disorder and providing education for individuals with disabilities, many of our constructs are designed around the implementation of language and cognition. Although not extensively studied by behavior analysts, understanding latent constructs that participate in behavior is important. If we can begin the identify the constructs that make up language and cognition we can use this information to further evaluate the effectiveness of tools and curriculum designed to aid individuals with disabilities. One area of recent research is in approaches to developing language and cognitive skills in children with autism (Belisle et.al 2016). By examining the different approaches to understanding the development of language such as Skinner's verbal operant theory (Skinner 1957), Sidman's Stimulus Equivalence Theory (Sidman 1971; Sidman and Tailby 1982), Relational Frame Theory (Hayes, Barnes-Holmes, and Roche 2001) which includes derived relational responding, one could determine the different ways that people learn, but do these theories represent a singular construct? Each of these model's center around an operant account of learning which in its complexity attempts to describe human language and cognition, however various research on Model Dependent Realism and Derived Relational Responding (Belisle 2020) also suggest that there is still a lot to be

learned. Approaching the 1950's, validity theorists began to recognize that validity is related to the specific purposes to which the test is designed and, "thus, can be high for some uses and low for others" (Newton and Shaw 2014)

Peak and Cronbach, purported that individual details of construct validity pertain to different measurements and the characteristics that are being measured (Slaney 2017). With the development of committees to regulate test procedures and available information regarding validity, reliability, administration, and norms, came a preliminary proposal of validity regulations outlined by the American Psychological Association (APA 1952). "The four types of validity identified in this document were predictive, status, content, and congruent" with congruent later being identified as construct validity, intended to "measure a construct arising from some theory but for which no criterion is available as a "trustworthy" measure of the attribute (state, quality, trait) in question" (APA 1952, p. 268; Slaney 2017, p. 62).

Construct Validity and Language

Since the primary development of theories regarding construct validity and internal consistency, numerous researchers have sought to develop systematic representation of test measurements and tools that can guide the practice and research of behavior analytics. A test is said to have validity if it can measure what it is intended to measure, and construct validity is a core component to measuring test results that are according to Devitt are in keeping with expectations (Devitt et al. 1998). PEAK attempts to address language and cognitive functioning using all three learning models described above. The PEAK Curriculum provides four modules of behavior analytic training which target skill deficits among children with disabilities. The PEAK modules are easy to incorporate within the classroom setting, and the PCA provides an

amiable, easy to use assessment tool to identify age-norm targets for each student. The PEAK Direct Training Module (PEAK-DT; Dixon, 2014a) is the first model of the PEAK relational training system and consists of 184 programs which use discrete trial training to teach skills ranging from basic learning skills such as modeled play and turn taking to more advanced language and learning skills such as intraverbal emotions and receptively labelling time (McKeel et al. 2015). Of the 184 curriculum items found within the Direct Training Module, 64 items are contained in the PCA. The PEAK-DT Module demonstrates the effectiveness of a packaged curriculum grounded in Applied Behavior Analysis which contains a "wide range of early learning skill programming that should be considered the foundational building blocks of language and cognition" (Dixon 2019). PEAK provides tools with packaged techniques, delivering ease of use for individuals who may have little prior experience with behavioral protocol implementation.

PEAK's Generalization Module (PEAK; Dixon 2014b) provides guidelines and properties for the training of 184 distinct skills ranging from generalized motor imitation to intraverbal interests of others. The PEAK Generalization Module demonstrates an effectiveness of teaching more complex language skills by emphasizing a slightly higher level of difficulty (Dixon 2019). The skills targeted in the PEAK-G Curriculum go beyond Skinner's basic verbal operants and focus on promoting a generalization of skills by stimulating the recognition of targets across stimuli, responses, and environments. The PEAK-Generalization module is unique in that it begins to fade out the paired reinforcement with each correct response. By incorporating a training-testing approach, the student begins to recognize when they have completed a task correctly or answered a question correctly. This sets the stage for a greater likelihood of self-

initiated praise and positive feedback for completing a task accurately. "Like the DT module, the G module is represented on the PCA with 64 items" (Dixon 2019).

The third of four PEAK modules, PEAK-Equivalence (PEAK-E; Dixon 2015) incorporates the use of equivalence technology to support the development of simple and complex categorical classes (Dixon et. al 2016). The Equivalence module is an expansion of the first two PEAK modules and is structured to provide assessment and curriculum guidelines which support the emergence of equivalence class formation using the procedures of Sidman (1971), and relational frame theory (Hayes et al. 2001). "Stimulus equivalence has provided a model for category formation and the development of human language whereby not all possible relations among stimuli and their various categories need to be directly trained" (Sidman 1994). The PEAK-E Module demonstrates the acquisition of equivalence skill targets promoting the emergence of more complex verbal behavior operants. Using the PEAK-E Curriculum to demonstrate the emergence of complex verbal behavior demonstrates "untaught relations" or "derived relations" as they are based on the explicitly taught relations (Hayes, Barnes-Holmes, and Roche 2001). The emergence of new, more complex skills, without being directly taught, is a skill that could benefit any learner, and particularly those who fall under the diagnosis of developmental disability or autism spectrum disorder. The PEAK-E curriculum incorporates teaching of skills such as reflexivity matching and advances to more complex skills such as ordering equivalent items from a menu (Dixon 2016). By incorporating stimulus equivalence into a school-based curriculum, you are allowing the student to "learn how to learn" where previous skills may have not been present. Stimulus equivalence skills can be demonstrated using a teaching strategy called Multiple Exemplar Training, which systematically provides examples of related stimuli until the student no longer needs to be taught "each symmetrical or

transitive relation" (Dixon 2016). In order to successfully complete this type of teaching, individuals must promote and arrange an environment in which such stimuli can be easily accessible and observed within the natural environment. Regarding the PCA, "24 completely novel test items designed to determine how complex a client's abilities are" were chosen to represent the Equivalence module and evaluate the relation of abstract concepts which ultimately guage learning capacity, intelligence, and a deeper understanding of social behavior (Dixon 2019).

The PEAK- Transformation Module (PEAK-T; Dixon, 2016) "provides a standardized curriculum and instruction for how to teach deictic and other relational skills (Belisle et al. 2016). Perspective taking skills, or theory of mind (Premack and Woodreff 1978), which are often absent among individuals with ASD, allow individuals to understand that the beliefs and perceptions of others may be different from their own views or observations. Relational frame theory (Hayes, Barnes-Holmes, and Roche 2001) "provides a behavioral account of perspective taking as deictic relational responding or responding relationally to events in terms of I and You, Here and There, Now and Then (Belisle et. al 2016). The PEAK-T assessment provides an intervention tool for evaluating an individual's relational abilities and includes programming that ranging from simple skills such as vocal imitation to more complex skills such as following complex directions (Belisle et.al 2016). The skills learned within the PEAK-T curriculum target concept formation and cognitive capacity and are essential behaviors that allow individuals to better interact with the world around them. "The T section of the PCA contains 192 test items broken down into an expressive and receptive sub-section" (Dixon 2019) which evaluate how an individual may understand diverse relations among stimuli including opposite, difference, comparisons, hierarchies, and perspective taking (Dixon 2019).

Evaluating Construct Validity of the PCA

With greater ease of implementation concerning the PCA, one should begin to evaluate the construct validity and internal consistency measures which target factors and relations contained within the PEAK Relational Training System, and more specifically the PEAK Comprehensive Assessment. There are currently several studies that support the convergent validity and reliability of the PEAK-DT and PEAK-G indirect assessments as well as the PEAK-E pre-assessment (Dixon et al. 2017). The purpose of this current study is to evaluate the internal validity of the PCA by examining correlations throughout the various assessments, evaluating measurements of different dimensions of the same construct, language and cognition. By examining the internal validity of the assessment, we can investigate the ways in which the PEAK Comprehensive Assessment works with a large autism sample. Many studies have evaluated the comparison of normative samples with those of individuals diagnosed with ASD, however, little research is published on using and autism sample alone. Secondly, while there are many articles demonstrating the intervention validity and convergent validity of the PEAK Relational Curriculum, little has been published on evaluating the internal constructs of language and cognition for the PCA.

In order to properly evaluate the internal consistency and construct validity of the 344 item PCA one must first understand the constructs of the PEAK Curriculum, what is being measured, and why this is important for the autism population. The PEAK curriculum was developed for the implementation and strengthening of language and cognition skills among those who have been diagnosed with deficits in this area. PEAK is an evidence-based implementation tool that has allowed individuals the opportunity to make inferences and connections by training skills which show up as deficit on the PEAK-CA. "The PEAK:

Relational Training System is a comprehensive approach to ABA Therapy, which embraces traditional verbal behavior accounts of basic language and incorporates contemporary behavior analytic strategies for promoting relational responding (a broad repertoire of learning meaning through relations between stimuli) which are responsible for our ability to understand and use abstract language" (Belisle and Dixon 2018). The internal measurements of the PEAK-CA relates directly to language and cognition posing the idea that human language (verbal behavior) derives its power from specific links to human cognition.

In addition to understanding the constructs that are being measured, one must also consider the standards of internal consistency and validity that are presented throughout the updated standards of educational and psychological testing. While previous accounts of internal consistency measures regarded 4 initial types of validity, the newest standards were developed in 1999 to address changes and the metamorphosis of the validity definition and application. "Evolutionary changes in the meaning of validity have occurred since the 1940s. The newest edition of the Standards (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education 1999) contains a dramatic shift in the definition and description of validity: the elimination of the content, criterion-related, and construct types of validity (Goodwin and Leech 2003). In a recent article that addresses the meaning of validity in the new standards of educational and psychological testing the older view of validity is replaced with a new view that "focusses on five types of validity evidence: evidence based on test content, evidence based on response processes, evidence based on internal structure, evidence based on response processes, evidence based on internal structure, evidence based on relations to other variables, and evidence based on the consequences of testing" (Goodwin and Leech 2003). The notion that validity is a complex measure and therefore requires

a variety of evidence, is welcomed within the scope of behavioral, psychological, and educational practices. Validity as described in the newest edition of standards provides the following definition and demonstration of the validation process, "validity refers to the degree to which evidence and theory support the interpretations of test scores entailed by proposed uses of tests" (Goodwin and Leech 2003). The standards continue, to include that validity is therefore one of the most valuable components to evaluating and developing psychological and educational tests. "The process of validation involves accumulating evidence to provide a sound scientific basis for the proposed score interpretations" (AERA, APA, and NCME 1999, p.9). Thus, it is the interpretation of test scores that is most significantly evaluated, not the test itself. Validity is a unitary measure meaning that it is "the degree to which all of the accumulated evidence supports the intended interpretation of test scores for the intended purposes (AERA, APA, and NCME 1999, p. 11).

The expectation of this particular measurement is that all subcomponents of the PEAK Comprehensive Assessment measures the same general construct of language and cognition. While validity can be assessed using different methods, it is important to note, in order for an assessment to be valuable to the population in which you are working, practitioners must have access to instructions for administration as well as research guiding the development and implementation of the tool being used. Many practitioners benefit from understanding the constructs being measured as well as how these are related when looking at a correlation of item-by-item responses paired with comparisons among populations, particularly among populations of Autism Spectrum Disorders. "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" (Dixon et al. 2017, p. 19).

By examining the operating characteristics of behavioral assessment tools, such as the PEAK Comprehensive Assessment (PCA), we are facilitating the use of tools to individuals who may only operate in the realm of measurements that have been psychometrically evaluated. Two core principles guiding the decision making process of intervention implementation are the documentation of socially-valued changes in which reliability and validity are required to verify that the measurement tool yields non-arbitrary results, and the regulation that many insurance companies will pay only for assessments that are empirically supported (Dixon et al. 2017). When considering behavior analytic tools developed for addressing the challenges of Autism Spectrum Disorder, The PEAK Relational Training System is unique in its level of psychometric support (Dixon et al. 2017).

Practical Implications: Language and Cognition Deficits in Children with Disabilities

Autism Spectrum Disorder, as defined by the Center for Disease Prevention and Control (CDC 2017) is "a group of developmental disabilities that can cause deficits in social, communicative, and behavioral interactions as well as restricted, repetitive patterns of behavior, interests, or activities that can persist throughout life" (CDC 2017). Two categories of behavioral characteristics among those with ASD take the form of behavioral excesses and behavioral deficits. Behavioral deficits include but are not limited to deficits in receptive language, expressive language, communicative intent, social skills, self-care skills, vocational skills, and academic skills. Behavior excesses include but are not limited to tantrums, screaming, aggression, echolalia, repetitive behaviors, and refusing to follow directions (Schuermann,

Webber, and Lang 2019). Behavioral characteristics among individuals diagnosed with autism may be portrayed through limited social interactions, self-stimulatory behavior, insistence on sameness, restricted routines, splinter skills, abnormal obsessions over particular interests, and low cognitive development which can display itself through delayed or nonverbal speech. The rise in autism prevalence is seen by many as a major area of concern. The CDC "reported the prevalence of autism, including those with high-functioning autism, at 1 in 150 children, in 2002; and it reported a prevalence of 1 in 59 children in 2018", including an astounding increase of 124% of school aged students served under the label of autism from 2005 to 2014 (CDC 2018; Department of Education 2016 a).With an overwhelming increase in autism diagnoses it is essential that behavior analysts develop tools and assessments that aid in progressive and effective programming for behavioral treatments geared toward increasing the overall efficacy and implementation of training skills that promote the social repertoire of individuals with ASD. It is also important that these tools and assessments be empirically valid and reliable through research-based support. One major component of behavior analysis that has focused on the emergence of new skills is based in relational frame theory of which teaching appropriate communicative, social and functional behavior is a primary goal. "PEAK Relational Training System (PEAK) is an assessment and curriculum package designed to promote language and cognitive skills using verbal behavior approaches along with advances in RFT, in application with individuals with disabilities" (Dixon, Whiting, Rowsey, and Belisle 2014). PEAK research has shown advancements in establishing a variety of verbal operants among individuals with ASD which concludes that where evaluated, the PEAK curriculum is efficiently creating the effects it was designed to create (Dixon et al. 2017). Among the relevant domains of functioning, individuals with ASD have shown markedly increased performance among

foundational learning skills, perceptual learning skills, verbal comprehension skills, verbal reasoning, memory and math skills, basic social skills, symmetry relational skills, non-arbitrary coordination, comparison, opposition and hierarchical relational skills (Dixon et al. 2017). The PEAK Relational Training System directly and indirectly targets a variety of language and cognition proficiencies, of increasing complexity, which aid in the social repertoire of individuals with ASD. Findings which empirically support the efficacy of PEAK display the ability to establish a variety of advanced operants, proposing that PEAK is with high consideration as an autism intervention package.

As regulated by the No Child Left Behind Act (NCLB 2001) and the Individuals with Disabilities Education Act (IDEA 2004) educational goals must contain requirements in which academic experiences be enriched with high-quality, research-based instructional strategies. The overarching principle is for students to reach a level of proficiency in core academic areas, regardless of a race, gender, socioeconomic status, or identified disability (Cusumano 2007). It is important for clinicians and behavior analysist within the school system to identify research-based strategies that can be implemented with reliability, validity, and consistency. Incorporating the PEAK curriculum as an intervention tool for individuals with disabilities adheres to the requirements of the Curriculum Based Measurement (CBM) which mandates using data to identify why students are not meeting learning trajectories. Furthermore, the CBM requires that clinicians and education specialists evaluate instructional and environmental variables which lead directly to intervention, seeking to remove barriers which impede learning (Cusumano 2007). "An apparent need entwined in this approach is a metric for monitoring student skill acquisition in basic academic domains. Tools for collecting these data must, first be sensitive to small changes in skill acquisition—and these assessments bust be quick and easy to administer"

(Cusumano 2007, p. 24). Regulations among the evidence-based incorporations of practices within the school setting provides all the more reason to continue to develop and integrate research-based assessments and interventions within the school setting. Skinner's (1957) Analysis of Verbal Behavior has served as a conceptual basis upon which many applied behavior-analytic interventions for promoting language acquisition are based (Carr and Firth 2005; Sundberg and Michael 2001). Numerous examples exist within behavior analytic research suggesting the utility of conceptualizing language as verbal behavior units.

PEAK Relational Training System

The PEAK Curriculum provides four modules of behavior analytic training which target skill deficits among children with disabilities. The PEAK modules are easy to incorporate within the classroom setting, and the PCA provides an amiable, easy to use assessment tool to identify age-norm targets for each student. The PEAK Direct Training Module (PEAK-DT; Dixon 2014a) is the first model of the PEAK relational training system and consists of 184 programs which use discrete trial training to teach skills ranging from basic learning skills such as modeled play and turn taking to more advanced language and learning skills such as intraverbal emotions and receptively labelling time (McKeel et al. 2015). Of the 184 curriculum items found within the Direct Training Module, 64 items are contained in the PCA. The PEAK-DT Module demonstrates the effectiveness of a packaged curriculum grounded in Applied Behavior Analysis which contains a "wide range of early learning skill programming that should be considered the foundational building blocks of language and cognition" (Dixon 2019). PEAK provides tools with packaged techniques, delivering ease of use for individuals who may have little prior experience with behavioral protocol implementation.

With promotion and emergence of the PEAK Curriculum, many advancements have been made in the research and delivery of the assessment protocol which plays a vital role in program selection for students in areas of language and cognition. While each module began with its own individual assessment containing 184 skill targets, researchers such as Mark Dixon and Jordan Belisle have successfully evaluated and compressed the 4 individual assessments from each module into one comprehensive assessment for ease of implementation. The PEAK Comprehensive Assessment (PCA), Copyright 2019, was developed to allow practitioners the opportunity to conduct a "multifaceted assessment of an individual within a relatively practical period of time" (Dixon 2019). However, the PCA was designed to complement many of the tools found within the four modules rather than replace the previous protocols for obtaining information on individualized programming. "In each module, and perhaps most importantly in the Direct Training and Generalization modules, an indirect assessment of all 184 items found within that module curriculum is provided (Dixon 2019). The indirect assessments are still helpful tools that can be given to caregivers who daily interact with the individuals being assessed, which provides a comprehensive assessment of skills that are or are not present in the student's repertoire.

Despite previous limitations such as assessment variation and general guidelines for administration of the assessment as well as open interpretation for clinical interpretation of the assessment scores, robust research has been conducted to continue to investigate the efficacy, reliability, validity and treatment outcome measurements for the standardized PEAK curriculum. Development of the PCA has provided the first standardized assessment administration in the ABA world (Dixon 2019). The new standardized platform requires that clinicians and administrators adhere to specific rules when administering the assessment. The PCA has a

"verbatim script that must be read to the client and administrators are not allowed to deviate from the passages of text and must score a client's response specifically under the strict test conditions" (Dixon 2019). Another advantage to the implementation of the PCA is that the protocol can be completed by a skilled clinician within one hour, which drastically reduces the time it takes to identify deficits in language and cognition skill sets. The PCA is also "linked directly to the PEAK curriculum, thus providing extremely detailed step-by-step directions on how to teach deficit skills (Dixon 2019).

Why is this important for Behavior Analysts?

Construct validity for the PEAK Comprehensive Assessment is an important component for clinicians to consider. When measuring language as a construct, which research has proven is an area of deficit for students with disabilities, the main way that we can target these constructs and help students develop ways to learn is by measuring the overall skill set of the learner. Once skill targets have been identified, we then can develop behavior plans and implantation protocols that help strengthen these areas thus teaching students how to learn. With the PEAK Curriculum having its basis in relational frame theory one major component of teaching students comes from derived relational responding. With the PCA clinicians are able to identify skills through a direct assessment, while also evaluating the presence of these skills through and indirect assessment. If the skill is not present, the clinician can identify appropriate programing and skill targets that strengthen the skill. Evaluation of the construct validity of the PCA is important for clinician implementation. The BCBA ethics code mandates that behavior analysts be held to a standard that implements and evaluates the best research-based strategies for intervention. By evaluation the construct validity and internal consistency of the PCA we are providing behavior analysts

with a tool that is backed by research and can help students who struggle in areas of language and cognition.

METHODS

Procedure and Participants

Student sought, received and followed proper IRB guidance, see Appendix A. Human Subjects IRB Approval. The IRB for this study was issued on February 27, 2019. The IRB approval number is IRB-FY2019-576. The evaluation of the construct validity of the PEAK Comprehensive Assessment was evaluated using 55 PCA submissions from 6 different agencies (4 special education providers, and 2 ABA therapy clinics). All were in the midwestern United States. Each assessment was conducted one-on-one in a secure area to minimize distractions. There is a total of 55 participants ranging in age from 4 years to 16 years. Each participant had been diagnosed with developmental disability, prevalently autism spectrum disorder, which impedes learning of language and cognition. The PCA answer documents were submitted by behavior clinicians who have previous experience with administering the PEAK Comprehensive Assessment, and all assessment data obtained from the PCA was assessed by a graduate student studying Applied Behavior Analysis at Missouri State University. The PEAK Comprehensive Assessment consists of 344 items evaluating the display of language and cognition skills and each question ranges in terms of complexity. Correct responses are scored with a (+) signifying that the student can perform the skill and incorrect responses are scored with a (-) signifying that the student cannot complete the skill. This data collection method allows for an initial evaluation of the internal validity of the PCA using real-world clinical data to inform future studies on this tool in similar settings. We describe approaches to extend this work in the discussion section of the current thesis.

Materials: The PEAK Comprehensive Assessment

The PEAK Comprehensive Assessment (PCA) has evolved over the past five years as a result of robust research evaluating and implementing curriculum based in Relational Frame Theory, and more specifically PEAK Relational Training System (Dixon 2019). The total time to complete the PCA should be less than ninety minutes but can vary depending on skill set of the client, client age, performance level, behavior problems exhibited, and necessary breaks in assessing. Clinicians administer the assessment using one of two methods, the tabletop method or the floor corner method. The tabletop method is more appropriate for students who are familiar with sitting at a table, communicating with a clinician one on one. The floor corner method is helpful for students who exhibit severe problem behavior, display escape tendencies, or are not used to one-on-one setting for assessment. All assessment materials are to be placed between the client and the clinician and all administration challenges should be noted in the Behavioral Observation section of the PCA. The PCA manual provides step by step instructions for validity of implementation. The assessment materials are to be present before the assessment begins and it is recommended that the setting be appropriate for proper administration. All data obtained during the PCA is recorded in the client record booklet and this is to be uploaded via Qualtrics for data analysis. Each item of the PCA will be recorded and analyzed to evaluate the concurrent measurement of language and cognition for the PCA. Although several studies have evaluated assessments that comprise the PCA, no research to date has evaluated the entire PCA when administered directly with children with disabilities.

This assessment and integration protocol is an important component of the futuristic development and modification of principles of behavior analysis. Evaluating the construct validity of the PEAK Comprehensive Assessment provides a very large and robust sample of

clients, with the normative sample being individuals with developmental delays or autism spectrum disorder. This is significant for the future of behavioral research because it serves as a thorough evaluation of a larger sample size of individuals with developmental delays. The Direct Training Assessment of the PCA is divided into sections of Foundational Learning Skills (FLS), Perceptual Learning Skills (PLS), Verbal Comprehension Skills (VCS), and Verbal Reasoning Memory and Math Skills (VMS). At the end of each section, the clinician computes the leaner score. If the student scores a zero in any section, the clinician will discontinue the rest of the direct training assessment. However, if the student receives a correct answer for one or more items, the clinician will continue to the next assessment module. It is important to note, if the student scores a zero on any of the direct training assessments, the clinician is only to discontinue the assessment for the remainder of the Direct Training Module while moving on to the Generalization Module for assessment.

The Generalization module of the PCA is divided into four sections of Foundational Learning Skills and Basic Social Skills (LLS), Basic Verbal Comprehension, Memory, and Advanced Social Skills (CMS), Advanced Verbal Comprehension, Basic Problem Solving, and Advanced Math Skills (CPM), and Verbal Reasoning, Advanced Problem Solving, and Advanced Reading and Writing Skills (RPR). Similar to the Direct Training Module Assessment, of the PCA, at the end of each section, the clinician computes the leaner score. If the student receives a score of zero, the clinician is to discontinue the Generalization Assessment. If the student gets one or more items correct the clinician is to move on to the next section within the Generalization Module. It is important to note that once completing the Generalization portion of the assessment the clinician is to move on to the Equivalence module of the PCA.

The Equivalence module is divided into four sections Reflexivity (REF), Symmetry (SYM), Transitivity (TRS), and Equivalence (EQU). Each section of Equivalence within the PCA has a discontinue criterion stating that if two consecutive items are scored incorrect, the clinician is to discontinue the section, proceed to the next section, scoring any remainder items for the current section, as incorrect. At the end of each section within the Equivalence module, the clinician is to compute the learner score and proceed to the next section within Equivalence. Once the Equivalence Assessment is completed, the clinician and student then move forward to the Transformation section of the Assessment.

The Transformation Assessment is divided into two different categories of Receptive and Expressive Assessment. Within these two categories are the 6 sub sections Coordination (COR), Comparison, (COM), Opposition (OPP), Distinction (DIS), Hierarchy (HIR), and Deictic (DTC). Contained in the instruction manual there is discontinue criterion for each sub section. If the student scores incorrectly on the two practice items (P1 and P2) or scores incorrectly on three consecutive test items, discontinue the current section and proceed to the next. At the end of each section the clinician is to compute each learner score and proceed to the next section of assessment. Each module of the PCA is administered separately, and clients are allowed breaks as needed. It is necessary while administering the PCA to evaluate reinforcers at the beginning of each module to ensure student engagement.

Data Analysis

Following the attainment of the student PCA scores, a correlation matrix, a descriptive output, and a factor analysis (principle component analysis) were computed to evaluate the overall relationship between each item on the PCA and the construct of language and cognition.

Since each module of the PCA contains a different number of questions, the percent of module items answered correctly were calculated for each student. To compute this percentage, divide the student score by the total number of questions in the module, then multiply this number by 100 to evaluate the percentage of correct items. The percentage of correct responses was taken from each module for the 55 students and compared with each other module that makes up the PEAK Comprehensive Assessment. Student percentage scores were also used to run a principle component analysis as well as used to run descriptive statistics on each of the components to examine overall averages of scores and standard deviation. This allows researchers to evaluate the relationships between skill sets that are represented and how these skill sets correlate with complex questions of the assessment. The greater distance that the standard deviation is from zero, the greater spread of the average scores for each module. Potentially, examining students with different skill sets would give a greater spread of skills depending of the types of questions that were answered correctly. It is purported that each question on the PCA corresponds to the overall construct of language and cognition among individuals who display learning disabilities particularly those identifiable with autism spectrum disorder. Beginning with the Direct Training and Generalization subtests of the PCA, individual factors are presented to evaluate a cluster of skills that reportedly develop around the same time among individuals with and without disabilities (PCA Manual). The importance of the factor analysis in this evaluation is to examine the data and to determine if certain items within the data cluster together. These items help the assessor determine the appropriate programs for the client by identifying skill excesses as well as skill deficits. The two modules of Direct Training and Generalization contain 16 items for each of the four factors which identify expressive, receptive, and generative tasks (PCA Manual). The Equivalence and Transformation subtests of the PCA are made up of relations instead of factors.

Each of these relations become increasingly complex as the student progresses through the assessment. The Direct Training and Generalization modules examine skills that the student has previously learned while the Equivalence and Transformation modules serve as a predictor for current and future learning.

The last component to data evaluation for the Construct Validity of the PCA, is the Principle Component Analysis. This was computed using a Varimax Rotation (Rowsey, Belisle, and Dixon 2014). This produces several sources of data including a correlation matrix, to assess the construct validity of the PEAK Comprehensive Assessment. This method was chosen because the primary purpose of this study was to examine the underlying constructs of the PCA using an exploratory analysis of the individual modules and components.

RESULTS

The PCA is a tool designed to help clinicians better understand cognitive deficits including language and cognition and to help develop skill driven curriculum that specifically targets skills that are absent from the student repertoire. By evaluating the measurement of the PCA clinicians are able to better understand the tool being used to measure language as a function of verbal behavior. These results suggest moderate to strong correlations among each of the four modules which suggest that language and cognition are being measured.

Moderate to strong correlations suggest that each of the modules are measuring different dimensions of the same construct which is language and cognition. All subcomponents of the PEAK-CA had a correlation coefficient greater than 0.67, suggesting that each was associated with the same general construct. Table 1 shows the correlation matrix results. Among the four modules Direct Training and Generalization have the highest correlation of 0.92837445. Generalization and Transformation have the second highest correlation of 0.86716454, Generalization and Equivalence have the third highest correlation of 0.7557248. Direct Training and Transformation have the fourth highest correlation of 0.7287478, Equivalence and Transformation have the fifth highest correlation of 0.72687013, and Direct Training and Equivalence have the sixth highest correlation of 0.67100978.

The results of the correlation matrix suggest that each of the four modules evaluate skill deficits in the areas of language and cognition, which is the general construct under evaluation with the PEAK-CA. Among individuals who have been diagnosed with autism spectrum disorder or a developmental delay, the greatest deficits display themselves in the form of language and cognition as well as poor social skills. The DSM V identifies individuals with autism as having

deficits in nonverbal communicative behaviors including, developing, maintaining, or understanding relationships, and poorly integrated verbal and non-verbal communication (Ciccarelli and White 2014).

Table 1. Correlation Matrix of the PEAK Comprehensive Assessment

	DT	G	E	T
DT	1			
G	0.92837445	1		
E	0.67100978	0.7557248	1	
T	0.72687013	0.86716454	0.72381531	1

Figure 1 displays a scatterplot for each correlation evaluating language and cognition among the four modules. The highest correlation representation is between the direct training and generalization modules, suggesting that the questions on these to assessments significantly related in measuring the general construct. The second highest correlation is between the Direct Training and the Equivalence modules followed by the Direct Training and Transformation modules, the Generalization and Equivalence modules, and lastly the two modules with the least amount of correlation between responses are the Equivalence and Transformation modules.

Figure 2 displays the results of the descriptive analysis which suggest that the highest scores are in the Direct Training module with a reported average of 40.8275, the second highest scores are within the Generalization module with a mean of 20.9780, the third highest scores are

in the Equivalence module with an overall average of 16.7438 , and the fourth highest scores are

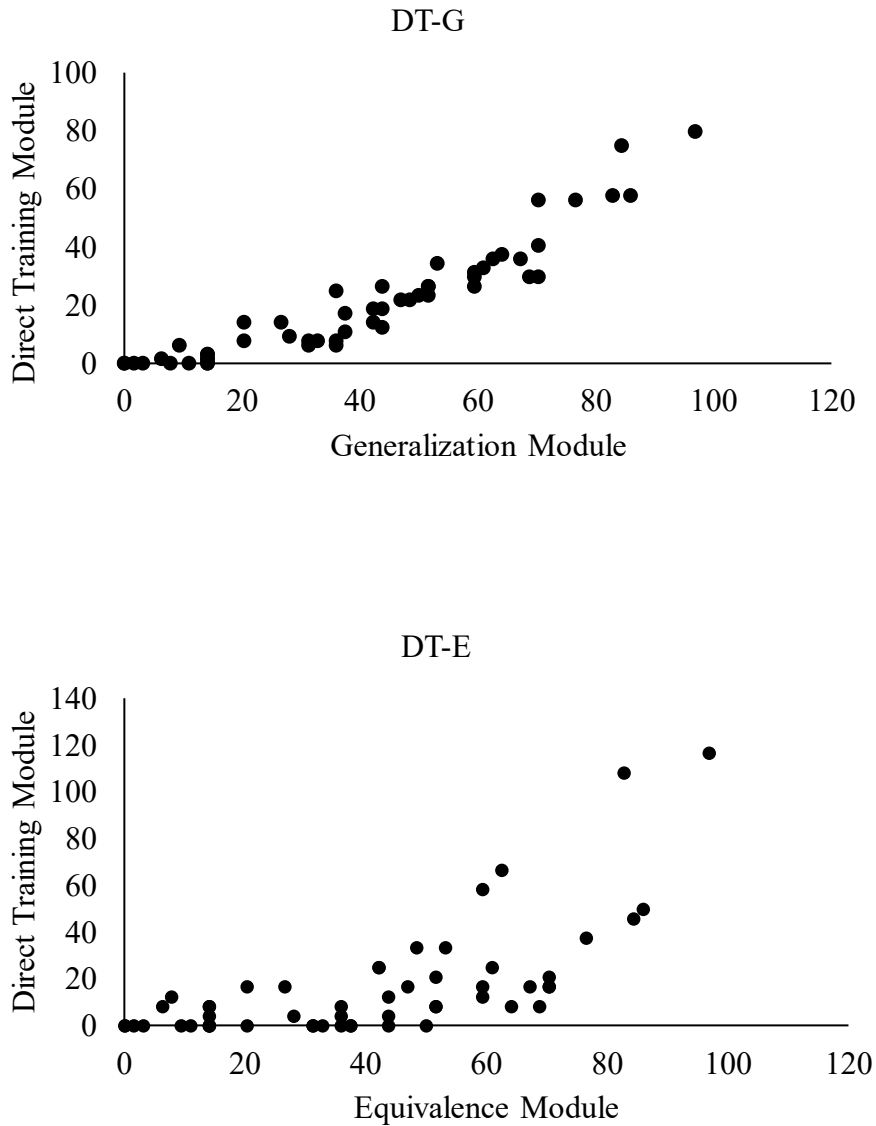in the Transformation module with an average of 8.1404.

DT-G

DT-E

Figure 1. Scatterplot for each of the correlations among the four PEAK Modules. Each percentage of correct responding for each module was correlated with the correct responding from another module to evaluate clusters which make up the construct evaluation of language and cognition.
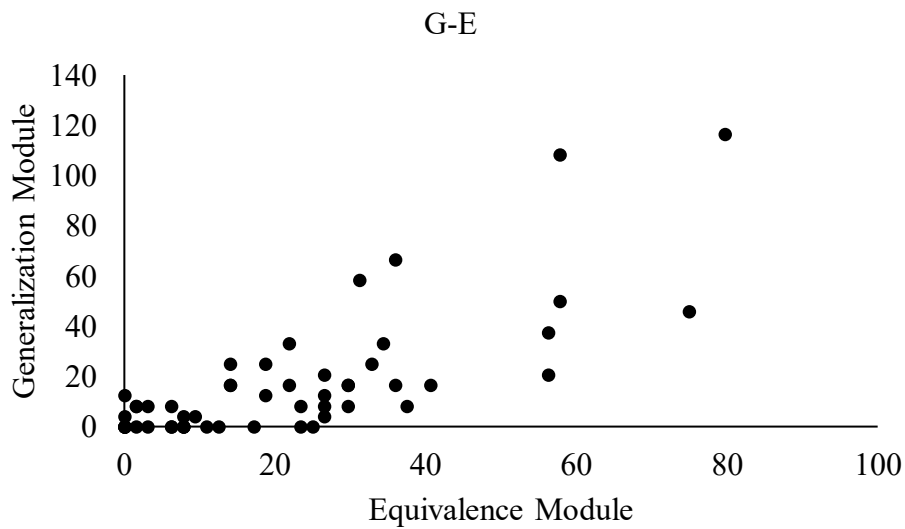
DT-T

G-E

Figure 1 (continued). Scatterplot for each of the correlations among the four PEAK Modules. Each percentage of correct responding for each module was correlated with the correct responding from another module to evaluate clusters which make up the construct evaluation of language and cognition.
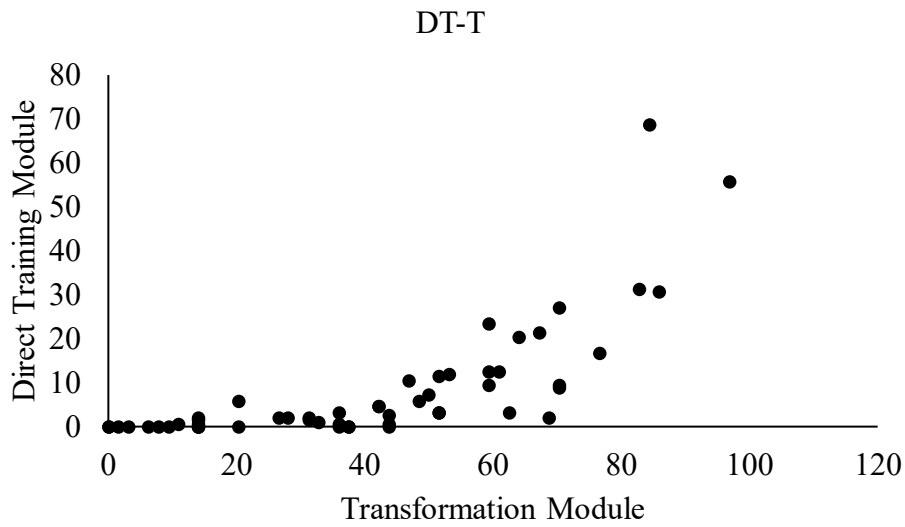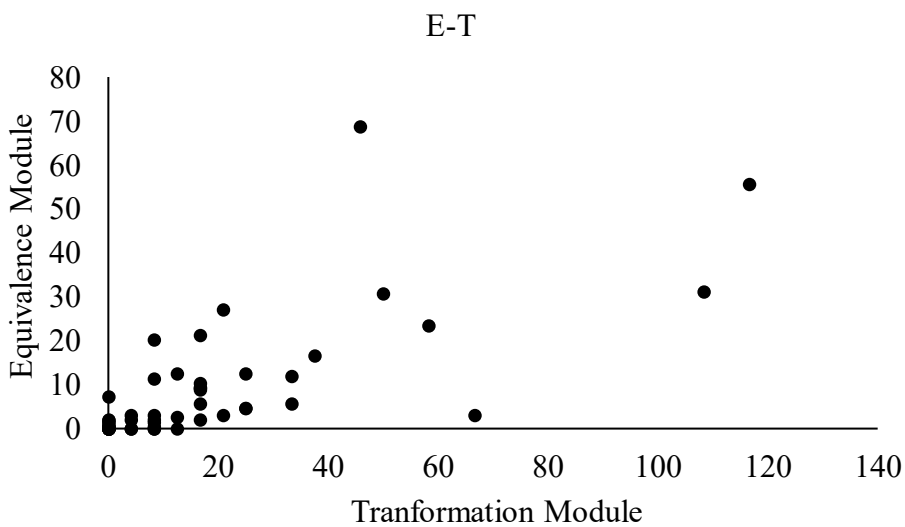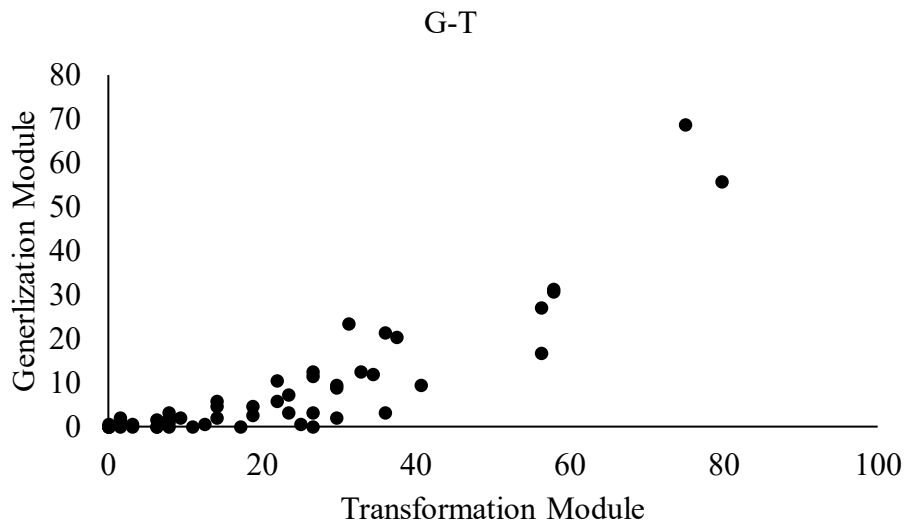
Figure 1 (continued). Scatterplot for each of the correlations among the four PEAK Modules. Each percentage of correct responding for each module was correlated with the correct responding from another module to evaluate clusters which make up the construct evaluation of language and cognition.

These results suggest that each module of the PCA increases in complexity from the Direct Training module to the Transformation module and furthermore suggest that the range of complexity positively correlates with the overall measurement of language and construct in the presentation of skills ranging from least difficult to most difficult. The standard deviations are reported for each module with Direct Training having a standard deviation of 25.20012, Generalization having a standard deviation of 19.70963, the Equivalence module having a standard deviation of 24.59700, and the standard deviation for the Transformation module being 13.58894. The variability of the higher standard deviation suggests that the data points for the average scores of each student are spread out across modules. This is consistent with data findings which suggest that each score on the PCA is correlated with a range of complexity depending on the skill set of the individual.
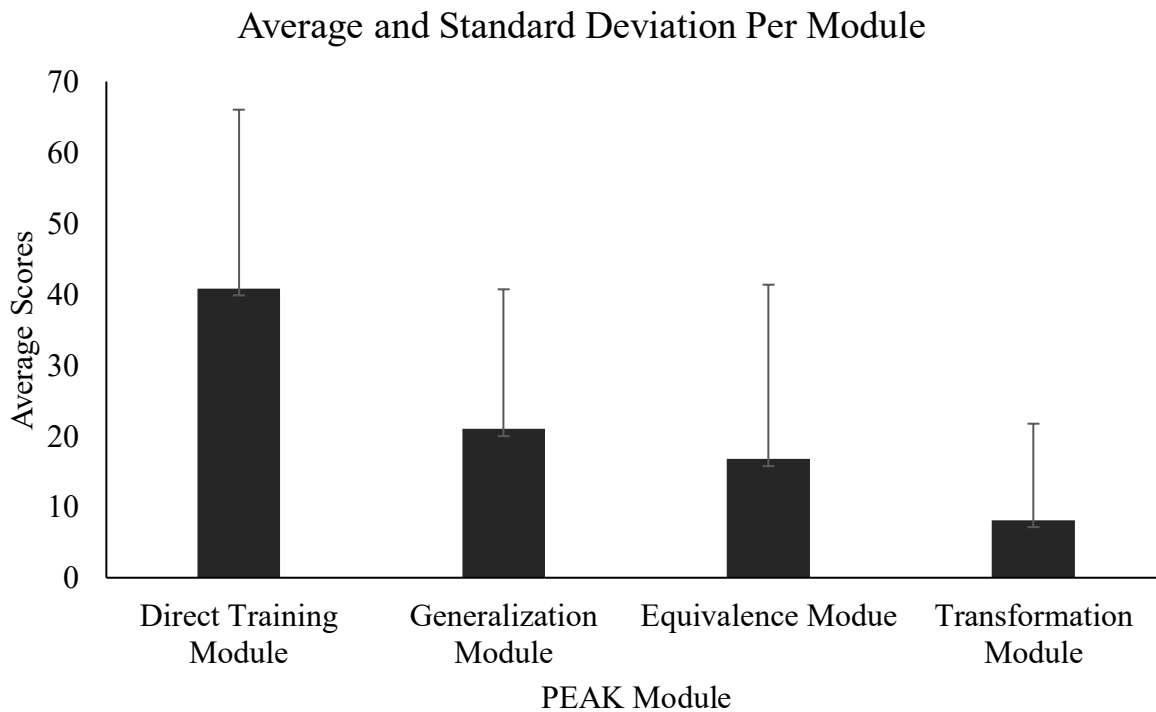

Figure 2. Average responding for each module as well as standard deviation for each module. The bars on the graph represent the mean, and the error bars represent the standard deviation.

A principal component analysis was conducted to evaluate whether certain items of the PCA cluster together. This method was chosen to further evaluate research conducted by (Rowsey, Belisle, and Dixon 2014). The results of the principal component analysis as seen in Figure 3 suggest that each module of the PCA examines one general construct, language and cognition. Reports from this analysis also show an eigenvalue > 1 for only one factor suggesting that there is one factor examining the general construct among the autism sample. The initial values suggest that the first component accounted for 83.627% of the variance, the second component 9.093% of the variance, the third component 6.359% of the variance, and the fourth component 0.922% of the variance. This suggests that overall only one factor accounts for the variance among the percentage of correct responding that was examined. This data analysis further purports that the PEAK Comprehensive Assessment is a valid tool for evaluating the general construct of language and cognition among individual who display a variety of disability and particularly those who display characteristics of Autism Spectrum Disorder. These results could further suggest that the overall percentage of correct responding examined on the PCA results is a good indicator of the overall skill set of the individual. Rather than simply examining verbal behavior, the PEAK CA allows the clinician to evaluate derived relational responding in a way that can be a predictor of future learning for the student.
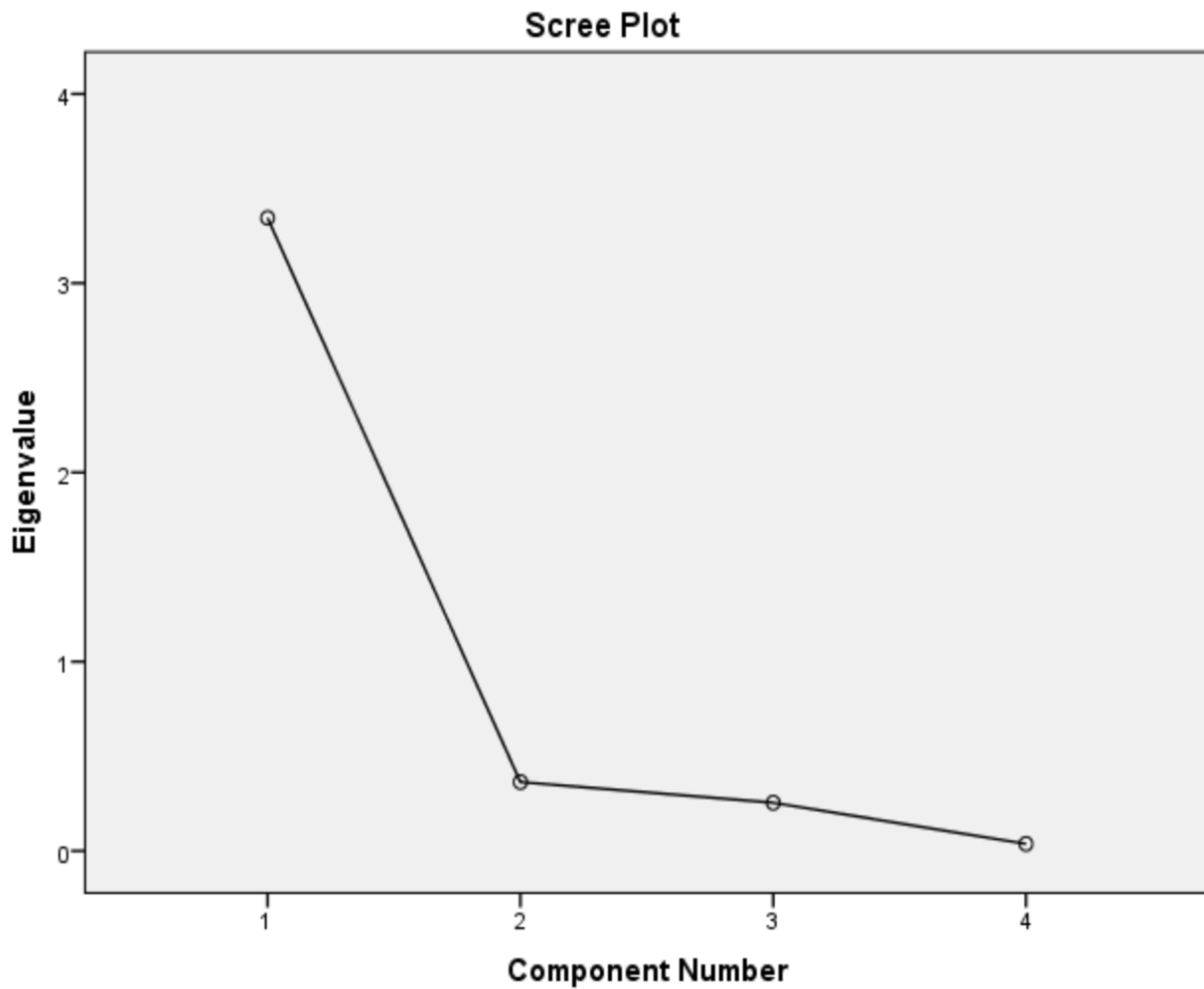
Figure 3. Principle Component Analysis shows that one factor displays the majority of the variance for the individual items evaluated within the PCA. These results suggest that only one factor is responsible for the variance as a single variable, suggesting that factor one accounts for as much variance as a single variable. These results are consistent with previous research that examined the principle component analysis of the PEAK Relational Training System modules.

DISCUSSION

The results of the current study evaluate the overall constructs examined within the PEAK Comprehensive Assessment. The PCA is designed to provide a time effective evaluation tool for identifying skill deficits among individuals with developmental disabilities. Among advancements made to identify the underlying constructs of the PCA are the developments designed to aid in curriculum for skill deficits and intervention that can target individual client skill sets determined by overall complexity of performance on the assessment. The PEAK CA evaluates a variety of skills from basic verbal behavior to complex relations. Results from this study show that each of the four modules measure language and cognition as a construct and that this single construct is responsible for any variance in skill performance on the PEAK CA.

This data helps support the PCA as an effective tool at developing curriculum to teach students with disabilities operational life skills and language skills using components based in relational frame theory. Future research in this area of study could further examine the relationship between the questions on the PEAK -Equivalence assessment component and the questions on the PEAK Transformation assessment component. There are many similarities in the data presented in the correlation and it could be presumed that after further evaluation many of the questions may equally evaluate future learning complexity and diversity for the student.

Implications

There are several pragmatic implications to the results of this evaluation. By evaluating the prevalence of a unifying construct within the PCA researchers can begin to examine the necessity of assessing each skill item by item. Theoretically, if each item on the PCA correlates

with language and cognition and there is a variety of advancement pertaining to skill evaluation, these results could affect the overall administration of the PCA. If a student displays skill in one area, it can be likely that the student will display skills in another area pertaining to skills of similar difficulty level. Thus also, if a student does not display a skill set related to more advanced assessment questions it can be likely that the student will not display other more advanced skills. These results, however, should be considered a pilot investigation to inform future research on the construct validity of the PCA or similar tools. In particular, a larger clinical sample should be sought to ensure greater external validity of the findings. In addition, a larger sample would allow for an analysis of the individual items contained within the PCA, similar to prior research on the various PEAK assessments. Doing so may be more successful in identifying latent constructs underlying operant learning processes.

Limitations

This study was a pilot study used to evaluate the underlying constructs of the PEAK Comprehensive Assessment. One limitation of this study is the small sample size. Future research in this area of study should involve financial incentive for clinicians who are willing to share the data of the PEAK Comprehensive Assessment. A greater sample size for this evaluation would be ideal in order to further examine the correlations among the PCA modules and responses. Another limitation to the study is the inability to complete Interobserver Agreement to examine fidelity on the administration of the PEAK Comprehensive Assessment. Future avenues of research could involve centers were IOA is available. Another future avenue of research could be evaluating fidelity for administration of the PCA via telehealth or web communications.

REFERENCES

American Educational Research Association., American Psychological Association., National Council on Measurement in Education., & Joint Committee on Standards for Educational and Psychological Testing (U.S.). (2014). *Standards for educational and psychological testing*.

American Psychological Association Committee on Test Standards. (1952). Technical recommendations for psychological tests and diagnostic techniques: preliminary proposal. *American Psychologist*, 7, 461–475

Belisle, J. (2020). Model dependent realism and the rule-governed behavior of behavior analysts: applications to derived relational responding. *Perspectives on Behavior Science*. doi: 10.1007/s40614-020-00247-x

Belisle, J., Dixon, M. R., Stanley, C. R., Munoz, B., & Daar, J. H. (2016). 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*(4), 965–969. doi: 10.1002/jaba.324

Belisle, J., & Dixon, M. (2018). Home. Retrieved January 1, 2020, from https://www.peak2aba.com/.

Carr, J. E., & Firth, A. M. (2005). The verbal behavior approach to early and intensive behavioral intervention for autism: A call for additional empirical support. *Journal of Early and Intensive Behavior Intervention, 2*(1), 18–27. https://doi.org/10.1037/h0100297

Centers for Disease Control and Prevention (CDC). Autism spectrum disorder (ASD). Retrieved December 28, 2019, from https://www.cdc.gov/ncbddd/autism/data.html

Ciccarelli S.K., & White, J.N. (2014). *Psychology: Dsm 5*. Boston: Pearson

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, *52*(4), 281–302. doi: 10.1037/h0040957

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*(1), 24–34. doi: 10.1037/h0100099

Devitt, J. H., Kurrek, M. M., Cohen, M. M., Fish, K., Fish, P., Noel, A. G., & Szalai, J.-P. (1998). Testing internal consistency and construct validity during evaluation of performance in a patient simulator. *Anesthesia & Analgesia*, *86*(6), 1160–1164. doi: 10.1097/00000539-199806000-00004

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. (2019). *Peak comprehensive assessment: administration manual*. Carbondale, IL: Shawnee Scientific Press.

Dixon, M. R., Whiting, S. W., Rowsey, K. undefined, & Belisle, J. undefined. (2014). Assessing the relationship between intelligence and the PEAK relational training system. *Research in Autism Spectrum Disorders*, *8*(9), 1208–1213. doi: 10.1016/j.rasd.2014.05.005

Dixon, M. R., Belisle, J., Stanley, C. R., Speelman, R. C., Rowsey, K. E., Kime, D., & Daar, J. H. (2016). Establishing derived categorical responding in children with disabilities using the PEAK-E curriculum. *Journal of Applied Behavior Analysis*, *50*(1), 134–145. doi: 10.1002/jaba.355

Dixon, M. R., Belisle, J., Mckeel, A., Whiting, S., Speelman, R., Daar, J. H., & Rowsey, K. (2017). An internal and critical review of the PEAK relational training system for children with autism and related intellectual disabilities: 2014–2017. *The Behavior Analyst*, *40*(2), 493–521. doi: 10.1007/s40614-017-0119-4

Goodwin, L. D., & Leech, N. L. (2003). The meaning of validity in the new standards for educational and psychological testing: *Measurement and Evaluation in Counseling and Development*, *36*(3), 181–191. doi: 10.1080/07481756.2003.11909741

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.

Individuals with Disabilities Education Act, 20 U.S.C. § 1400 (2004)

Laura D. Goodwin & Nancy L. Leech (2003) The meaning of validity in the New standards for educational and psychological testing:, measurement and evaluation in counseling and development, 36:3, 181-191, DOI: 10.1080/07481756.2003.11909741

Mckeel, A. N., Rowsey, K. E., Belisle, J., Dixon, M. R., & Szekely, S. (2015). Teaching complex verbal operants with the PEAK relational training system. *Behavior Analysis in Practice*, *8*(2), 241–244. doi: 10.1007/s40617-015-0067-y

Newton, P. E., & Shaw, S. D. (2014). *Validity in educational & psychological assessment*. Los Angeles: SAGE.

No Child Left Behind Act of 2001, P.L. 107-110, 20 U.S.C. § 6319 (2002).

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences*, *1*(4), 515–526. doi: 10.1017/s0140525x00076512

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*(1), 15–23. doi: 10.1007/s10882-014-9398-9

Scheuermann, B., Webber, J., & Lang, R. (2019). *Autism: teaching makes a difference*. Boston, MA: Cengage.

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

Sidman, M. (1994). *Equivalence relations and behavior: a research story*. Boston, MA: Authors Cooperative.

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*(1), 5–22. doi: 10.1901/jeab.1982.37-5

Skinner, B. F. (1957). *Verbal behavior*. New York: Appleton-Century-Crofts.

Slaney, K. L. (2017). *Validating psychological constructs: historical, philosophical, and practical dimensions*. London, United Kingdom: Palgrave Macmillan. doi: 10.1057/978-1-137-38523-9_3

Sundberg, M. L., & Michael, J. (2001). The benefits of skinner's analysis of verbal behavior for children with autism. *Behavior Modification*, *25*(5), 698–724. doi: 10.1177/0145445501255003

U.S. Department of Education (2016a). *Thirty-eighth annual report to congress on the implementation of the individuals with disabilities education act, 2016*. Washington, D.C.: Office of Special Education and Rehabilitative Services, Office of Special Education Programs.

APPENDIX A. HUMAN SUBJECTS IRB APPROVAL

Date: 5-21-2020

IRB #: IRB-FY2019-576
Title: Evaluating the Efficacy of the PEAK Relational Training System in Active Clinical Settings - Existing Data
Creation Date: 2-27-2019
End Date:
Status: Approved
Principal Investigator: Jordan Belisle
Review Board: MSU
Sponsor:

Study History

Submission Type	Initial	Review Type	Expedited	Decision	Approved
Submission Type	Modification	Review Type	Expedited	Decision	Approved

Key Study Contacts

Member		Role		Contact
Member	Jordan Belisle	Role	Principal Investigator	Contact jbelisle@missouristate.edu
Member	Jordan Belisle	Role	Primary Contact	Contact jbelisle@missouristate.edu
Member	Taylor Lauer	Role	Investigator	Contact tl45911@live.missouristate.edu
Member	Nicole McDonald	Role	Investigator	Contact nc88@live.missouristate.edu
Member	Lindsey Dennis	Role	Investigator	Contact lam1014@live.missouristate.edu
Member	Annalise Giamanco	Role	Investigator	Contact annalise2015@live.missouristate.edu