Promoting the Emergence of Advance Knowledge (Peak) Relational Training System: Exploring the Practical Implementation of the Generalization Training and Direct Training Modules.

Marthinus Barnard (B.A. Hons)

National University of Ireland, Maynooth.

Research Supervisor: Dr. Carol Murphy

Co-Supervisor: Michelle Kelly

Thesis Presented in Part - Fulfilment of the Requirements for the Doctorate in Psychological Science (Behaviour Analysis and Therapy)

Department of Psychology, Faculty of Science and Engineering

Index

Chapter	Description	Page
	Acknowledgements	iii
Chapter 1:	Review of the Current Relevant Literature	1
Chapter 2:	Study 1: Exploring the implementation of PEAK in a Practical Setting	28
Chapter 3:	Study Two: Exploring the Generalization Training Protocol's Effect on related PEAK Direct Assessment Targets.	70
Chapter 4:	General Discussion	84
	References	94

Acknowledgements

I’d like to thank my Supervisor Dr. Carol Murphy, my Co-Supervisor Dr. Michelle Kelly, and everyone who supported and guided me on this long road. Without their help, I may have never reached this point. I’d also like to thank my wife Rachel, my parents, Marlize and Kobus, my brother Pieter, my entire extended family and circle of friends who believed in me, encouraged me, and never gave up on me.

Chapter One

Review of the Current Relevant Literature

Abstract

The present research sought to explore the implementation of the PEAK Relational Training System's Generalization and Direct Training modules in a practical setting across two studies. Study 1 investigated whether a relationship exists between the PEAK-G and PEAK-D Training protocols, if the skills mastered through these would generalize to the natural environment, and whether staff would rate these protocols as socially valid. In Study 1, participants included 12 children aged 4-5 years, 5 with diagnosed autism, 7 with neuro-typical development, and 6 members of staff. Assessments were conducted with all participants using PEAK-G and PEAK-D Modules; subsequently 5 target relational skills from the PEAK-G were taught, followed by a second PEAK-G and PEAK-D assessment. Results showed that following PEAK-G training, children derived target skills in the PEAK-D module that were absent in the first assessment which involved no exposure to PEAK-D training. Mastered PEAK skills were also found to have generalised to the appropriate natural environment. Staff rated these PEAK modules as socially valid and acceptable language training protocols. Study 2 aimed to partially replicate results from Study 1 using a multiple-probe design across behaviours to explore if PEAK-G training would result in related PEAK-D scores being mastered without PEAK-D training. Participants were 4 children aged 4 with diagnosed autism. Baseline conditions involved probes of targeted skills from the PEAK-G and PEAK-D modules. This was followed by PEAK-G training and subsequent tests of related target skills in PEAK-D to determine if learning PEAK-G relational skills resulted in acquisition of PEAK-D relational skills in similar target areas. Results showed that relational PEAK-D target skills that were absent prior to PEAK-G training were evident for 5 children with autism. Procedural fidelity measures also indicated that high levels of fidelity were maintained by staff over an extended period following training via Behaviour Skills Training. Implications of results and suggestions for future research in refining the efficiency of PEAK training are discussed.

Autism spectrum disorders (ASD) are neurodevelopmental conditions characterized by severe impairment in reciprocal social interactions, communication skills and the presence of restricted, stereotypical behaviours (APA, 2000). The prevalence of Autism in Ireland is currently estimated to be at a rate of 1% for the entire population (Sweeny, & Staines, 2016). Several intervention approaches for ASD and related problem behaviours exist (Peters-Scheffer, Didden, Korzilius, & Sturmey, 2010). These include floor time, holding, medication, options, Picture Exchange Communication System (PECS), sensory integration, speech and music therapy, special education and visual schedules (Green et al., 2006; Hess, Morrier, Heflin, & Ivey, 2008). However, there is little empirical evidence for the effectiveness of many of these approaches and available evidence shows mixed results (Foxx, 2008; Howlin, 1997; Schechtman, 2007; Smith, 1999). Amongst all the available interventions, those based on Applied Behaviour Analysis (ABA), (Cooper, Heron, & Heward, 2007), have emerged as a commonly recognized effective treatment within the area of autism services (Dixon, Carman, et al., 2014; Grey, & Hastings, 2005; Virués-Ortega, 2010; Foxx, 2008; Remington et al., 2007; see also Larsson, 2012, 2013 for a complete review).

The field of ABA was derived from the work of B.F Skinner (1938, 1957) by researchers such as Don Baer, Todd Risley, Ivar Lovaas, and Montrose Wolf (Baer, 1993). As an applied approach towards behaviour it has proven efficacy in a wide range of areas and behavioural problems, including AIDS prevention in healthcare staff (DeVries, Burnette & Redman, 1991), gerontology (Gallagher, & Keenan, 2006), seatbelt use (Van Houten, et al., 2005), medical procedures (Hagopian, & Thompson, 1999), industrial safety (Foxx, Hopkins, & Anger, 1987), and education (Dardig, et al., 2005). The applications of ABA that are employed by behaviour analysts to alter behaviour aim to increase socially appropriate behaviour, reduce inappropriate behaviour, facilitate the learning of new

functional skills, teach the skill of generalizing responses across different settings, people, topographies and functions, and promote stimulus control overall (Sulzer-Azaroff and Mayer 1991). These aims are achieved by using ABA techniques which are validated through rigorous use of experimental methods (Hess, Morrier, Heflin, & Ivey, 2008). Research has demonstrated that as well as being effective, among the currently utilized treatments for individuals with developmental disabilities such as autism, those based on the principles of ABA are the most widely used (Hess, Morrier, Heflin, & Ivey, 2008; National Autism Center, 2009). For example, the ABA based ABC sequence of behaviour acquisition technique or Discrete Trial Training method (DTT), is amongst the most frequently used and involves the presentation of a discriminative stimulus, such as ‘‘what is this’’ (A), waiting for a response on the part of the learner (B), and then providing an appropriate consequence such as positive reinforcement or a corrective prompt (C) as well ensuring short intervals between discrete trials (Brown-Chidsey & Steege, 2004; Koegel, Russo, & Rincover, 1977). When this technique is combined with other commonly used techniques such as Early Intensive Behaviour Intervention (EIBI) research has demonstrated that this results in increases in academic achievement, adaptive behaviour skills, and social skills (Myers & Johnson, 2007) as well as specific verbal and language skills in participants with intellectual disabilities (i.e.., Greer, Yaun, & Gautreaux, 2005; Yamamoto & Mochizuki, 1988).

In addressing verbal and language deficits, behaviour analytic approaches commonly focus on functional communication training (FCT) to increase language abilities and subsequently reduce problematic behaviours (see Tiger, Hanley, & Bruzek, 2008 for a review of literature on FCT). Research has demonstrated that addressing the language skill deficits typically found in children with ASD is of critical importance as success in this approach has been shown to be a strong predictor of improved overall outcomes (Szatmari, Bryson, Boyle, Atreiner, & Duku, 2003; Frampton, Robinson, Conine, & Delfs, 2017; Venter, Lord, &

Schopler, 1992). Approaches such as intensive language instruction are derived from psycholinguistic (Lovaas, 2003) and verbal operant conceptualizations of language learning (LeBlanc et al. 2006). Although the former approach to language instruction, psycholinguistics, is primarily focused on receptive and expressive language, the latter approach is derived from B.F. Skinner's (1957) conceptual analysis in his text *Verbal Behavior*. Skinner posited that language learning was best described not by topography, but by the functional relationship between stimulating events (A), the response of the speaker (B), and subsequent actions on the part of the listener (C) which are referred to as Verbal Operants (Skinner, 1957). For example, Skinner suggested that requesting or ''manding'' was best described in the context of a learning history in which a response was more likely to occur in the future if it was emitted in the presence of appropriate motivation and resulted in the delivery of the desired item (Skinner, 1957). Skinner defined six primary operants in this manner: the mand (requesting), tact (labelling), echoic (imitating another's verbal behaviour), intraverbal, textual, and transcription responses (writing or reading in response to a discriminative stimulus). Skinner also described various types of audience controls and relational responses, called autoclitics, which modified the form of responses and modulated the meaning of language. One of the primary tenets of Skinner's account of language is that each of these Verbal Operants must be considered a functionally independent, yet inter-related, product of the relevant environmental variables that control when and if the operant will be emitted (Skinner, 1957). Empirical support for Skinner's account of language has increased in recent times (Sautter, & Le Blanc, 2006). Many research studies have supported and reviewed Skinner's account of language. For example,: Bailey, and Wallander, (1999) explored Skinner's complex operants and treatment of higher cognitive processes such as language, thinking, and understanding and found that Skinner treated such notions as complex forms of operant behavior that can be connected with stimuli and consequences and

has been underappreciated by behaviour analysts and almost completely ignored by the rest of psychology and philosophy. Sautter, and LeBlanc, (2006) conducted a quantitative review of 44 published literature which examined the studies on verbal behavior that were empirical in nature, concerned with human verbal behavior, and addressed at least one verbal operant (e.g., mand, tact, echoic, autoclitic, and/or intraverbal behavior) within the experiment, and found a growing body of research exists to support many of the tenets of Skinner's conceptualization and taxonomy but many areas of verbal behavior research have yet to be addressed. Johnson, Kohler, and Ross, (2016) described the history of the theory of verbal behaviour, defined the key features of the theory, and presented research supporting its application to the communication development of children with autism. Frampton, et al., (2017) assessed the efficiency of tact and listener training for participants with ASD, and found these forms of training to be efficient means of training target skills. Other studies include Michael, (1984); and Braam and Polling, (1983).

However, despite the broad scope of Skinner's conceptualization of language learning, empirical research upon Skinner's analysis has been generally limited to mands, tacts, intraverbals, and echoic responses (Dixon et al. 2007; Dymond et al. 2006). Nonetheless, these four verbal operants have translated well into ABA-based intervention strategies as they are fundamental to the language-learning process and are often observed to be absent in individuals with language deficits (Sundberg and Michael 2001).

Applied Verbal Behaviour (AVB) or Verbal Behaviour Therapy (VBT) is an example of an ABA based approach to language learning that has traditionally focused on Skinner's verbal operants. In gene.ral, AVB assessment protocols have been used to demarcate the verbal repertoires of individuals across the independent verbal operant categories and inform instructors as to the types of skills that need to be taught. The application of the verbal

behaviour approach to improving language has surged over the last 30 years (see Dymond, O'Hora, Whelan, & O'Donovan, 2006; Reed, DiGennaro Reed, Jenkins, & Hirst, 2014) namely due to the success of this approach in children with ASD (Sundberg & Michael, 2001). The AVB approach to early intervention for autism also gained widespread use after the publication of *Teaching Language to Children with Autism or Other Developmental Disabilities* by Sundberg and Partington (1998) and a manualised assessment protocol, *The Assessment of Basic Language and Learning Skills: An Assessment, Curriculum Guide, and Tracking System for Children with Autism or Other Developmental Disabilities* (ABLLS; Partington & Sundberg, 1998). The manualised nature of the ABLLS enabled clinicians and parents alike to implement ABA based language-acquisition programmes to help learners with autism acquire communication skills. Other similar assessment programmes such as *Shaping Knowledge through Individual Life Learning Systems* (SKILLS: Dixon, Tarbox, Najdowski, Wilke, & Granpeesheh, 2011), the revised ABLLS (ABLLS-R; Partington, 2008), and the *Verbal Behavior Milestones Assessment and Placement Program* (VB-MAPP; Sundberg, 2008) have also gained widespread popularity and use since being published (Reed, & Luiselli, 2016). From these, two of the most successful assessment-guided curriculums which incorporate behaviour analytic principles and concepts in their design, are the Assessment of Basic Learning and Language Skills-Revised (ABLLS-R; Partington, 2008) and Verbal Behavior Milestones Assessment and Placement Program (VBMAPP; Sundberg, 2008). The ABLLS-R serves as an assessment-guided curriculum and skill tracking guide for children with developmental disabilities, consisting of 25 different skill sets comprised of skills for basic learning and language acquisition (Partington, 2008). The VB-MAPP is similar insofar as it also features an assessment-guided curriculum and skill tracking guide, but targets are organized developmentally from 0 to 48 months of age (Reed, & Luiselli, 2016). Additionally, the VB-MAPP includes Barriers and Transition

Assessments and Task Analysis and Skills Tracking tools which address other areas of treatment.

These behavioral packages all make strong use of the science and principles of ABA, incorporating techniques such as positive reinforcement strategies to increase verbal skills, and both have been used extensively in research to provide a relative measure of an individual's verbal repertoire, especially the VB-MAPP (Charania et al. 2010; Grannan, and Rehfeldt 2012; Gunby et al. 2010). Despite this, there are some limitations that exist with these packages, as highlighted in the existing PEAK literature.

Dixon, Whiting, Rowsey, and Belisle (2014), claim that a lack of empirical research demonstrating the reliability or validity of these curricula and assessment tools limits the confidence with which these packages may be implemented (Dixon, Whiting, Rowsey, & Belisle, 2014). For example, it is reported that no empirical data exists demonstrating meaningful outcomes following completion of either program, and subsequently little is known about whether training using these curricula contributes to any meaningful advances in learning or academic skills for participants. Also, that comparisons with renowned measures of functioning and intelligence have not been established, so whether the subject matter of these packages represent cognitive skills and abilities is left to the assumption of the implementers (Dixon, Whiting, et al., 2014).

Dixon, Carman, et al., (2014), asserts that these verbal behaviour assessment packages are further limited in that they do not fully utilize many of the concepts presented by Skinner (1957), and often stop with the most basic units of verbal language utilized in the previously discussed study: mainly tacts (labelling), mands (requesting), intraverbals (explaining, discussing, or describing an item or situation that is not present, or not currently happening), and echoics (verbal imitation) (Dixon, Carman, et al., 2014). The researchers

further state that although establishing these types of responses may promote language where it has scarcely developed, in order to be truly comprehensive and produce meaningful effects, more complex verbalizations must be targeted (Dixon, Carman, et al., 2014). One final limitation highlighted is that these assessment packages are further limited in that they neglect many of the important advances in the study of language that have occurred since Skinner's verbal operants from 1957 which are important for children with autism such as Stimulus Equivalence and Derived Relational Responding (Dixon, Carman, et al., 2014).

In Dixon and Belisle (2015), it is further claimed that another limitation is that no published data exists which tackles the relationship of scores on these assessments and the real-world outcomes, progress, or functional skills of children with autism, and that evidence of concurrent improvement on other frequently used assessments of language and functional skills or educational goals is scant. Furthermore, Dixon and Belisle (2015) also asserts that currently, in the absence of formalized assessments of language skills, decisions regarding the selection of target skills for language skill interventions are left to the discretion of the care provider or program supervisor, and despite this approach being common place in ABA settings, it is not a sufficiently scientific approach in treating the severe language deficits frequently experienced by individuals with ASDs (Dixon, Belisle, 2015).

In reviewing these limitations it is important to note that several stem from direct comparisons between the PEAK system and existing systems, such as the VB-MAPP and ABLLS-R. This presents a difficulty as to compare PEAK, which is an assessment, curriculum and out-of-the-box type training system, to an assessment guided curriculum would lead to an invalid comparison. Instead it would be of more benefit if the PEAK system was viewed as a more a modern, advanced, and comprehensive behavioural assessment package that not only assesses participants and guides implementers, but also provides ready-made task-analysis style programs for implementers. These train participants using methods

based on important advances in the study of language that have occurred since Skinner's verbal operants from 1957, such as Stimulus Equivalence, Relational Frame Theory, and Derived Relational Responding.

From the above, two important limitations of the existing AVB assessment protocols to consider are that these verbal behaviour assessment packages do not fully utilize many of the concepts presented by Skinner (1957) and often stop with the most basic units of verbal language, and that these assessment packages neglect many of the important advances in the study of language that have occurred since Skinner's verbal operants from 1957 which are important for children with autism such as Stimulus equivalence and Derived Relational Responding (Dixon, Carman, et al., 2014). Despite the immense success of Skinner's verbal operants for both conceptual and practical accounts of language, there is a growing debate about whether this account is complete (Gross, & Fox, 2009). This debate over the completeness of Skinner's account of verbal behaviour is being questioned in contemporary research for several reasons. Firstly, Skinner's account of verbal behaviour on which these assessment packages are based, fails to explain Derived Relational Responding (especially Stimulus Equivalence) involving arbitrary stimuli, and secondly, it relies on socially mediated learning (Hayes, Barnes-Holmes, Roche, 2001; Luciano, Barnes-Holmes, & Barnes-Holmes, 2001; Reed, & Luiselli, 2016). Stimulus equivalence, first described by Sidman (1971; 1982; 1994), is an empirical phenomenon that attempted to account for language acquisition that was found to occur outside of direct operant contingencies via Derived Relational Responding (Barnes, 1994; Rehfeldt, 2011). For example, being able to respond to numerous exemplars taken from the colour category "blue" as being "blue" despite very few shades of blue being directly taught as belonging to said colour category (Sidman, 2009). According to Sidman, instead of all stimuli or behaviours such as language acquisition being learned as a result of direct operant contingencies, equivalence relations and

equivalence responding also play a vital role (Sidman, 1971). Sidman states that for a behaviour, such as language acquisition, to be described as an example of equivalence responding it must demonstrate reflexivity, symmetry, and transitivity in the absence of differential reinforcement. Reflexivity describes when a stimulus is matched to itself, i.e. matching to sample A=A, B=B, C=C. Symmetry describes when the sample-comparison relations that are established during training can be reversed, i.e. A=B means B=A. Transitivity describes when a third stimulus is introduced and trained into the already existing network, i.e. A=B, B=C. An introduction like this would typically result in the derived relations A=C and C=A, despite no direct training occurring between A and C. When all three derived relations are present like this the stimuli are described as being part of an equivalence class, as each stimulus is equivalent and can be substituted for one of the others. The phenomenon of Stimulus Equivalence and its account of learning through equivalence relations has been found to serve as a behavioral model for explaining the emergence of rudimentary comprehension and reading skills, and the development of generative syntactic repertoires. However, some common problems with Stimulus Equivalence is that the definitions of the relations are too narrow to describe further relations which have been demonstrated within reflexivity, symmetry, and transitivity; and researchers encounter great difficulty in establishing or demonstrating Stimulus Equivalence in non-human organisms or very young babies (Augustson, & Dougher, 1992; Green, 1990; Lipkens, Hayes, & Hayes, 1993).

A modern approach that address the deficits in Skinner's account of language and expands Stimulus Equivalence is Relational Frame Theory (RFT). This is a comprehensive approach to understanding how humans learn using verbal behaviour through Derived Relational Responding (Hayes, 1994). According to Relational Frame Theory, humans learn patterns of relational responding at an early age during their interactions with the verbal

community and with an appropriate history of exemplar training (Barnes, 1994). Four processes that are considered defining characteristics of RFT are: relational responding, mutual entailment, combinatorial entailment, and transformation of stimulus function. Relational responding describes the process of discriminating relationships between stimuli and responding in a way that is based on those relationships between these stimuli, for example, buying a particular brand of shampoo because it smells better than the other brands of shampoo on the market. Mutual entailment indicates that when a stimulus is related in a specific way to another stimulus, then the second stimulus is related in a complementary way to the original stimulus, for example, if this particular shampoo smells better than another brand of shampoo then that brand of shampoo doesn't smell as good as the first brand of shampoo. Combinatorial entailment refers to the reciprocal relationships that exist between two stimuli by virtue of how these stimuli are related to other intermediary stimuli. An example of how combinatorial entailment occurs is as follows: if a participant has learned that 1 cent is less than 5 cents, and 5 cents is more than 1 cent (mutual entailment), but the participant has also learned that 5 cents is less than 10 cents and 10 cents is more than 5 cents, then combining these two sets of mutually entailed relations may result in a combinatorically entailed relationship being derived between the 1 cent and 10 cents, where 1 cent is less than 10 cents and 10 cents is more than 1 cent, despite no direct training having occurred. The final process, transformation of function, occurs when relational responses between stimuli are made. When two stimuli are related, functions of each stimulus change according to what stimulus they are related to, and how these stimuli are related. For example, if a child who is afraid of sharks due to sufficient experience with sharks, either directly or indirectly, is told that there are sharks in the water at their favourite beach, the hierarchical relationship this establishes between their favourite beach and sharks results in a transformation of their favourite beach's function. Whereas previously their favourite beach

would have been “fun”, “relaxing,” and a “great place to go for a swim,” it is now “dangerous” and an object of fear by virtue of its newly established relationship with sharks, and the events and experiences the child usually frames in coordination to sharks.

Over time and through multiple relations across many stimuli, these relations become sufficiently abstracted so as to be arbitrarily applied to any stimuli (Barnes, 1994). The term *relational framing* describes the process in which stimuli are coordinated into different relational frames through operant learning under contextual control (Reed, & Luiselli, 2016). In fact, all four of these RFT relations are considered to occur in a manner that is consistent to operant processes (Barnes, 1994). Relational responding to separate stimuli might be one of the first pieces learned by verbal children, learning to mutually entail and combinatorically entail relations between stimuli is also likely to be directly shaped through differential reinforcement, and transformation of stimulus functions is also likely an operant process (Blackledge, 2003). It has been argued that during language acquisition through RFT processes, learners are essentially reinforced for responding to specific verbal stimuli as though they possess the stimulus functions of other, related stimuli (Blakcledge, 2003), and thus, these operant processes (relational responding, mutual and combinatorial entailment, and transformation of function) are shaped until all come under increasingly complex and specific contextual control (Blackledge, 2013). These emergent derived relations are advantageous to skill and language acquisition because they are time and resource-efficient means of teaching language. Derived Relational Responding seems to be the most comprehensive approach to understanding language and cognition while it remains conceptually systematic with behaviourism (Reed, & Luiselli, 2016).

Promoting the Emergence of Advanced Knowledge Relational Training System (PEAK; Dixon 2014a, b), is an assessment and curriculum guide designed to assess and improve language and cognitive skill deficits experienced by special populations is a modern

addition to the AVB literature (McKeel, Dixon, et al., 2015). It consists of a series of assessments and curriculum guides that does not stop short on basic language units but incorporates the traditional Skinnerian verbal operants (Skinner 1957) with contemporary behaviour analytic concepts such as Relational Frame Theory (Hayes et al. 2001; McKeel, Dixon, et al., 2015). The PEAK system capitalizes on the advances of RFT in understanding and promoting complex human behaviour (i.e.., language and cognition) as well as the efficiency of training due to Derived Relational Responding. (Reed, & Luiselli, 2016). The PEAK system consists of four modules, each including a separate 184-item criterion-referenced assessment and corresponding in-built curriculum programs. When taken together; these four modules have been noted to represent the apex of the science's current understanding of language development (Dixon, Belisle, et al., 2014).

The first PEAK manual, the PEAK Direct training manual (PEAK-D), is both a measure of the language and cognitive skills and a skill-based curriculum (Dixon, Belisle, et al., 2014). The assessment is comprised of 184 items, or skills, that are typically directly taught to children with and without disabilities. The skills assessed include prerequisite learning skills (i.e., eye contact and keeping hands still), vocal skills (i.e., imitating sounds and imitating words), writing skills (i.e., copying text and transcribing speech), conversational skills (i.e., telling jokes and answering questions), basic and advanced math skills (i.e., addition and subtraction), and advanced conversational skills (i.e., identifying metaphors and communicating uncertainty). Each of the 184 items that comprise PEAK Direct training module are listed on the assessment, with check boxes to indicate whether each item is in the participants' repertoire. The structure of the PEAK-D assessment alpha-numeric levels can be visualized as a triangle called the PEAK performance matrix. The first level of the performance matrix is at its highest point and contains two skills, 1A and 1B. Each descending level of the triangle has a greater number of programs as the triangle

expands (i.e., level 14 has 26 items) (Please see appendix 1). The PEAK-D training manual (Dixon 2014a), contains 184 curriculum programs that mirror the 184 assessment items of the PEAK Direct Assessment. Each program includes information relating to the program's goal, a list of stimuli typically used in the program, instructions on how to arrange and present stimuli, and a place for recording the stimuli used. PEAK programs are arranged in an alphanumeric order that corresponds to the PEAK-D assessment and thus can be used in conjunction with the assessment to determine appropriate programs for instruction. The PEAK-D training manual is designed to assess and teach language skills according to the traditional ABC design in which each response is reinforced in the presence of an appropriate discriminative stimulus (McKeel, et al., 2015).

This is then followed by the second PEAK Manual, the PEAK Generalization training manual (PEAK-G), which has an identical structure to PEAK-D of 184-point assessment and matching curriculum guide. The PEAK Generalization assessment expands upon the skills assessed by PEAK-D, by assessing the emergence of generalized responding. Generalized responding is a basic behaviour analytic process whereby individuals respond to untrained stimuli in the same way as they were taught to respond to trained stimuli that are formally similar (Dixon, Belisle, et al., 2014). Generally, typically developing individuals will only require training across a few exemplars of a stimulus before being able to correctly identify a variation of that stimulus that they have never seen before (Dixon, Belisle, et al., 2014). Thus, through generalization, individuals are able to expand their language repertoire without requiring Direct training of each skill (Dixon, Belisle, et al., 2014). The generalized skills that are assessed and targeted by the PEAK Generalization training manual (PEAK-G), are generalized motor imitation (i.e.., being able to perform any motor action after observing a model), vocal imitation (i.e.., being able to produce any vocal sound after hearing a model), categorization based on formal similarity (i.e., being able to label any variation of an animal

as a specific subcategory of animal), simple and complex mathematics (i.e. being able to add or subtract any two digit numbers), and complex language-based problem solving skills (i.e. using a logical syllogism to solve any simple logical problem) (Dixon, Belisle, et al., 2014). By targeting these skills PEAK-G aspires to compound upon the directly trained skills of individuals with autism by promoting the emergence of generalized learning. Items on the PEAK-G assessment are arranged identical to those on the PEAK-D assessment, along 14 alphanumeric levels and target the same receptive and expressive skills, as well as the verbal operants targeted by the PEAK-D assessment.

The subsequent PEAK Equivalence Module (PEAK E; Dixon, in press-a) and the PEAK Transformation Module (PEAK T; Dixon, in press-b) are concerned with learning through relations between stimuli. Although limited information about these manuals were available at the time of writing, due to their content being as yet unpublished by the author of the PEAK Relational training manuals, from the information available in newly published research articles the PEAK-E manual is described by that author as follows, "PEAK-E is comprised of an assessment of participants' relational abilities and a collection of 184 programs designed to progress participants' relational skills based on their results on the assessment."(Dixon, Speelman, Rowsey, Belisle, 2016). The PEAK-T manual is also described by the author as "[PEAK-T] provides a standardized curriculum and instructions for how to teach deictic and other relational skills" (Belisle, Dixon, Stanley, Munoz, & Daar, 2016).

Recently, the PEAK Direct training assessment has gathered considerable psychometric support as a valid and reliable assessment of the directly trained language repertoire of individuals with autism. Summaries, limitations and suggestions for future research of several research studies of interest across several domains of psychological research will be discussed below.

In study of convergent validity, Dixon, Whiting, Rowsey, & Belisly (2014) assessed the relationship between intelligence, as measured by IQ scores, and performance on the PEAK Direct training assessment with children with autism or other developmental and intellectual disabilities (*n*=50). Each child was administered the PEAK-D assessment from the PEAK Direct training manual. Scores from this assessment were compared to IQ scores for all participants to assess the relationship between the two measures. Results indicated a strong, significant correlation between scores on standardized IQ tests and scores on the PEAK-D assessment ($r = .759$, $p < .01$). The results demonstrated strong convergent validity and indicate that the PEAK may be a useful assessment and curriculum guide for training language and learning skills to individuals with autism and other developmental disabilities. Limitations of this research were as follows. First, the use of several IQ tests for comparison with the PEAK resulted in a comparison with a general measure of IQ and Peak as opposed to a specific measure of intelligence. Thus, more direct comparisons are warranted to exact the relationship between PEAK and particular IQ tests for the purposes of more accurate estimation and prediction. Secondly, IQ tests taken from the participants' records varied in the amount of time that had passed since their evaluation, with some participants having up to five years of time since their last IQ evaluation, which may have resulted in a score that is not fully representative of their current abilities. Despite this limitation, the authors reported that only 10 (20%) of the IQ scores collected were greater than 2 years old, and a strong correlation was found on both measures of ability nonetheless. Suggestions for future research calls for further examination of the psychometric properties of PEAK, including predictive validity of PEAK on outcomes such as educational and vocational achievement; investigation of the test-retest validity of PEAK and the internal consistency of the measure; an analysis of the requirements to implement the assessment and curriculum tools provided by the PEAK; and finally, replications of the current research should include other

populations, such as typically developing populations with average intelligence to assess the generalisability of the PEAK protocol.

Dixon, Carman, et al (2014) demonstrated in a study of convergent validity that PEAK-D correlated significantly with the Peabody Picture Vocabulary test (Dunn, & Dunn, 2007), and the Illinois Early Learning Standards Test (Illinois State Board of Education, 2013). Additionally, staff performances were evaluated on reliability of delivery of the PEAK-D assessment. In this study, children diagnosed with autism or related disorders (n=13) were exposed to an initial PEAK-D assessment designed to evaluate skill deficits within their repertoire, the Peabody Picture Vocabulary Test, and the Illinois Early Learning Standards Test. To evaluate staff performance, measures of Inter Observer Reliability (IOR) were calculated for all assessment results across 25% of students chosen at random from the sample. Results yielded significant positive correlations among the obtained PEAK-D assessment scores, the Peabody and the Standards assessments. Results for IOR measures was 96 % agreement for PPVT assessments, 96 % for Illinois Early Learning Standards assessments, and 85 % for scores on the PEAK-D assessment indicating high IOR for all measures. Implications for evidence-based discrete trial training based on the research findings were: 1) PEAK-D assessment represented an initial step toward an evidence-based behavioral assessment and training curriculum for verbal and academic skills for those with developmental disabilities. 2) PEAK may be a particularly useful tool in education settings or for families who wish to improve a child's verbal and academic skills. 3) PEAK-D assessment advances through much of the material and areas typically targeted in schools for development and has demonstrated reliability between novice assessors and trained behaviour analysts thus further suggesting that school personnel will be able to administer the PEAK-D assessment accurately to obtain a measure of a child's skills in the absence of formal, extended training. Some limitations of this study were the small sample size and the limited

age range of participants. However, despite the shortcomings of these limitations the data set accurately reflects the skill level of each individual participant, a feature that rapidly diminishes in a larger sample size. Suggestions for future research include studies of test-retest reliability, which would provide confidence in the stability of the assessment scores; gathering typically developing data, to provide a comparison to age matched typically-developing peers to facilitate age-appropriate goal setting and placement decisions; and examinations of how assessment scores converge with IQ scores to provide further confidence in completeness and meaningfulness of assessment scores.

In a two-phase study, Dixon, Belisle, Whiting, and Rowsey (2014) first administered PEAK-D to typically developing education students (*n*=208) to acquire normative data by which "specific cognitive and language deficits experienced by individuals with autism can be identified as they deviate from typical cognitive and language development" (p. 1598). Results indicated a strong positive relation between the PEAK-D total score and age. What the Researchers found was that as age increased, the PEAK-D total score rose equivalently until the age of approximately 8 years in the typically developing sample. In the second phase of the study, students with autism and developmental disabilities (*n*=94) received PEAK-D testing. However, contrary to the typically developing population's results, these results did not find any significant correlation between the PEAK-D total scores and age. The Researchers also found that, PEAK-D assessment total scores were significantly lower than the typically developing group. One relevant implication from this research is that typical developmental age norms can serve as benchmarks for targeted performances in children with autism and other developmental disabilities, as in similar measures of language ability. Furthermore, the data suggest that the PEAK-D can be applied with learners based on their language and cognitive functioning, independent of chronological age. Limitations of the current study were as follows. First, the

study had a small sample size relative to typically developing samples for predicting IQ. However, the authors justify this limitation by stating that PEAK was never designed as an intelligence test, instead, the main purpose of PEAK is to be an assessment and curriculum guide (Dixon, et al., 2014c). A second limitation reported was the relative homogeneity of the sample, which the authors state could be improved upon in subsequent investigations by using a demographically diverse population that matches the distribution of the participants with autism. Recommendations for future research suggest several ways in which the study could be extended. First, via the development of typically developing populations for the subsequent modules of PEAK, and second, future research must show how behaviour analytic interventions based on the PEAK program actually change via pre-post intervention analysis.

Rowsey, Belisle, and Dixon (2015) conducted a principal component analysis after administering PEAK-D to children (n=98) who had autism and other developmental disabilities. The results indicated that PEAK-D targets four factors related to language and learning abilities: foundational learning skills; perceptual learning skills; verbal comprehensive skills and verbal reasoning; memory and mathematical skills. Rowsey et al. viewed these findings as further empirical validation of an ABA-based instructional protocol, which also moves away from the traditional approached used in ABA to define behaviour operationally. Suggestions for future research were that it should be investigated if any differences exist from the current results using different samples, as there is a high probability that such a difference would exist; and that larger participant numbers were required in future replications as the research on sample sizes (Osborne, & Costelloe, 2004) states that larger samples tend to minimize the probability of errors, maximize the accuracy of population estimates, and increase the generalisability of the results from principal component analyses and other forms of statistical research.

As the first intervention-outcome evaluation of the PEAK-D, McKeel, Dixon, Daar, Rowsey, and Szekely (2015) included students with autism and other developmental disabilities (*n*=27) in a randomized experimental-control group design that lasted for 1 month. The control group participants (*n* = 13) received treatment as usual through special education services at their schools. Participants in the experimental group (*n* = 14) were exposed to five programs from the PEAK-D training manual curriculum that were selected based on their pre-intervention assessment scores. Two discrete-trial training sessions were implemented each week. Pre-intervention and post-intervention comparison scores indicated that participants in the experimental group had significantly more gains in language skills than the control group cohort. Limitations of this study were as follows. Firstly, although the findings indicated a statistically significant difference, this statistically significant difference between groups should be interpreted relative to the small sample size and do not necessarily translate to clinically significant outcomes. Second, the authors reported that the wide range of pre-experimental variables such as functioning level, presence of disruptive behaviour, and diagnosis may have also influenced statistical outcomes. Recommendations for future research are as follows. Future research should include a larger sample size and appropriate counterbalancing to reduce the potential influence of such nuance variability. Also, because this study took place across a very brief period, the frequency of tracking and reassessing the participants may not adequately reflect the application of the PEAK Relational Training System as it would be implemented in a completely applied setting. With regards to procedural fidelity, future research should include specific observations designed to measure the degree to which assessors and trainers remain faithful to the specified procedures throughout the course of the research and in-situ training provided where needed.

In Belisle, Rowsey, & Dixon, (2016), researchers evaluated the effectiveness of an in situ Behavioural Skills Training (BST) programme for improving staff implementation of the PEAK-D training manual. Members of direct care staff (*n*=3) were paired with students (*n*=3), all naive to PEAK and BST training, were provided with BST training in order to improve staff implementation of the PEAK-D training manual. The research data suggests that in situ BST is effective at increasing direct care staff implementation fidelity when conducting PEAK with participants with autism. All staff achieved a greater than 90% treatment fidelity score over three consecutive blocks following training, regardless of initial score or level of experience before training. Maintenance data further suggest that the staff were able to maintain strong implementation fidelity even when training was discontinued. Research limitations were as follows. First, the effectiveness of in situ BST was not compared to that of other staff-training models. Second, increasing the duration of the maintenance phase would have allowed for a stronger demonstration of the maintenance of staff implementation fidelity over time. Finally, inter-observer agreement (IOA) was not conducted on learner performance, as the data collected from the staff and IOA were conducted on data entry. Several recommendations for future research were made. First, it was suggested that one type of extension may involve using continuous data recording, rather than probe data, to allow for a more comprehensive evaluation of the effect of the training program. Future research may also evaluate how other staff-training programs compare to the procedures presented here, both at a single case and group design level, to ensure that participants are receiving training that can be considered best practice for the benefit of the clients who are served by the PEAK-D training protocol.

The interobserver reliability of the PEAK-D training protocol has ranged from 85% to 90% when scores between practitioners and a highly skilled behaviour analyst (Dixon,

Carman, et al., 2014) and between two trained implementers (Dixon, Whiting, Rowsey, & Belisle, 2014; Rowsey et al., 2015) are compared.

A central aspect of the PEAK protocol is staff implementation. From a review of the PEAK literature to date there is a clear emphasis on the importance of staff training and maintaining high levels of treatment fidelity (Belisle, Rowsey, & Dixon, 2016; McKeel, Rowsey, Dixon, & Daar, 2015; Dixon, Stanley, Belisle, & Rowsey, 2016; Dixon, et al., 2014; Dixon, Whiting, Rowsey, & Belisly, 2014). Despite this, a commonality in a majority of the existing PEAK literature is that the procedures were conducted by trained behaviour analysis graduate students (Dixon, Whiting, Rowsey, & Belisly, 2014). However, the PEAK protocol was designed with the intention that the procedures could also be implemented by behaviour therapists, direct care staff, or family members (Belisle, Rowsey, & Dixon, 2016). Given the emergent literature on PEAK, there may therefore be utility in evaluating the effectiveness of staff training procedures when conducted with staff that lack the formal training of behaviour analysis graduate students. Behavioral skills training (BST) is a staff training method that is behaviourally grounded, has resulted in strong empirical demonstrations of effectiveness (i.e., Nosik, Williams, Garrido, & Lee, 2013), and has been demonstrated to be an effective staff training method to accompany the PEAK-D training protocol (Belisle, Rowsey, & Dixon, 2016). Therefore, it is currently the training model of choice when training staff in the implementation of PEAK as it has provided consistent results of high treatment fidelity once applied in the literature (Belisle, Rowsey, & Dixon, 2016). However, these findings could be further explored through conducting training and implementation fidelity ratings with staff that have little experience with implementing ABA using only the information contained within the PEAK manuals within the BST model. This could then provide a further exploration of how effective the manuals are at training those with very minimal knowledge

of the tactics used in ABA such as the parents or care staff for which PEAK was also developed to be implemented by.

Another important aspect of the PEAK, which compliments this, is staff ratings of the social validity and acceptability of PEAK as a training protocol. Peak has a firm base of support for its reliability and validity, however, one form of validity which has yet to be measured and reported is social validity. *Social validity* is a term coined by behaviour analysts to refer to the social importance and acceptability of treatment goals, procedures, and outcomes (Foster, & Mash, 1999). Various researchers have made the assertion that there is a need in ABA to show that an intervention or treatment package will be accepted and viable if implemented in a community setting, that treatment addressed one or more meaningful or important problems in the client's life, and that it produced clinically important changes in those problems (Schwartz & Baer, 1991; Kazdin, 1977; Wolf 1978, 1978; VanHouten, 1979). Most current approaches (Gresham, & Lopez, 1996) follow the leads of the aforementioned researchers and define three distinct but related elements of intervention that can be assessed for their social validity: (a) the goals of treatment, (b) the treatment procedures, and (c) the outcomes produced by treatment procedures (Foster, & Mash, 1999). Goals can be assessed for both their importance (i.e.., what justifies working toward the particular treatment goals?) and their acceptability (i.e.., does "society" find the treatment goal to be worthwhile or desirable?). Treatment procedures are usually assessed solely for acceptability, whereas outcomes are assessed for their social importance (i.e.., does the degree of client change represent an important improvement for the client?) (Foster, & Mash, 2000). Schwartz and Baer (1991) described two ways in which social validity was important to the field of ABA. The first way, involved the advancement and survival of the field. They stated that it is important for to understand what programs are liked and disliked by the public so that behaviour analysts could continue to be a resource to that consumers would value. The

second way involved understanding why some programmes were liked and some disliked. They indicated that understanding why this occurred would eventually allow program developers to predict whether the programs they developed would be accepted or rejected. In 1968 when describing the dimensions of ABA Baer et al., made the distinction between applied and non-applied research, with applied being one of the core principles of ABA. They stated that applied research chooses target behaviours because of their relevance to, importance to and interest to society. This need to determine the degree to which society appreciates the methods and outcomes produced by applied research then influenced Wolf (1978) to define the need to measure the impact on society and has influenced several other to develop models and conceptualizations of social validity and its components. Schwartz, and Baer suggested that these evaluations of social validity be completed by potential clients, by significant others, by professionals and staff who are likely to implement the treatment, and by vocal spokespersons in the community as all these may be important in whether a treatment is likely to be implemented or accepted. As Carter (2010) detailed,

> In order for a treatment program to survive, it must not only be effective, it must also address problems considered relevant by consumers, it must do so in a manner that customers can tolerate or even possibly enjoy, and it must produce outcomes that are considered valued. (Carter, 2010, pp. 9)

Therefore, as no previous PEAK research has reported any of these elements there is a requirement to further explore them as their importance to the PEAK protocol cannot be understated.

Dixon and his colleagues have devoted extensive time and resources to PEAK research, producing sound preliminary findings that should promote additional studies (Reed, & Luiselli, 2016). Throughout the analyses of the available research it is clear that where

limitations or suggestions for future research are made, these are addressed in subsequent research and extensions to current research which again reflects the desire of PEAK researchers and researchers globally to further refine the PEAK training protocol and establish it as an empirical measure of language assessment and a subsequent teaching tool. However, from this review of the combined research published to date it is evident that several limitations still exist. First, there needs to be more demonstrations of intervention effectiveness (McKeel et al., 2015) across all modules of PEAK. Second, there is a need for further analyses that compare the PEAK-D training protocol with other ABA-based instructional methodologies, i.e. naturalistic teaching approaches (NTAs) (Rowsey, Belisle, Dixon, 2014; LeBlanc et al., 2006). Third, there is a need to establish if practitioners who are using the PEAK programme would rate it as socially valid or of practical use within applied settings, where material resources and excess time for arduous assessments are not always readily available. The final and overarching limitation of all the published PEAK research available at the time this review was conducted, is that the author of the PEAK Relational Training System manuals, Dr. Mark Dixon, has been involved in every study on PEAK that has been published to date. Although this limitation does not affect the validity of any outstanding PEAK research findings, it does highlight the need for PEAK-based research to be conducted by researchers that are unaffiliated with the author, as doing so will further legitimise research findings using the PEAK system.

The PEAK Relational Training System and the PEAK-D training manual have emerged as conceptually sound, psychometrically robust, and an innovative advancement of conventional ABA tactics for teaching children and youth who have autism and other developmental disabilities (Reed,& Luiselli, 2016). However, further research is required to address the key limitations previously discussed. Thus, the current research will aim to address these limitations across two studies. In Study One the research questions were: will

the implementation of the PEAK Generalization training Module result in the participants acquiring the targeted skills; will the acquisition of these skills result in an increase in participant's scores on the assessment for the Direct training Module without having received any Direct training via the PEAK-D training manual; will an assessment based on another form of ABA based teaching, Natural Environment Teaching (NET), detect the targets that were taught via the Generalization training manual and the Direct training manual; and will staff implementing the PEAK training protocols rate them as a social valid and acceptable protocol for training the target skills? Study Two explored: whether training targets from the PEAK Generalization training manual using a Multiple Baseline Design (MBD) resulted in related PEAK-D assessment targets being mastered without any exposure to the PEAK-D training protocol; whether PEAK-G mastered skills will maintain over time; and whether staff would still display a high level of implementation fidelity 6 months after receiving their initial training.

Chapter Two

Study One: Exploring the Implementation of PEAK in a Practical Setting

The current research will seek answers for several key questions regarding the implementation of the first two training modules, the Generalization Training Module (PEAK-G) and the Direct Training module (PEAK-D) taken from the Promoting the Emergence of Advanced Knowledge (PEAK) Relational Training System. Answering these questions has been highlighted by previous research (i.e. Belisle, et al., 2015; Dixon, et al., 2014c; Reed, & Luiselli, 2016), to be important for expanding the available knowledge base of PEAK and doing so will provide an initial exploration of various issues involved in implementing PEAK within a practical setting. To answer these questions involved implementing a series of interventions and assessments with the final aim of providing an initial empirical exploration of the implementation of PEAK training protocols in a practical setting, across a typically developing sample, a sample with a clinical diagnosis of Autism Spectrum disorders (ASD), and a sample of staff members.

Study One will consist of three interventions that will explore: (a) will the implementation of the PEAK Generalization training Module result in participants gaining the targeted skills, and will this result in increased participant scores on the assessment for the Direct Training Module (PEAK-D) without any exposure to Direct Training via the PEAK-D training protocol; (b) will an assessment based on another form of ABA-based teaching, Natural Environment Teaching (NET), detect the targets that were mastered via the PEAK-G training and the PEAK-D training; and (c) will staff implementing the PEAK training protocols rate them as a socially valid and acceptable protocol for training the target skills?

Intervention One sought to explore whether implementation of the PEAK-G training protocol would result in participants acquiring the targeted skills, and whether the acquisition of these skills would then result in an increase in participant's scores on the PEAK-D assessment without any exposure to the PEAK-D training protocol. As stated, sought to expand upon the findings reported in McKeel et al., 2015, where a Randomized Controlled

Trial of the PEAK-D Training module was conducted with children with autism. In this study, the results supported the researcher's assertion that exposure to the curriculum portion of PEAK-D was functionally related to post-intervention score increases on the PEAK-D assessment. However to date, PEAK researchers are not able to clearly state that these same results will hold true for the remaining PEAK training modules, such as the Generalization Training Module, as no published data currently documents group or individual child performances after intervention. Conceptually, this change in scores would be demonstrated via pre- and post- intervention PEAK-G assessments, given that PEAK-G is based on the same behavioural principles as PEAK-D. However, PEAK-G training involves a slight departure from the traditional discrete trial training (DTT) ABA teaching approach utilised in the PEAK-D training protocol. Instead, PEAK-G training utilises a train/test approach, where learners are exposed to randomly alternating training trials (using traditional DTT methods to promote generalization) and test trials, where novel stimuli are presented under assessment conditions to prove for the emergence of independent generalization without any feedback, and thus whether the same pre- post- intervention score increase will hold true remains to be explored. Therefore, in Intervention One, full PEAK-G assessments were conducted, both pre- and post- the PEAK-G training intervention, with twelve participants, five with diagnosed autism, and seven which were typically developing, to explore whether individual PEAK-G assessment scores changed following PEAK-G intervention, which would imply that participants had been trained to exhibit generalisation in areas where they were previously found to not. This would provide valuable data for future researchers considering the use of the complete set of PEAK training protocols as it would demonstrate empirically whether the delivery of PEAK training would result in the acquisition of the targeted skills regardless of whether the module uses the more traditional DTT approach or the novel PEAK-G train/test strategy.

Another aspect of Intervention One was to explore whether exposure to the PEAK-G Training protocol increased participant's scores on the PEAK-D assessments without any exposure to the PEAK-D training protocol. In the current PEAK training protocols, the authors suggest running PEAK-G concurrent with PEAK-D (Dixon, 2014b), as PEAK-D targets prelearner skills that the more advanced PEAK-G builds upon and trains into natural generalization (Dixon, 2014b). However, no previous research had been conducted on whether it would be possible to only train the more advanced PEAK-G without the PEAK-D and still increase PEAK-D assessment scores. If the results indicate that PEAK-G will increase PEAK-D skills without PEAK-D being directly implemented, then this could possibly decrease the time necessary to implement PEAK in cases or settings where factors such as time and available resources are limited. Thus, Intervention One also involved conducting the full PEAK-D assessment both pre- and post- the implementation of the PEAK-G intervention with all twelve participants in order to determine if there is an increase in scores on the PEAK-D assessment following PEAK-G training.

Intervention Two sought to build upon the results from Intervention One. Having already been assessed using the PEAK-G and PEAK-D assessments, trained via the PEAK-G training intervention, and assessed again using the PEAK-G and PEAK-D assessments post the PEAK-G training, participants in Intervention Two then received PEAK-D training to criterion for five PEAK-D targets each. This was done to both partially replicate the effect reported in McKeel et al., (2015), with both ASD and typically developing participants, and to allow a more detailed examination of PEAK-G and PEAK-D. Thus, inafter receiveing PEAK-D training, mastered PEAK targets from all participants for both PEAK-G and PEAK-D modules were then assessed to determine if any compatibilities existed between the skills taught to mastery through PEAK training and another form of ABA-based teaching which also focuses on cognitive and language development: Natural Environment Teaching (Stokes,

& Baer, 1977). After participants completed their Peak training in both PEAK-G and PEAK-D modules they were administered a Natural Environment Teaching assessment in order to determine if the skills which they had mastered from PEAK had also generalised and were being emitted by the participants in their natural environment. This was important because if found to have occurred, this could result in increased access to various new forms of reinforcement which were previously unavailable due to a lack of the necessary skill repertoires. This assessment consisted of taking a list of the PEAK skills which the child had demonstrated as mastered in the structured instructional setting and then assessing if these skills are emitted by the participant in the appropriate natural context when this context was presented. The results of this assessment may determine if PEAK training has resulted in outcomes which could be measured and elicited using another form of teaching.

Intervention Two aimed to directly address the suggestion for future research highlighted by Rowsey, Belisle, & Dixon, 2014, which called for examination of how various ABA approaches and treatments compare to PEAK through empirical investigation. This type of comparison would not only benefit the PEAK literature greatly, as yet another possible successful comparison between the PEAK system and another ABA approach, but as PEAK is based in the field of applied verbal behaviour (AVB) and approaches such as NET are based in the field of naturalistic teaching approaches (NTAs) an exploration of the results might yield interesting findings regarding similarities and differences between these approaches, such as those found in LeBlanc, et al., (2006) where researchers found an already existing compatibility between the AVB approach and NTAs.

Intervention Three sought to explore whether the staff implementing the PEAK training protocols rate them as a socially valid and acceptable protocol for training the target skills. As detailed in Chapter One, treatment programs which are based on the principles of ABA have a responsibility to not only be effective, but must also address problems which are

considered relevant by consumers in a way that is tolerable or enjoyable, and must produce outcomes which are considered valued by the society or community in which they occur (Carter, 2010). As a science, ABA has seven core criteria which every study requires to be considered an example of applied behaviour analysis (Carter, 2010). The first of these dimensions is that, for research to be considered an example of ABA, it must be applied (Baer, Wolf, & Risley, 1968). This criterion refers to how behaviours are selected for intervention. In applied research target behaviours must be selected because of their importance to society, whereas in non-applied research, target behaviours may be chosen for reasons that provide convenience to the researcher (Carter, 2010). This concept of choosing target behaviours which are highly relevant to society and that typically reveal immediately important relationships between the behaviour and certain stimuli forms the basis of the concept of social validity. *Social validity*, as defined in Chapter One, is a term coined by behaviour analysts to refer to the social importance and acceptability of treatment goals, procedures, and outcomes (Foster, & Mash, 2006). Currently, there has been no research conducted, to the knowledge of the researcher, which has explicitly examined whether staff who are implementing the PEAK protocols would rate it as a socially valid treatment package. Thus, the rationale for this intervention was to train staff in the implementation of the PEAK-D and PEAK-G assessments and training protocols, then let staff implement both the PEAK-D and PEAK-G training, measure the level of treatment implementation fidelity and integrity at two points in time, and to subsequently ask that staff anonymously rate the overall protocol using a measure of social validity to provide an exploration of PEAK as a socially valid assessments and training packages. This intervention ran parallel with Intervention One and Two as staff received their implementation training prior to the commencement of Intervention One.

Method

Participants

Twelve children, four girls and eight boys aged 4-5, were recruited from a typically developing and an ABA class within the same Montessori pre-school. Five of the children, one girl and four boys, had previous diagnoses of Autism Spectrum Disorder by clinical psychologists independent of the current research in accordance with the diagnostic criteria set out in the Diagnostic and Statistical Manual of Mental disorders (5th ed.; *DSM–5*; American Psychiatric Association, 2013). Of the five children with ASD two participants, both boys, were considered pre-verbal and had very low-levels of expressive language. However, as the PEAK protocol was developed for use with children that have autism, research has indicated that this may not interfere with their progress, but simply limit the number of skills which could be targeted (Dixon, 2014a). Of the remaining seven participants, three girls and four boys, there were no formal diagnoses for developmental or intellectual disabilities and these participants formed the typically developing sample for the study. Of the typically developing children one girl and three boys were learning English as a second language throughout the pre-school year. This only presented as a difficulty with the girl, as her expressive English was still in the early stages of developing. Although no formal assessment of her language abilities existed at the time of participation, her receptive English was determined by the researcher and her class teacher to be of a sufficient level that she could participate without any negative effects on her progress. Due to this, similar provisions were made when selecting her programme targets as those made for the other pre-verbal participants in the ASD sample. Participants remained the same throughout experiment one. A total of six staff participants, four female and two males aged 20-45, were also recruited from the teaching staff at a Montessori School. All participants have a primary degree in either education or psychology and limited experience with ABA practices but extensive

experience in the Montessori style of teaching. Staff participants work across a typically developing class and a class of children with ASDs.

Throughout Study One, five children with ASD and seven typically developing children completed the PEAK-D assessment, PEAK-G assessment, PEAK-G training, and a follow-up PEAK-G and PEAK-D assessment; four children with ASD and six typically developing children completed the PEAK-D training and NET assessments; and six teaching staff completed staff training and the assessment of social validity.

Ethics

It was not anticipated that participants would suffer any physical discomfort as a result of participation and children were appropriately supervised and monitored throughout the procedures in accordance with usual teaching regimes used at the school. All procedures were conducted in-class and with the consent and supervision of the school Director and staff. The researcher is trained in the principles and application of ABA and was therefore competent to work with participants, and participants were provided with frequent short breaks and positive reinforcement throughout procedures. This was in compliance with code 1.02 of BACB ethical codes, code 2.2.3 PSI and code 2.01 of APA whereby the researcher should practice within the limits of their own competence.

Verbal assent was sought from each child before commencing each session and was obtained by asking participants if they would like to work with the researcher or if they would prefer to continue working with another teacher on academic goals. Participants were free to respond by opting to work with the researcher or continuing with other Montessori work. This is in accordance with ethical principles 10.04 (BACB), 1.3.3 PSI and 3.10 APA of consent and informed consent. Furthermore, written consent was obtained from the parents/guardians of the children involved and parents/guardians were made aware of the

type of procedures involved throughout the current research, and were also made aware that participation could be withdrawn at any point, 10.10-10.13 (BACB), 1.3.4 (PSI) and 8.02 (APA). Written consent was also obtained from the Directors at the school prior to the commencement of the research.

The researcher was present and available throughout all PEAK procedures and it was planned that procedures were to be terminated if children showed signs of distress. Physical indicators of distress were defined as increased stereotypy or other problem behaviour, or verbalised dislike of procedures. Prior to commencement of the researcher it was determined that there were no anticipated risks in participating apart from those that are incurred in the normal school day for participants. However, research has shown that children diagnosed with Autism Spectrum Disorders (ASD) display far more maladaptive and problem behaviours than their typically developing peers, with as many as one third displaying clinically significant problem behaviours (Hartley, Sikora, & McCoy, 2008). Therefore, any research that involves participants from a population of children diagnosed with ASD must consider the risks involved with problem behaviours. Thus, during the course of this research every effort was made by the researcher and anyone involved in the research to control for the risks involved in the occurrence of problem behaviours such as self-injurious behaviours and escape/aversive behaviours. This was in compliance with code 2.03 of Behaviour Analyst Certification Board (BACB) code of ethics, code 3.1.4 of the code of professional ethics of the Psychological Society of Ireland (PSI) and code 3.04 of ethical principles of psychologists and code of conduct of the American Psychological Association (APA) whereby the researcher must protect the wellbeing and health and safety of the research participant at all times. Sessions were conducted during school hours usually once a day for ASD participants, and twice per week for typically developing participants. Duration of individual sessions was never more than 30 minutes. The higher frequency of teaching

sessions with ASD participants was considered justifiable because the PEAK targets for teaching corresponded with those in the children's Individualised Educational Programme and teaching schedules for these.

For staff participants, the major ethical considerations were regarding informed consent and anonymity of data collected from staff participants. All staff recruited into the current study signed informed consent forms. These forms included; relevant contact details, qualifications of researcher and supervisor, a detailed outline of expected duration of participation, procedures set in place for data collection, protection and how to request the release of their individual participant data. Informed consent forms were also accompanied by information sheets which describe the research in greater detail. Staff were asked to confirm that they have read the information sheet on the informed consent form. Staff were also informed that participation was on a voluntary basis and that no penalty would be exacted for non-participation.

With the ever-growing importance of data protection, the following ethical provisions were made to ensure absolute data protection throughout the research process. All fidelity data gathered from the PEAK assessments, PEAK training and subsequent assessments were gathered under pseudonyms, and the link between the pseudonyms and the participant's identities was destroyed once the raw data was entered into the researcher's computer. Participants were informed that once their raw data had been processed in this way that retrieving individual data would not be possible as data would have been completely anonymised. All assessments of social validity were completely anonymous. All participants were given a copy of the assessment as well as brief instructions in an unmarked white envelope. Once the assessment was completed participants were instructed to place it back in the envelope and to post the envelopes in a specially designated post-box style box. These

provisions were to ensure that all social validity assessment data are completely anonymous to prevent any unintended observer effects or skewed results due to bias.

Setting

All training and assessments took place in the participant's classroom or on a one-to-one basis in an empty classroom with a trained member of staff or observer present at all times. Training sessions were no longer than 30 mins per session and assessment sessions ranged from 30 mins to 120 mins. Various stimuli needed to run the training programs and administer the assessments were gathered from the classrooms as per the PEAK manual's instructions (i.e., picture vocabulary cards, toys, common items, and blocks). Arrays of preferred edible and tangible items were selected for each individual based on indirect assessment with classroom staff who frequently worked with that child. A brief preference assessment was also conducted at the beginning of each session by presenting the array of available items to the child and asking the child, ''What do you want to work for?'' The item selected by the child, either vocally or through gesture, was used as a reinforcing stimulus for the remainder of the session. Staff training and implementation of the PEAK protocol took place across two classrooms where staff teach.

Materials and measurements

Peak: direct training assessment and curriculum.

The PEAK Direct Training assessment and subsequent training programs as found in the PEAK Relational Training System: Direct Training Module (Dixon, 2014a) were implemented with all participants. Please refer to Chapter One for a detailed description of the PEAK Direct Training assessment and curriculum. During the assessment phase all six of Skinner's verbal operants are assessed (i.e. mands, tacts, intraverbals, echoics, textual and

transcription responses), as well as prerequisite learning skills (i.e., eye contact and keeping hands still), vocal skills (i.e., imitating sounds and imitating words), writing skills (i.e., copying text and transcribing speech), conversational skills (i.e., telling jokes and answering questions), basic and advanced math skills (i.e., addition and subtraction), and advanced conversational skills (i.e., identifying metaphors and communicating uncertainty). Hence, as the curriculum's training programs and the assessment are linked, these skills are then targeted throughout the implementation of the training programs.

PEAK: Generalization training assessment and curriculum.

The PEAK Generalization Training assessment and subsequent training programs as found in the PEAK Relational Training System: Generalization Training Module (Dixon, 2014b) were also implemented with all participants. Please also refer to Chapter One for a detailed description of the PEAK Generalization Training assessment and curriculum.

Natural Environment Teaching Assessment.

An assessment based on techniques employed during Natural Environment Teaching was formulated by the researcher prior to commencement of PEAK training. To date, there does not appear to be any empirical assessment of skill generalization apart from the PEAK PEAK-G assessment. Thus, the researcher formulated an assessment of skill generalization under the supervision of a Board Certified Behaviour Analyst (BCBA) with extensive experience with the implementation of NTAs such as NET. The assessment featured a list of all the targets each participant had mastered during PEAK-D and PEAK-G training and followed the basic structure of the PEAK assessment protocol to maintain the consistency and accuracy of assessment results.

Social Validity and Acceptability Assessment.

Staff completed an anonymous assessment of the social validity and treatment acceptability levels of the PEAK protocols as assessment and intervention tools packages. This assessment was based on the Intervention Rating Profile-15 (IRP-15) by Martens, Witt, Elliot, & Darveaux (1985) which was modified, with the permission of the original researchers for the purpose of this research. The assessment consisted of 15 items which were rated using a six-point Likert-type rating scale and ranged from 1(*strongly disagree*) to 6 (*strongly agree*) and asked that staff participants also to identify which sample, whether ASD participants, typically developing participants, or both samples of participants, they had implemented the PEAK protocols with by ticking the appropriate box. During the assessment staff were asked to circle whichever response they agreed with most and each staff participant completed a separate assessment for both PEAK-G and PEAK-D. Total scores were obtained by summing all items, with higher summed scores indicating greater levels of acceptability. This instrument has an internal consistency of 0.98, with research by Von Brock, and Elliot (1987) reporting that a total mean score rating of 58.7% or above being indicative of an acceptable level of perceived social validity by participants.

Staff fidelity measure.

Implementation fidelity was assessed using a modified version the PEAK Implementation Checklist (PEAK-IC; see Appendix 2). Items on the original PEAK-IC are taken directly from the PEAK-D training manual's instructions on how to run PEAK-D trials accurately (Dixon, 2014a). However, to also assess the implementation fidelity of PEAK-G the original PEAK-IC items were modified to include the train/test protocol used in PEAK-G where train and test trials are randomly alternated within training blocks.

Experimental design and measurements

Study One employed a within-participants AB design comparing pre- and post-intervention scores across three interventions for both typically developing and ASD participants. The independent variables (I.Vs) and dependent variables (D.Vs) differed for each intervention. For Intervention One, the independent variable was the implementation of PEAK Generalization training for all participants on selected PEAK-G targets based on the PEAK-G assessment outcomes for each participant, please see table 1 for a full list of PEAK targets for each participant. Operants targeted during Intervention One included tacts, intraverbals, textual, and transcription responses. The D.Vs were the total PEAK-G and PEAK-D assessment scores pre- and post- the application of the I.V, i.e. did PEAK-G training raise participant's PEAK-G and PEAK-D assessment scores, without participant exposure to the PEAK-D training protocol. Both PEAK-G and PEAK-D assessments targeted mands, tacts, echoics, intraverbals, textual, and transcription responses, in conjunction with autoclitics, (please see Chapter One for descriptions of these verbal operants).

In Intervention Two, the I.V was the application of the PEAK-Direct Training intervention to increase the current participant repertoire of PEAK-D skills, (please see table 1 and 2 for full list of targets), and the D.V was the NET based assessment to determine if generalization of mastered PEAK skills post the application of the I.V had also occurred to the appropriate natural environment (please see table 1 and 2 for a list of the D.V targets for participants). Operants targeted during both PEAK-D training, as well as the NET assessment included tacts, intraverbals, textual, and transcription responses. For Intervention Three, the I.V was the implementation of staff training in administering the PEAK protocols, and the D.Vs were staff scores of PEAK as socially valid and acceptable, and staff treatment fidelity scores at 6-month follow-up. Analysis of data were primarily within-participant; as a small *n* design precluded a valid within-group comparison and limited generalisability of findings.

Table 1

Natural Environment Assessment targets (mastered PEAK skills) for typically developing participants

	NET Targets	
Participant	PEAK-G Targets	PEAK-D Targets
1	8C Metaphorical Tact Extensions	14A Addition
	10D Transcription: Connecting the Dots	10R Tact Planet Names
	5G Generalized Patterning	10B Receptively Label Coins
	12L Delayed Receptive Picture ID	10L Tact Coins
	4E Matching Numbers and Letters	11M Tact Values of Coins
2	8C Metaphorical Tact Extension	10B Receptively Label Coins
	5G Generalized Patterning	10L Tact Coins
	4E Matching Numbers and Letters	11M Tact Values of Coins
	10D Transcription: Connecting the Dots	10R Tact Planet Names
	3D Receptively ID Shades of Color	14A Addition
3	8C Metaphorical Tact Extension	14A Addition
	5G Generalized Patterning	10R Tact Planet Names
	4E Matching Numbers and Letters	10B Receptively Label Coins
	10D Transcription: Connecting the Dots	10L Tact Coins
	11B Picture Sequences with Delay	11M Tact Values of Coins
4	Metaphorical Tact Extension	9F Receptively Label Letters
	5G Generalized Patterning	9O Tact Letters
	4E Matching Numbers and Letters	9G Receptively Label Numbers
	10D Transcription: Connecting the Dots	9P Tact Numbers
	3D Receptively ID Shades of Color	14A Addition
5	4E Matching Numbers and Letters	10B Receptively Label Coins
	5G Generalized Patterning	10L Tact Coins
	10D Transcription: Connecting the Dots	10R Tact Planet Names
	6E Receptively ID Shapes	11M Tact Values of Coins
	11B Picture Sequences with Delay	14A Addition
6	6E Receptively ID Shapes	14A Addition
	4E Matching Numbers and Letters	10R Tact Planet Names
	10D Transcription: Connecting the Dots	10B Receptively Label Coins
	11B Picture Sequences with Delay	10L Tact Coins
	5G Generalized Patterning	11M Tact Values of Coins
7	6E Receptively ID Shapes	-
	5G Generalized Patterning	-
	11B Picture Sequences with Delay	-
	10D Transcription: Connecting the Dots	-
	12L Delayed Receptive Picture ID	-

**Note. Participant 7 did not partake in Intervention Two, hence no PEAK-D targets were selected.*

Table 2

Natural Environment Assessment targets (mastered PEAK skills) for ASD participants

	NET Targets	
Participant	PEAK-G Targets	PEAK-D Targets
1	2B Tact: Non-Identical Animals	11A Receptively Label Sensory Feelings
	6E Receptively ID Shapes	11F Delayed Picture Identification
	5G Generalized Patterning	11B Receptively Label Coin Values
		11C Receptively Label Rooms
	10D Transcription: Connecting the Dots	
	11B Picture Sequences with Delay	11E Receptively Label Body Part Function
2	10D Transcription: Connecting the Dots	10B Receptively Label Coins
	6E Receptively ID Shapes	10C Receptively Label Actions
	3D Receptively ID Shades of Color	11A Receptively Label Sensory Feelings
	5G Generalized Patterning	10M Tact Actions
	11B Picture Sequences with Delay	11B Receptively Label Coin Values
3	2B Tact: Non-Identical Animals	11A Receptively Label Sensory Feelings
	3D Receptively ID Shades of Color	11F Delayed Picture Identification
	4C Exclusion: Feature	11B Receptively Label Coin Values
	10D Transcription: Connecting the Dots	11C Receptively Label Rooms
	5G Generalized Patterning	11E Receptively Label Body Part Function
4	6E Receptively ID Shapes	8A Receptively Label Toys
	4E Matching Numbers and Letters	8E Receptively Label Clothing
	7H Problem Solving: Packing a Container	8C Receptively Label Food
	3D Receptively ID Shades of Color	8B Receptively Label Common Items
	13H Receptively ID Non--- Identical Vehicles	8D Receptively Label Animals
5	13H Receptively ID Non--- Identical Vehicles	-
	7H Problem Solving: Packing a Container	-
	3D Receptively ID Shades of Color	-
	4E Matching Numbers and Letters	-
	6E Receptively ID Shapes	-

**Note. Participant 5 did not partake in Intervention Two, hence no PEAK-D targets were selected.*

Procedures

Staff training.

Study One commenced with approximately one hour of individual staff training for all staff members who were going to be directly assessing and implementing PEAK-G and PEAK-D. Training was conducted under the supervision and guidance of a Board-Certified Behaviour Analyst (BCBA) by an experienced PEAK instructor and the primary researcher across the two classrooms in which the PEAK assessments and training were going to be implemented. Training followed the four steps of BST – instructions, modelling, rehearsal, and feedback. The researcher and experienced instructor encouraged staff to ask questions during training and feedback and answered any questions throughout training, as well as during in-situ implementations. Training consisted of asking staff to read a 30-page excerpt taken from the introductory chapters of the PEAK-D training manual, supplemented with information on the train/test strategy from the PEAK-G training manual, a week before the training session. This provided staff with adequate time to become familiar with how to run PEAK and aided in reducing the amount of time required during training to explain the basic concepts and processes of PEAK. After a brief recap of the main points contained within the excerpt, training then progressed onto didactic instruction, modelling and demonstration of how to accurately conduct the PEAK-D and PEAK-G assessments and implement programmes from both curricula. These steps were followed by feedback during role-play and rehearsal of chosen PEAK-G and PEAK-D targets. Another important aspect of staff training was the use of the PEAK IC. The PEAK IC was used to measure if staff were implementing the various elements of PEAK correctly during modelling for research purposes, and served as a tool for providing accurate feedback on implementation fidelity throughout modelling and feedback sessions.

PEAK assessments.

Both the PEAK Direct Assessment and the PEAK Generalization Assessment are 184-item, criterion referenced, subtests of the PEAK curricula designed to assess an individual's ability to learn and respond to verbal stimuli. Both assessments were laid out in identical fashion and assessed using the same methods. The only difference between assessments was the skills being assessed. Intervention One commenced with the PEAK-D assessment and followed with the PEAK-G assessment.

The Direct Training assessment specifically evaluates an individual's ability to learn language skills through direct contingencies (i.e.., through reinforcement of specific verbal responses). Each item begins with the examiner providing an instruction along with relevant stimuli. The participant was allowed up to 3s to respond. If the participant responded correctly, a ''yes'' was recorded for that skill; incorrect responses were recorded as a ''no.'' If a response was not made within 3s or a disruptive behaviour was emitted, a block of 10 trials was run; if the child was able to correctly respond to nine of 10 trials, a ''yes'' was recorded for that item. Because disruptive behaviour or non-attending on the part of the participant can lead to false positives and false negatives, the 10-trial probe sequence is recommended by the PEAK authors in order to ensure that the particular skill being evaluated is consistently present in the learner's repertoire (Dixon, 2014a). For example, If the assessor asked the participant to receptively point to a colour but it was determined the participant was not looking at the array presented; the assessor would then further assess the skill using the 10- trial probe. This would eliminate the chance of the participant correctly guessing nine out of ten trials without looking. PEAK-D assessment scores range from 0 to 184 and were calculated by adding the total number of items responded to correctly (Dixon 2014a). Reinforcement was not provided during assessment sessions for correct responding, but was provided to participants based on compliance with

the assessment procedure. Specifically, participants were provided access to the chosen reinforcer after every 10 trials to prevent frustration from lack of reinforcement.

The PEAK-G assessment differs from the PEAK-D assessment in that it focuses on the learner's ability to extend learned responses to similar but non-identical stimuli. For the PEAK-G assessment, assessments were administered and total scores calculated in the same way as the PEAK-D assessment. As in the PEAK-D assessment, reinforcement was again not provided during assessment sessions for correct responding, but was provided to participants on the basis of compliance with the assessment procedure. Specifically, participants were again provided access to the chosen reinforcer after every 10 trials to prevent frustration from lack of reinforcement.

PEAK Generalization training.

For each ASD and TD participant, five programs from the PEAK-G curriculum were selected based on the individual's results from the PEAK-G assessment. For the ASD group, programs were selected from the alphanumeric assessment items (i.e. 1A, 2A, 2B, 3A, 3B) that the participant incorrectly responded to during the PEAK-G assessment. However not all targets were selected this way, and those that corresponded with goals highlighted in participant's individualised education plans (IEPs) were given preference, as some of the skills tested in the PEAK-G assessment were considered too far beyond their existing repertoires to teach effectively but occurred at a low alphanumeric level, for example programme 3A. Flexible Textual Behaviour requires that participants be able to find words in any direction when presented with a word search (Dixon, 2014). This programme would have been developmentally inappropriate to teach pre-schoolers with ASD. In the typically developing group, programs were selected from the alphanumeric assessment items that the participant incorrectly responded to during the PEAK-G assessment and that corresponded with age-appropriate Montessori goals using a similar rationale as previous.

Training sessions were conducted one to five times a week and consisted of discrete trial training using the train/test strategy in which the learner is taught to generalize skills using a specified set of stimuli but tested for mastery using a different set of stimuli (Dixon, 2014b). This is an essential piece of the PEAK-G protocol because as previously stated, generalization is a type of learning that is demonstrated when a learner is presented a novel stimulus and subsequently presents an appropriate response (Dixon 2014b). In the train trials, a discrete trial begins with the presentation of a discriminative stimulus or question. The individual was then allowed up to 3s to respond. If a correct response was emitted, a reinforcing consequence was provided (i.e., praise, edible, and preferred activity). If an incorrect response or no response was emitted, a series of prompts were presented in order to evoke the appropriate response. The prompting sequence and prompt-level decisions followed the PEAK prompt scoring system. The PEAK system uses a novel prompt-based scoring system when running PEAK training. This specialised scoring system was designed to allow for the detection of increases in independent responding and subtle changes in prompts required for a participant to emit the target response during training. Whereas traditional ABA data collection frequently only records correct or incorrect responses regardless of the prompt level used, the PEAK scoring system allows the implementer to measure the level of prompt necessity in a quantifiable way by specifying scores for different levels of prompting, please see figure 1 for an example of this scoring system. During training sessions, as each trial is presented, the level of prompt that was required to achieve a correct response is recorded by circling the corresponding number on the data sheets. As previously stated, during the test sessions employed in the PEAK-G training, prompt scoring reverts back to the traditional ABA scoring methods where a response is either emitted independently or is marked as incorrect. PEAK recommends that if participants are struggling to make progress in their training programs, researchers should adjust the complexity of the stimulus array as well as start at a prompt level that is more suitable for the individual. These

adjustments are then faded back as the participants makes progress during their training sessions. In the test trials a discrete trial also begins with the presentation of a discriminative stimulus, and the individual then allowed 3s to respond. However, although test stimuli are generally similar to the Train stimuli, in that they share the relevant features that the program specifies; they are never presented with feedback for incorrect or correct responses. This lack of feedback during test trials is to ensure that the learner's responses are solely under the control of a generalization and not from Direct Training with the stimuli used. The PEAK-G manual leaves it to the practitioner to decide how to split train and test trials. Either the practitioner can have one trial block of 10 train trials followed by one trial block of 10 test trials, or 5 train and 5 test trials can be randomly interspersed throughout the same trial block consisting of 10 trials. This is due to the fact that in certain populations, such as those with ASD, not all learners would be able to transition without problem behaviours from a trial block of 10 train trials with accompanying reinforcement, to a block of 10 test trials with no reinforcement provided whatsoever. Thus, in the mixed format reinforcement may not be delivered with every trial in a block of 10, but there is guaranteed delivery of reinforcement for at least half of the trials in the trial black. For the purposes of this research this mixed format of 5 train trials and 5 test trials per trial block was utilised.

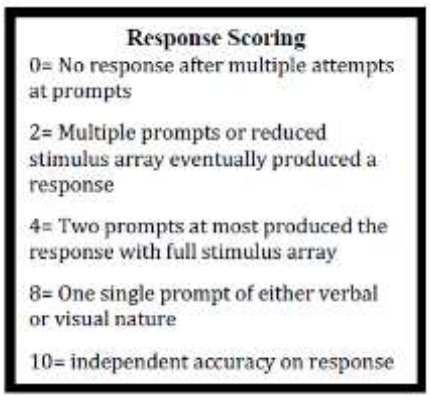

Figure 1.

PEAK Response Scoring System.

Throughout training, each participant was required to respond to at least one trial block for each program assigned per session. Trial blocks consist of 10 trials covering five train targets and five test targets (as specified in the program). These were presented as 10 consecutive trials from each program. Mastery criteria for all programs were set at 90 % for two consecutive trial blocks as per the recommendation in the PEAK-G manual. When a program was mastered, the trainer would remove that program from the participant's current programmes section of their folder and move the programme to where the participant's mastered skills are kept. Mastered programmes may be revised throughout the day for maintenance of skills over time. If participants responded correctly to 100% of trials on the first presentation of a program, that program was considered mastered. Session lengths varied per participant performance (i.e.., participants who frequently responded correctly completed the session more quickly). Each participant was exposed to up to a minimum of three trainers throughout the training phase of the intervention. Participant's progress during the training phase was monitored on a frequent ongoing basis based on their PEAK scores during trial blocks in order to identify when modification to programs would became necessary due to a lack of progression. The PEAK-G manual presents two options on how to conduct program revisions Continuous Assessment or Periodic Assessment. Continuous Assessment involves the programme implementer being present and replacing programmes as learners reach criterion, whereas, Periodic Assessment involves an outside manager assessing progress during scheduled sessions and determining if targets should be altered at that time, (please see Dixon, 2014b, pp. 41-42 for further details on these).

As the researcher was present for all PEAK trials the Continuous Assessment procedure was used in this intervention to monitor for participants that were experiencing difficulty with progressing in their PEAK-G targets. Where it was determined that a lack of progress was occurring due to the level of difficulty of the programme, this was addressed by

simplifying the program. This involved various strategies such as using less complicated stimuli, reducing the number of test trials per trial block, re-evaluating reinforcers, or revising prompting procedures during trials. In extreme cases where these changes did not make a difference, the program was suspended until the related pre-requisite skill was mastered. The complete suspension of programmes only occurred on three separate occasions throughout the course of the research after it was determined, under guidance from the supervising Board Certified Behaviour Analyst (BCBA) supervisor, that the pre-requisite skills required for the two participants to make progress in these programs were not in the participant's repertoires. Thus, three programmes were suspended in total, contingent upon the participants first mastering the required pre-requisite skills required to progress and reach criterion on these programs. Due to time constraints these programmes were not continued within this research.

PEAK Direct Training.

As this intervention commenced immediately after the PEAK-G training intervention was completed, the PEAK-D training made use the post-intervention PEAK-D assessment carried out in the PEAK-G training intervention as the most current PEAK-D assessment of the current skill repertoire of the participants. Thus, as with PEAK-G training, five programs from the PEAK-D curriculum were selected for each participant in the both groups. For both groups, programs were selected that correspond to the five lowest alphanumeric assessment items that the participant incorrectly responded to during the PEAK-D assessment. Training sessions were also conducted between one to five times a week and consisted of traditional discrete trial trainings and followed the same structure as employed in the PEAK-G training. Mastery criteria for all PEAK-D programs were the same as during PEAK-G training. When a program was mastered, the trainer removed that program from the participant's current programmes section of their folder and moved the programme to where the participant's mastered skills were kept. Mastered programmes were available to be revised throughout the

day for maintenance of skills over time. Session lengths varied according to participant performance (i.e.., participants who frequently responded correctly completed the session more quickly). Each participant was exposed to at least three trainers. Trainers had previous training and practical experience conducting PEAK discrete trial training with participants. All trainers received a behaviour skills training session on how to deliver PEAK programs at the commencement of Study One and were required to revise the task analysis of the training procedure. Participant progress was also monitored on the same frequent ongoing basis using the same procedures as during PEAK-G training. No PEAK-D programs were suspended.

Post-PEAK training Mastery Assessment.

After reaching mastery criterion on all their PEAK programmes all twelve participants were re-assessed using both the PEAK-D and the PEAK-G assessments in order to compare scores and determine if any score changes occurred in PEAK-G assessment scores after the implementation of PEAK-G training, which would indicate that participants had been trained in generalization, and also to determine if PEAK-D assessment scores had increased without any direct implementation of PEAK-D Training. The assessment procedure was identical to the initial PEAK-D and PEAK-G assessments, (please refer to previous section on PEAK assessments for exact procedure).

Natural Environment Assessment.

Following the mastery of all PEAK-D and PEAK-G programmes and their subsequent re-assessment, each participant was assessed again to determine if the skills that were mastered had also generalised to the natural environment using an assessment based on NET. Given that the skills targeted via PEAK were found to not be present in participant's skill repertoire there was no baseline phase conducted for the NET assessment. The assessment followed the basic structure of a PEAK assessment, i.e. prepare the assessment materials, establish learner motivation, present the participant with the stimuli and the SD, and allow

participant 3s to respond correctly across 10 discrete trials. Prior to conducting each assessment, the researcher arranged the natural environment, i.e. classroom or outside setting, of each participant in advance in such a way that the multiple PEAK skills that were taught could be rapidly evoked by the researcher in a loosely-structured manner as is typical in Natural Environment Teaching. Conducting the assessment required the assistance of a second trained observer to record the assessment data as during NET implementers are typically focussed on contriving motivation in the targeted task and rapidly alternating between teaching programs to maintain high levels of participant interest and build behavioural momentum. This makes data recording difficult as having to interrupt this process to record data may interrupt the behavioural momentum and thus impair participant motivation to complete the targeted tasks. The assessment was conducted using untaught novel stimuli that were related to the mastered tasks. For example, to evoke and assess the mastered skill "Receptively ID Shades of Color" (Peak DT – program 3D), the participant was presented with pre-arranged novel stimuli in their play environment such as toy cars and asked to show the assessor "the light blue car" out of an array of five different coloured cars. The participant would then be given 3s to identify the correct item as was required in the PEAK assessments. Similar to the PEAK assessments, participants were not reinforced for their responses to assessment trials however did receive reinforcement in the form of a free-play break after 10 consecutive assessment trials. This allowed the researcher to maintain the participant's motivation by pairing with the participant during this time. As previously stated, data collection involved a trained second observer that recorded results for each assessment trial for all mastered PEAK targets. This involved marking correct responses with a '+' and incorrect responses with a '- '. In order for a skill to be considered as having generalised to the natural environment the participant, as during the PEAK assessments, had to score 9 out of 10 trials (90%) correct if an initial delivery of the discriminative stimulus (SD) was missed, as may occur in any fast-paced assessment. This again reflects the assessment

procedures employed in PEAK assessments and is designed to accurately determine if the skill has successfully generalised to the natural environment as a leaner is unlikely to get 9 out of 10 by accident.

Assessment of Social Validity.

Subsequent to the completion of Study Two, all staff that had participated Study One in assessing and implementing the PEAK protocols was administered an anonymous rating scale designed to measure their scores of the PEAK protocols as socially valid and acceptable packages of language assessment and subsequent teaching. Assessments were put into white, unmarked envelopes and distributed amongst staff. These included brief instructions to remind staff that the assessment was completely confidential, and briefly explained how to complete the assessment, and provided instructions on how to submit completed assessments, please see materials section for further details. This assessment was based on the Intervention Rating Profile-15 (IRP-15) by Martens, Witt, Elliot, & Darveaux (1985) which was modified, with the permission of the original researchers for the purpose of this research (please see materials section for detailed description). For submission of assessments staff were required to place the assessments in the previously supplied unmarked envelopes and seal them. A special post-box style box was placed in the staff room for staff to return their assessments anonymously. This box had a slot cut into the side facing into the room for posting of completed assessment in their envelopes. Once assessments were in the box there was no method of identifying the identity of the staff member that completed an assessment. The box was left in the staff room for a week and checked daily until all assessments had been returned. At this point the assessments were removed and analysed. Research by Von Brock, and Elliot (1987) had previously indicated that a total mean score rating of 58.7% or above would be indicative of an acceptable rating of level of perceived social validity by

participants. After this stage staff were fully debriefed and thanked for their continual participation.

Assessment of Implementation Fidelity and Integrity

Implementation fidelity and integrity was assessed using a modified version the PEAK Implementation Checklist (PEAK-IC; see Appendix 2) across 5 practice PEAK-D and PEAK-G trial blocks. During assessment, each of the 5 practice trial blocks consisted of 10 PEAK trials, and were completed using both PEAK-G programmes and PEAK-D programmes separately. The researcher and trained instructor recorded the number of steps performed correctly and incorrectly by the staff in each trial block on the PEAK-IC for both PEAK modules by ticking the appropriate column, see Appendix 2 for a complete list of PEAK-IC steps. The number of steps performed correctly were divided by the total number of steps and multiplied by 100, in order to determine the percentage of implementation fidelity and integrity by each staff participant for each block. Percent correct was used as the metric because of an inequality in the number of steps required to correctly conduct each trial, as some steps were not possible in all trials. This process was conducted at both the initial staff training phase of the research and at a six month follow-up. Ongoing procedural integrity was measured through IOA sessions with staff.

Inter-observer Agreement

Throughout the implementation of staff training, the PEAK-D assessment, the PEAK-G assessment, PEAK-G training, PEAK-D training, and NET assessments, Inter-observer Agreement (IOA) was conducted for 30% of each respective assessment and intervention via a trained second observer. This second observer independently coded a data sheet for each trial item and this was then compared to the data sheet of the trainer. IOA was calculated by dividing the number of items scored in agreement by the total number of

items observed and then multiplying by 100 to gain a percentage of IOA. Levels of agreement for all PEAK assessments ranged from 91.3% to 100% (M= 97.32%), and were 100% for PEAK scoring across training trials. Levels of agreement for the NET assessments were 100% (M=100%), see table 3.

Table 3

Percentage of Inter-Observer Agreement for Study One.

IOA	Agreement	Disagreement
PEAK-G Assessment Time 1	91.03%	8.70%
PEAK-G Training	100%	0
PEAK-G Assessment Time 2	100%	0
PEAK-D Assessment Time 1	92.90%	7.1%
PEAK-D Training	100%	0
PEAK-D Assessment Time 2	100%	0
Natural Environment Teaching Assessment	100%	0

Results

PEAK-Generalization Training Results

For all participants receiving the PEAK-G intervention, five programmes from the PEAK-G curriculum were selected based on the individual's results from the PGA. These were trained to the mastery criteria of a score of 100% once, or 90% or above two times consecutively, as specified in the PEAK training manuals. All 7 TD participants and 3 of the ASD participants completed all 5 PEAK-G targets to mastery. Two ASD participants, had programmes suspended after it was determined, under guidance from the supervising BCBA supervisor that the pre-requisite skills required for the two participants to make progress in these programmes were not in the participant's repertoires. Three programmes were suspended in total, (see table 1 and 2 for a list of programmes that were taught to criterion).

PEAK Assessment Score Changes

Tables 3 and 4 display the increases in PEAK-Generalization Training Assessment (PGA) and PEAK-Direct Training Assessment (PDA) scores for typically developing (n=7) and pre-schoolers with ASD (n=5), pre- and post- participants receiving PEAK-G training on five learning targets each. All participants made gains in their PGA and PDA scores, without the direct implementation of PEAK-D training. From table 4, the increases in PGA scores for the TD participants ranged from 7 to 43. The increases in PGA scores for the ASD participants ranged from 3 to 23, (see table 3). Tables 3 and 4 also show the range of score increases in the PDA scores for both ASD and TD Participants (TD: 13 to 51; ASD: 15 to 50).

Mean score increase totals, table 5, were also compiled for PGA and PDA time 1 and time 2 for both ASD and TD participants. Results are as follows: TD participants achieved a mean PGA score increase of 25.6 (SD=16.4; 95 Cl, -40.71 to -10.42) and a mean PDA score increase of 24.7 (SD=13.67; 95 Cl, -37.35 to -12.07); ASD participants achieved a mean PGA score increase of 13.4 (SD=8.87; 95 Cl, -24.42 to 2.37), and a mean PDA score increase of 27.4 (SD=14.4; 95 Cl, -45.23 to -9.56). It is important to note the increase in PDA scores occurred without the implementation of training in PEAK-DT Module targets but may be a result of teaching similar learning targets selected from the PEAK-G Module. Following the time 2 PGA and PDA assessments one participant from each sample were excluded from remaining assessments due to reasons external to the research.

Paired-samples t-tests were conducted to compliment traditional visual analysis in evaluating the impact of the PEAK-G training on PGA and PDA scores for both sets of participants. There was a statistically significant increase in PGA and PDA scores for TD [PGA: t (6) = -4.131, p=.006; PDA: t (6) =-4.783, p=.003] and

ASD [PGA: *t* (4) =-3.375, p=.028; PDA: *t* (4) =-4.266, p=.013] participants from pre- to post-intervention, (see table 5).

Table 5

Paired samples t-tests examining pre-post intervention differences for PGA and PDA scores for TD and ASD participants

		Mean Diff	SD	*t*	*p*
TD					
	Pre-PGA vs. Post-PGA	-25.57143	16.37943	-4.131	.006
	Pre-PDA vs. Post-PDA	-24.71429	13.67131	-4.783	.003
ASD					
	Pre-PGA vs. Post-PGA	-13.40000	8.87694	-3.375	.028
	Pre-PDA vs. Post-PDA	-27.40000	14.36315	-4.266	.013

Table 3

PEAK-G and PEAK-D Assessment scores time 1, time 2, PEAK-G Targets, and a PGA and PDA (increase) score for ASD participants.

Participant	Assessment Time 1		PEAK-G Targets	Assessment Time 2		Change	
	PEAK-G	PEAK-D	(Taught to Criterion)	PEAK-G	PEAK-D	PEAK-G	PEAK-D
1	25	51	2B Tact: Non-- Identical Animals 6E Receptively ID Shapes 5G Generalized Patterning 10D Transcription: Connecting the Dots 11B Picture Sequences with Delay	44	101	19	50
2	26	55	10D Transcription: Connecting the Dots 6E Receptively ID Shapes 3D Receptively ID Shades of Color 5G Generalized Patterning 11B Picture Sequences with Delay	43	70	17	15
3	44	85	2B Tact: Non-- Identical Animals 3D Receptively ID Shades of Color 10D Transcription: Connecting the Dots 5G Generalized Patterning 11B Picture Sequences with Delay	67	118	23	33
4	3	11	13H Receptively ID Non--- Identical Vehicles 7H Problem Solving: Packing a Container 3D Receptively ID Shades of Color 4E Matching Numbers and Letters	8	32	5	21
5	11	10	4E Matching Numbers and Letters 7H Problem Solving: Packing a Container 6E Receptively ID Shapes	14	28	3	18

Table 4

PEAK-G and PEAK-D Assessment scores time 1, time 2, PEAK-G Targets, and a PGA and PDA (increase) score for TD participants.

Participant	Assessment Time 1		PEAK G Targets	Assessment Time 2		Change	
	PEAK G	PEAK D	(Taught to Criterion)	PEAK G	PEAK D	PEAK G	PEAK D
1	71	123	8C Metaphorical Tact Extensions 10D Transcription: Connecting the Dots 5G Generalized Patterning 12L Delayed Receptive Picture ID 4E Matching Numbers and Letters	114	158	43	35
2	106	151	8C Metaphorical Tact Extensions 5G Generalized Patterning 4E Matching Numbers and Letters 10D Transcription: Connecting the Dots 11B Picture Sequences with Delay	148	170	42	19
3	80	131	8C Metaphorical Tact Extensions 5G Generalized Patterning 4E Matching Numbers and Letters 10D Transcription: Connecting the Dots 3D Receptively ID Shades of Color	110	145	30	14
4	96	137	8C Metaphorical Tact Extensions 5G Generalized Patterning 4E Matching Numbers and Letters 10D Transcription: Connecting the Dots 3D Receptively ID Shades of Color	134	156	38	19
5	18	59	6E Receptively ID Shapes 5G Generalized Patterning 11B Picture Sequences with Delay 10D Transcription: Connecting the Dots 12L Delayed Receptive Picture ID	25	72	7	13
6	37	39	4E Matching Numbers and Letters 5G Generalized Patterning 10D Transcription: Connecting the Dots 6E Receptively ID Shapes 11B Picture Sequences with Delay	46	90	9	51
7	38	53	6E Receptively ID Shapes 4E Matching Numbers and Letters 10D Transcription: Connecting the Dots 11B Picture Sequences with Delay 5G Generalized Patterning	48	75	10	22

Trials to criterion data for TD and ASD Pre-schoolers.

Mean trials to criterion data were also analysed for matching programmes with TD and ASD pre-schoolers for both PEAK-Generalization Training and PEAK-Direct Training (Table 6). These data justified an independent-samples *t* test which was conducted to compare trials to criterion of all programmes for both PEAK-G and PEAK-D in both TD and ASD participants (Table 7).

In the PEAK-G trials to criterion data there was a significant difference found between results for the ASD participants (M=10.22, SD=7.55) and those for the TD participants (M= 2.29, SD=1.27; t (21.767) = 4.55, p< .000, two tailed. The magnitude of the differences in the means (mean difference =7.93, 95% Cl: 4.55 to 11.30) was large (eta squared = 0.72).

In the PEAK-D trials to criterion data there was also a significant difference found between results for the ASD participants (M=2.15, SD=0.67) and those for the TD participants (M= 1.63, SD=1.07; t (47.89) = 2.10, p =.04, two tailed. The magnitude of the differences in the means (mean difference =5.67, 95% Cl: .02 to 1.01 was large (eta squared = 0.29).

Visual analysis of the mean number of trials to criterion for PEAK-G learning targets in the participants with ASD reveals a much higher mean number of trials were required to meet criterion than those seen in the TD participants scores for PEAK-G and the number of trials needed for both ASD and TD participants throughout PEAK-D training.

Table 6

Mean Trials to Criterion for matching programmes for TD and ASD participants.

Programme	Mean Trials to Criterion ASD	Mean Trials to criterion TD	Mean difference
PEAK-G programmes			
3D Receptively ID Shades of Color	6.3	1	5.3
4E Matching Numbers and Letters	8	1.7	6.3
5G Generalized Patterning	6.7	2.3	4.4
6E Receptively ID Shapes	17.3	3.7	13.6
10D Transcription: Connecting the Dots	18	1.6	16.4
11B Picture Sequences with Delay	8.7	1.8	6.9
PEAK-D Programmes			
10B Receptively Label Coins	1	1.4	.4

Table 7

Independent t-test between trials to criterion for ASD and TD participants.

	ASD		TD		
	M	SD	M	SD	*t*-test
Trials to Criterion PEAK-G Training	10.38	7.71	2.29	1.268	20.670**
Trials to Criterion PEAK-D training	2.15	.067	1.63	1.07	47.89***

***p<.000*
****p=.04*

Natural Environment Teaching (NET) Assessment.

Following the mastery of all PEAK-G and 5 PEAK-D learning targets for all remaining participants (TD *n*=6; ASD *n*=4), a Natural Environment Teaching based assessment was conducted with all participants. This assessment was conducted by the researcher whilst a second trained instructor recorded the data (see Table 1 and 2 in the method section for a complete list of all NET targets for both TD and ASD participants). From table 8, 100% (PEAK-GT: 35/35; PEAK-DT: 30/30) of skills taught to criterion using

both PEAK-GT and PEAK-DT had generalized to the Natural Environment for TD participants. In the ASD participants 94.1% of skills taught to criterion (16/17) using PEAK-G training had generalised to the natural environment, and 90% (18/20) of the skills taught to criterion using PEAK-D training were also present. Note that as two ASD participants had PEAK-G programmes suspended, as per PEAK-G guidelines, due to non-progress in scores the suspended learning targets were not assessed with these participants via the NET assessment as the participants had not successfully learned the skills targeted.

Table 8

Natural Environment Teaching assessment

Population	PEAK Module	Total Percentage of Skills Present	Total Number of skills missing
ASD	Generalization Training Manual	94.5%	1
	Direct Training Manual	90%	2
TD	Generalization Training Manual	100%	0
	Direct Training Manual	100%	0

Staff Use of PEAK

During the rehearsal step of the BST initial implementation, fidelity data was taken to ensure staff were trained sufficiently in delivering PEAK assessments and training accurately, and performance feedback was provided to staff based on this. Implementation fidelity and integrity was assessed using a modified version the PEAK Implementation Checklist (PEAK-IC; see materials) across 5 practice PEAK-D and PEAK-G trial blocks. Items on the original PEAK-IC are taken directly from the PEAK-D training manual's instructions on how to run PEAK-D trials accurately (Dixon, 2014a). However, to also assess the implementation fidelity of PEAK-G the original PEAK-IC items were modified to include the train/test protocol used in PEAK-G, see method section for further details. Each of the 5 practice trial

blocks consisted of 10 PEAK trials, and were completed using both PEAK-G programmes and PEAK-D programmes separately. Staff implementation fidelity during training was found to be very high with a score of 100% (M= 100%) being obtained by all 6 staff members during the final stages training. Implementation fidelity was assessed again using the PEAK-IC at a 6-month follow-up to assess if the previously high levels of implementation fidelity were maintained. Implementation fidelity assessments were conducted for 30% of trials and was found to be maintained at a very high level of 100% (M= 100%) correct for all 6 participants, see table 9.

Table 9

Staff PEAK Implementation Checklist scores

Participant	PEAK- IC Score during BST	PEAK-IC Score at 6-month follow-up
1	100%	100%
2	100%	100%
3	100%	100%
4	100%	100%
5	100%	100%
6	100%	100%

Social Validity: Staff Ratings of PEAK

Six members of staff that directly implemented both PEAK protocols with either one of the participant samples completed an anonymous subjective rating of the social validity and treatment acceptability of the PEAK protocols as assessment and intervention protocols. This assessment, see Appendix 2, was based on the Intervention Rating Profile-15 (IRP-15)

by Martens, Witt, Elliot, & Darveaux (1985) which was modified, with the permission of the original researchers for the purpose of this research. The assessment consists of 15 items which are rated using a six-point Likert-type rating scale with ranges from 1(*strongly disagree*) to 6 (*strongly agree*). Total scores are obtained by summing all items, with higher summed scores indicating greater levels of acceptability, see table 10 for summary of results. Both PEAK-G (71.9%) and PEAK-D (77.6%) received a total mean rating score that was above the 58.7% subsequent researchers (Von Brock & Elliott, 1987) considered necessary for an intervention to be deemed acceptable, with only the PEAK-D receiving an individual rating slightly below this threshold (57.7%).

Table 10

Staff ratings of PEAK-G and PEAK-D training based on Intervention Rating Profile – 15.

		PEAK GT Total Validity Score		PEAK-D Total Validity Score	
Participant	Population	Numerical Score	Percentage Score	Numerical Score	Percentage Score
1	ASD	63/90	70%	52/90	57.7%
2	ASD	68/90	75.5%	57/90	63.3%
3	ASD	55/90	61.1%	55/90	61.1%
4	TD	78/90	86%	76/90	84.4%
5	TD	77/90	85.5%	77/90	85.5%
6	TD	78/90	86.6%	71/90	78.8%
Total score		388	71.8%	419	77.5%
Mean Score		64.7/90	71.8%	69.8/90	77.5%
SD		11.25		9.53	

Discussion

The data from Study One indicates that the PEAK Relational Training system is an effective and socially valid practical language assessment and training tool which is easily implemented by novice staff and teaches targets in a manner that promotes generalisation.

When pre- and post-intervention PEAK Generalization assessment scores were compared after PEAK-G training, it was found that all participants had achieved an increased PEAK-G assessment score following the intervention. Although participant scores were expected to have increased by a number loosely related to the number of PEAK-G targets that participants had trained to mastery, it was found that for 10 out of 12 participants a PEAK-G assessment score increase was achieved that was above the number the number of PEAK-G programs taught to mastery. Similar results were found when the PEAK-D assessment scores from pre- and post- PEAK-G intervention were analysed. It was found that despite having no exposure to the PEAK-D protocol, all 12 participants displayed an increase in PEAK-D assessment scores within a range of 13 to 51. The rise in PEAK-G assessment scores pre- and post- PEAK G training seems to indicate that PEAK-G training is an effective intervention for not only teaching participants the PEAK-G targets but also for promoting the generalisation of other skills.

Similarly, the results indicating an increase in PEAK-D assessment scores post- the PEAK-G intervention also suggest that PEAK-G targets several of the skills present in the PEAK-D assessment and training protocol, thus, to implement the PEAK-G training protocol first, over the easier PEAK-D protocol, may result in an overall positive gain of skills targeted by PEAK-D. Using a paired samples t-test, justified for the small sample sizes by the procedure described in Winter (2013) it was found that for all participants in either sample a statistically significant difference was present between the pre- and post- Intervention PEAKG and PEAK-D assessment scores. Thus, the conclusion that PEAK-G promotes generalisation of skills beyond those targeted by PEAK-G is also supported by these results as all participants made significant gains in their post-intervention PEAK assessment scores for both PEAK modules. These results go towards addressing the limitations discussed in

McKeel et al., (2015) where the authors reported a need for intervention effectiveness to be demonstrated for all the modules of PEAK.

An important limitation of these findings is that in Intervention One the experimental design that was employed lacked the level of experimental control that would allow strong conclusions from the data. Thus, any PEAK-G or PEAK-D score changes which occurred, not including those which occurred due to mastery of PEAK-G targets, cannot be solely attributed to the PEAK-G intervention. Due to the lack of experimental control extraneous variables such as maturation or may have also attributed towards the statistically significant score increases reported in both PEAK-G and PEAK-D scores for participants post- the PEAK-G intervention. The rational for using an AB design in Intervention One was to explore whether implementing the PEAK-G protocol in the same way it would be implemented in a practical setting would result in any findings which would warrant further investigation under stricter experimental contingencies. Therefore, these findings justified a further exploration of this effect by replicating Intervention One using an experimental design that would exert strong experimental control to determine if these results were due to the implementation of PEAK-G training or unknown extraneous variables that were not accounted for in the current research.

NET assessments revealed that skills taught using the PEAK-G and PEAK-D protocols had generalised to their appropriate natural environments for both the TD and ASD participants. TD participants displayed 100% generalisation of mastered PEAK skills to their appropriate natural environments and ASD participants displayed 94.1% of skills taught to mastery through PEAK-G, and 90% of skills taught to mastery through PEAK-D had also generalised. These findings are important because they highlight that the skills targeted by PEAK modules, especially the PEAK-G module, have a very high chance of generalising to the natural environment where they become functional and result in more opportunities for

participants to engage in natural reinforcement that was previously unavailable to them. This intervention address the suggestion for future research highlighted by Rowsey, Belisle, & Dixon, 2014, which called for examination of how various ABA approaches and treatments compare to PEAK through empirical investigation. A limitation of these findings was that there was no baseline assessment conducted for the NET assessment which would have demonstrated a clear effect of NET assessment score changes pre- and post- PEAK-G intervention. However, this omission was justified for several reasons. First, all participants had already undergone assessment via the PEAK-D and PEAK-G assessments, and had demonstrated that the skills which were targeted by the PEAK training and the NET assessment were not present in their skill repertoires. Therefore, to have assessed each target skill a second time via the NET assessment would have been unnecessary, and a waste of the participant's time. Second, as the PEAK assessments functioned as the primary assessment for both PEAK-G and PEAK-D targets in the structured work environment, the NET assessment was a secondary assessment which sought to determine if the skills also presented themselves in the appropriate natural environment. Thus, the NET assessment did not constitute the main assessment of mastered PEAK skills but simply functioned as a preliminary exploration into a possible interaction between AVB based-teaching approaches and Naturalistic teaching approaches.

Another trend that emerged from the data during visual analysis and a subsequent t-test was that there was a significant difference in the number of trials to criterion it required TD and ASD participants to achieve mastery of both matching targets and overall PEAK-G and PEAK-D targets. Although easily dismissed as an effect resulting from the fact that typically developing participants do not experience the same levels of difficulty with generalising newly mastered skills as their ASD counterparts, it is still vital that PEAK develop a strong base of normative data whereby, as stated in Dixon, Belisle, Whiting, and

Rowsey (2014), "specific cognitive and language deficits experienced by individuals with autism can be identified as they deviate from typical cognitive and language development" (p. 1598). Thus, this data offers an exploratory insight into the existence of this effect also within the PEAK-G protocol.

Results show that when provided with Behaviour Skills Training (BST) staff demonstrated through measures of implementation fidelity, both during training and at a 6-month follow-up, that PEAK can be implemented with a very high level of fidelity over a long period of time. This was further supported by IOA data that were collected throughout the course of Study One and ranged between 91.03% and 100%. The implications of the results from Study One are extensive. First, the high levels (100%) of implementation fidelity demonstrated by staff during assessments of implementation fidelity suggest that by using the BST model staff that have very limited experience of ABA and are naïve to PEAK training can be successfully trained to demonstrate very high levels of implementation fidelity. These findings are significant because they demonstrate that the PEAK protocols can be easily implemented in practical settings using BST and also that staff are very likely to display a high level of implementation fidelity over an extended period of time, despite limited past experience of implementing an ABA based assessment and training protocol. These conclusions are further supported by the high levels of inter-observer agreement (between 91.03% and 100%) that were recorded throughout Study One which demonstrated that data recording was also accurate overall. These findings extend findings reported in Belisle, Rowsey, & Dixon, 2016, where researchers found Behavioural Skills Training to be an effective means of training implementers that are naïve to PEAK in how to deliver the PEAK-D curriculum and assessment by replicating this effect with the PEAK-G curriculum and assessment. A limitation of these findings was that implementation fidelity measures were only conducted on two occasions with staff. Although, these measures indicated that

procedural fidelity and integrity had been maintained at very high levels at the 6-month follow-up, there is no data to indicate that procedural fidelity was maintained in the period between. Inter-observer agreement data was collected throughout the course of the research; however, this only increases the confidence that observers are accurately measuring dependent variables of interest, and does not measure the level of implementation fidelity. For further in-depth discussion of Study 1 findings, limitations and recommendations for future research please see Chapter Four.

Overall Study One has made some significant findings to supplement the existing research on the use of the PEAK Relational Training System in practical settings. These findings have provided some exploratory results on aspects such as differences in performance between typically developing participants, and participants with ASD, demonstrated that the implementation of PEAK-G resulted in positive assessment score changes for all participants on both the PEAK-G and PEAK-D assessments, that new skills mastered through PEAK-G and PEAK-D training generalise to their appropriate natural environment, that naïve staff can be trained to implement the PEAK protocols with high levels of implementation fidelity, and that staff would rate both PEAK-G and PEAK-D protocols as socially valid and acceptable language training and assessment tools within the populations that they were applied. These findings have also provided the rationale for further exploration in Study 2 of the effects that were observed between the implementation of PEAK-G training and the subsequent mastery of PEAK-D skills on the PEAK-D assessment without exposure to the PEAK-D training protocol. This further exploration will use an experimental design that will clearly demonstrate whether the mastery of PEAK-G skills has a direct effect on PEAK-D skills being mastered indirectly and may have important implications for the implementation of PEAK in practical settings moving forward.

Chapter 3

Study Two: Exploring the Generalization Training Protocol's Effect on related PEAK Direct Assessment Targets.

Study One found that when exposed to the Generalization training protocol (PEAK-G) of the Promoting the Emergence of Advanced Knowledge (PEAK) Relational Training System, five participants with diagnosed autism and seven typically developing participants made significant positive gains in their assessment scores on the PEAK Direct (PEAK-D) training assessment. Overall, only two participants did not obtain a follow-up PEAK-G assessment and PGA score that was higher (5+) than the number of skills mastered. However, the overall levels of increase were still statistically significant for both participants with clinical diagnoses of Autism Spectrum Disorder (ASD), and typically developing (TD) participants and justified a further exploration of this effect. Therefore, the current study aimed to partially replicate this effect through a multiple baseline design by targeting specific PEAK-G targets that have interrelated PEAK-D targets to explore whether a functional relationship existed under stricter experimental conditions than those employed in Study One. Participants for this study were four children with diagnosed autism that were naïve to PEAK training and the assessment protocols. Participants were initially assessed via the PEAK-G and PEAK-D assessments, specifically targeting only PEAK-G targets that had related PEAK-D targets. PEAK-G training was then implemented for five PEAK-G targets identified as being absent from participant's repertoires until all participants reached the mastery criteria. PEAK-D assessments were subsequently conducted again to explore whether exposure to PEAK-G training had resulted in the related PEAK-D targets being mastered, without exposure to the PEAK-D training protocol. Follow-up probes were also conducted towards the end of each participant's multiple baseline intervention on the mastered PEAK-G skills to explore if mastered PEAK-G skills would be maintained without any further explicit instruction via PEAK-G training targeting those skills. This research represents a

novel exploration into these effects and does not follow from specific research recommendations made by previous PEAK researchers.

Method

Participants

Four children, one girl and three boys aged four, were recruited from an ABA class within a Montessori pre-school. All four children had previous diagnoses of Autism Spectrum Disorder by clinical psychologists independent of the current research in accordance with the diagnostic criteria set out in the Diagnostic and Statistical Manual of Mental disorders (5th ed.; *DSM–5*; American Psychiatric Association, 2013). All four children completed the PEAK-D assessment, PEAK-G assessment, PEAK-G training, and follow-up PEAK-D and PEAK-G assessments.

Ethics

It was not anticipated that participants would suffer any physical discomfort as a result of participation and ethical provisions were identical to those made in Study One. Please see Study One for full details of all ethical provisions.

Setting

All training and assessments took place in the same setting as Study One, please see Study One for more detailed description.

Materials

PEAK: Direct Training assessment and curriculum

The PEAK Direct Training assessment and subsequent training programs as found in the PEAK Relational Training System: Direct Training Module (Dixon, 2014a) were implemented with all participants. Please refer to Chapter One for a detailed description of the PEAK Direct Training assessment and curriculum. During the assessment phase all six of Skinner's verbal operants are assessed (i.e. mands, tacts, intraverbals, echoics, textual and transcription responses) , as well as prerequisite learning skills (i.e., eye contact and keeping hands still), vocal skills (i.e., imitating sounds and imitating words), writing skills (i.e., copying text and transcribing speech), conversational skills (i.e., telling jokes and answering questions), basic and advanced math skills (i.e., addition and subtraction), and advanced conversational skills (i.e., identifying metaphors and communicating uncertainty). Hence, as the curriculum's training programs and the assessment are linked, these skills are then targeted throughout the implementation of the training programs.

PEAK: Generalization training assessment and curriculum

The PEAK Generalization Training assessment and subsequent training programs as found in the PEAK Relational Training System: Generalization Training Module (Dixon, 2014b) were also implemented with all participants. Please also refer to Chapter One for a detailed description of the PEAK Generalization Training assessment and curriculum.

Experimental design and measurements

Study Two implemented a multiple-probe across behaviours and introduction of the next PEAK-G target behaviour was contingent upon the previous target being trained to mastery. The Independent Variable was the application of the PEAK-G training intervention for 5 PEAK-G skills determined absent from participant's repertoires through the PEAK-G assessment. Operants targeted during the PEAK-G training intervention included tacts, intraverbals, textual, and transcription responses. The primary DVs were participant's PEAK-G and PEAK-D baseline assessment scores and post-intervention PEAK-D assessment scores for the PEAK-D skills that were related to the PEAK-G skills targeted, whereas a secondary dependent variable was the maintenance of mastered PEAK-G skills over time. Please see table 1 for a list of PEAK targets in Study Two.

Table 1.

Participant's PEAK-G and PEAK-D targets.

Participant Name	PEAK G Targets	PEAK D Targets
James	9G Receptively ID Rooms by Function	11N Tact Rooms
	5G Generalized Patterning	14O Delayed Recall and Audience
	6E Receptively ID Shapes	9K Tact Shapes
	2B Tact: Non- Identical Animals	8M Tact Animals
	10C Tact: Pictures with Delay	11Q Tact Following Delay
Séan	3D Receptively ID Shades of Color	9L Tact Colors
	5G Generalized Patterning	14O Delayed Recall and Audience
	2B Tact: Non-- Identical Animals	8M Tact Animals
	6E Receptively ID Shapes	9K Tact Shapes
	14A Tact: Non-- Identical Clothing	8N Tact Clothing
Sarah	9G Receptively ID Rooms by Function	11N Tact Rooms
	5G Generalized Patterning	14O Delayed Recall and Audience
	6E Receptively ID Shapes	9K Tact Shapes
	2B Tact: Non- Identical Animals	8M Tact Animals
	10C Tact: Pictures with Delay	11Q Tact Following Delay
Luke	4E Matching Numbers and Letters	7C Match Numbers and Letters
	14K Tact: Non-- Identical Vehicles	12N Tact Vehicles
	3D Receptively ID Shades of Color	9L Tact Colors
	2B Tact: Non-- Identical Animals	8M Tact Animals
	6E Receptively ID Shapes	9K Tact Shapes

Procedure

PEAK assessments

The study intervention commenced with the PEAK-D assessment and followed with the PEAK-G assessment, please see Study One for a detailed description of the PEAK assessment protocols. For the purposes of this study only the items on the PEAK-G assessment that had related targets on the PEAK-D assessment were assessed. This was so that it could be explored whether participant mastery of PEAK-G skills resulted in the mastery of their related PEAK-D skills on the PEAK-D assessment without any exposure to the PEAK-D training protocol.

PEAK Generalization Training

All participants had five PEAK-G and five related PEAK-D targets chosen from the results of their PEAK-G and PEAK-D assessments. Once these targets were selected the PEAK-G training commenced immediately on a chosen programme. As the assessments demonstrated that all the programmes targeted were not in participant's skill repertoires, an extended baseline was not conducted on the first skill targeted for each participant as doing so would have been unethical as it would waste the participant and implementer's time. Training sessions were conducted one to five times a week and consisted of discrete trial trainings using the train/test strategy in which the learner is taught to generalize skills using a specified set of stimuli but tested for mastery using a different set of stimuli (Dixon, 2014b), please see Study One for detailed description of the PEAK-G training protocol. No programs were suspended during this study.

Inter-observer Agreement

Throughout the implementation of the PEAK-D assessments, the PEAK-G assessments, PEAK-G training, and staff implementation fidelity assessments, Inter-observer Agreement (IOA) was conducted for 30% of each respective assessment and intervention via a trained second observer and calculated by dividing the number of items scored in agreement by the total number of items observed and then multiplying by 100 to gain a percentage of IOA. Levels of agreement for all PEAK assessments and PEAK scoring across training trials were 100% (M=100%). Levels of agreement for the fidelity assessments were also 100% (M=100%).

Results

A multiple-probe multiple-baseline design across participants, and across behaviours (PEAK-D learning targets) was implemented with four pre-schooler participants with ASD. All participants received PEAK G and PEAK D assessments, and five PEAK G programs and five related PEAK D programs were identified as absent from their repertoires, see table 14 for an example of related PEAK G and PEAK D programs targeted. Figures 1 to 4 show the multiple probe baseline results for each participant which were at zero levels. All four participants successfully learned the skills selected from their PEAK-G assessments upon exposure to the PEAK-G training protocol as outlined in the PEAK Generalization Training Manual.

Upon reaching the learning criterion for each individual PEAK G programme, the participants underwent the PEAK-D assessment procedure to probe for correct responding for the 5 PEAK-D targets selected by the researcher based on the similarity of targets learned using PEAK-G training methods. Data for each participant are represented in figures 1 to 4. These results demonstrate that only after participants completed PEAK-G training and subsequently reached the learning criterion did they display mastery of the related PEAK-D skills. This was

demonstrated by all four participants across all the related PEAK-D targets. These results further support and refine the findings reported in Study One where it was demonstrated that application of the PEAK-G training protocol resulted in participant's PEAK-D assessment scores increasing by a statistically significant number.

Once all PEAK-G targets were trained to the required learning criteria and the related PEAK-D targets were re-assessed, the PEAK-G targets were also re-assessed via the PEAK-G assessment to explore whether mastery of PEAK-G targets was maintained. The results indicate that participants maintained all PEAK-G skills that were trained to mastery regardless of the order in which they were mastered during the implementation of the multiple baseline design.

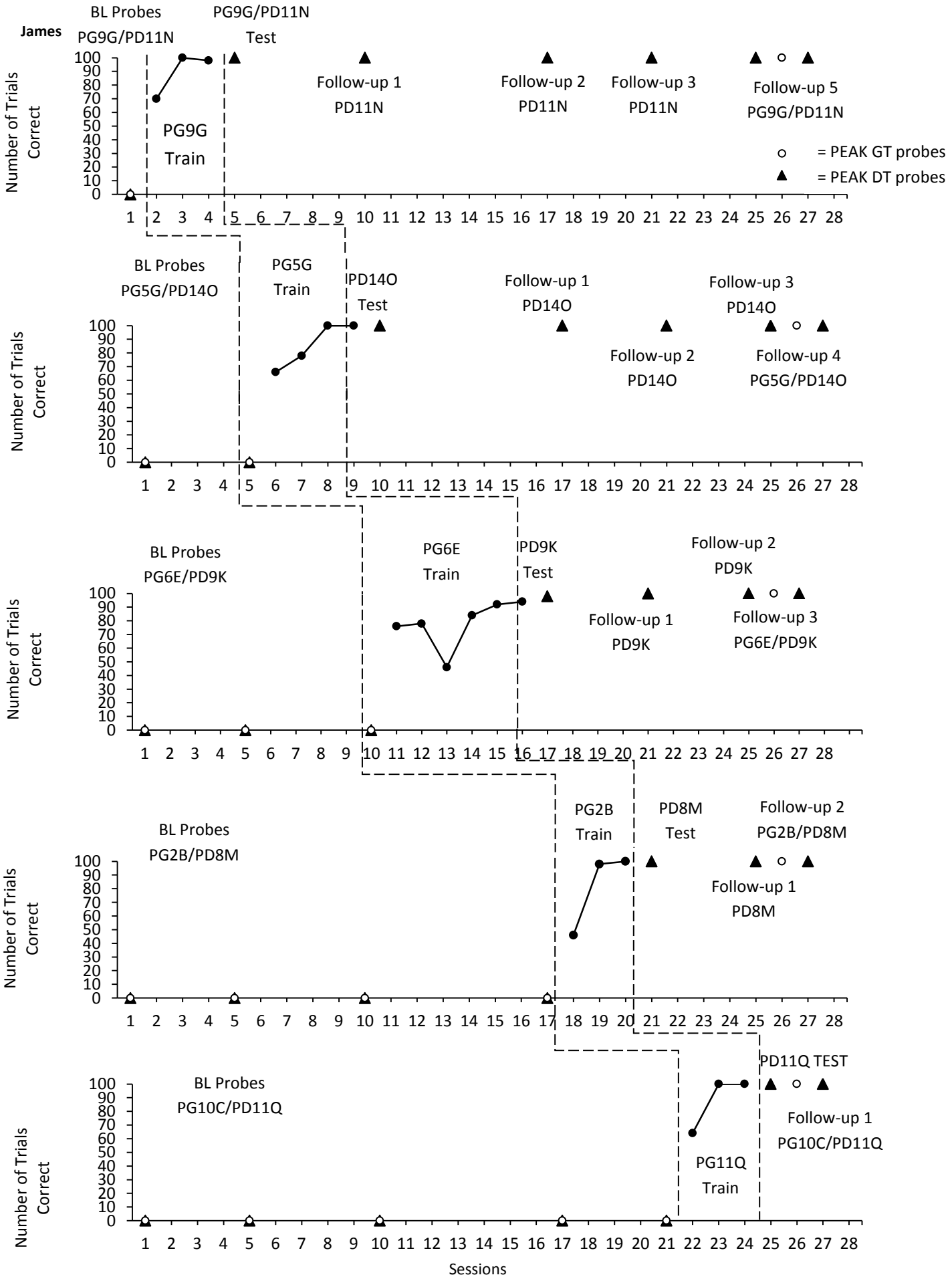

Fig. 1 Baseline probes for PEAK-G and PEAK-D targets; training trials for PEAK-G targets 9G, 5G, 6E, 2B, and 10C; subsequent test for collateral acquisition of PEAK-D targets. BL: Baseline G9G: PEAK Generalisation Module, Programme #9G; D11N: PEAK Direct Module, Programme #11N

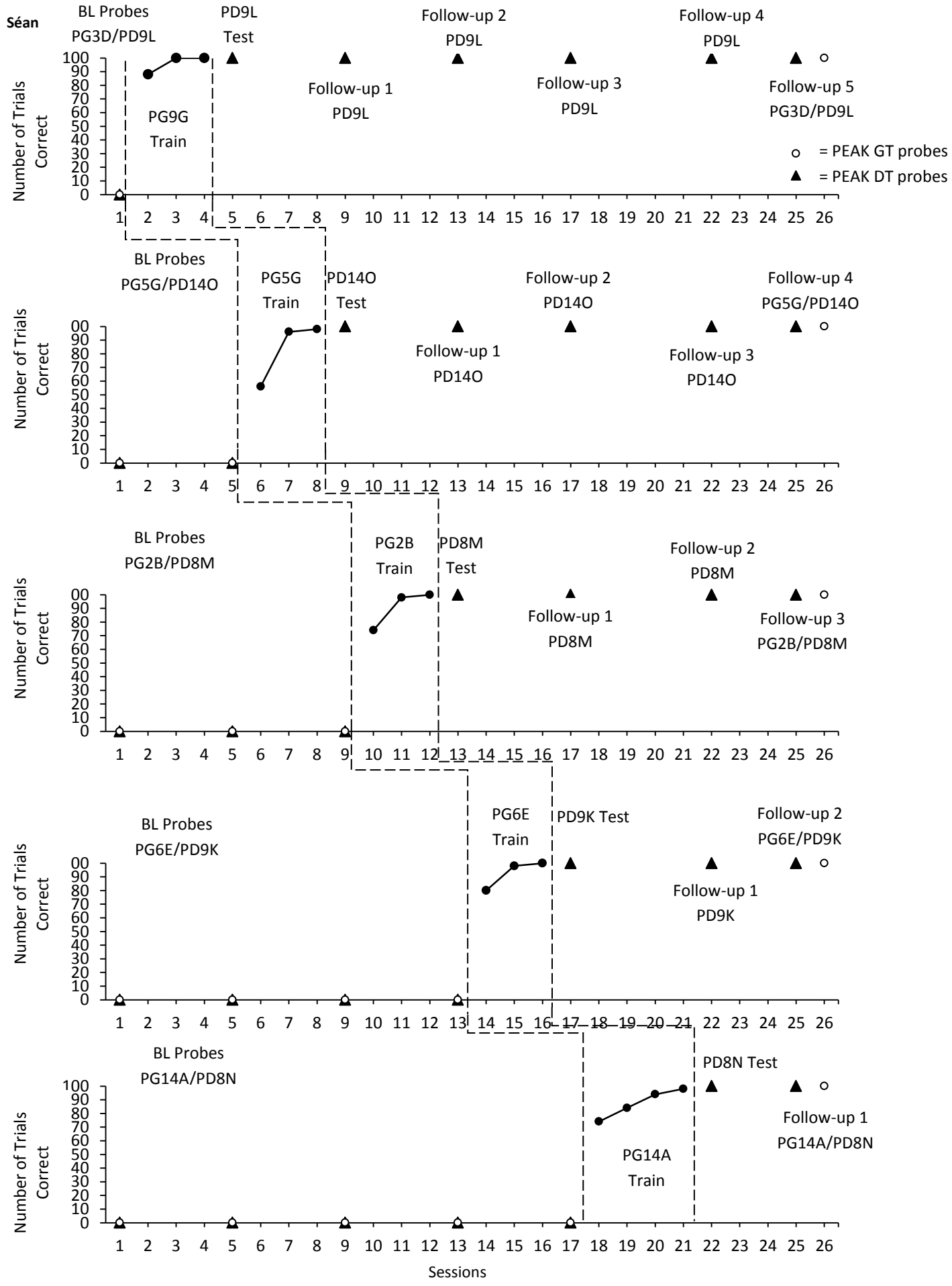

Fig. 2 Baseline probes for PEAK-G and PEAK-D targets; training trials for PEAK-G targets 9G, 5G, 2B, 6E, and 14A; subsequent test for collateral acquisition of PEAK-D targets. BL: Baseline G9G: PEAK Generalisation Module, Programme #9G; D11N: PEAK Direct Module, Programme #11N

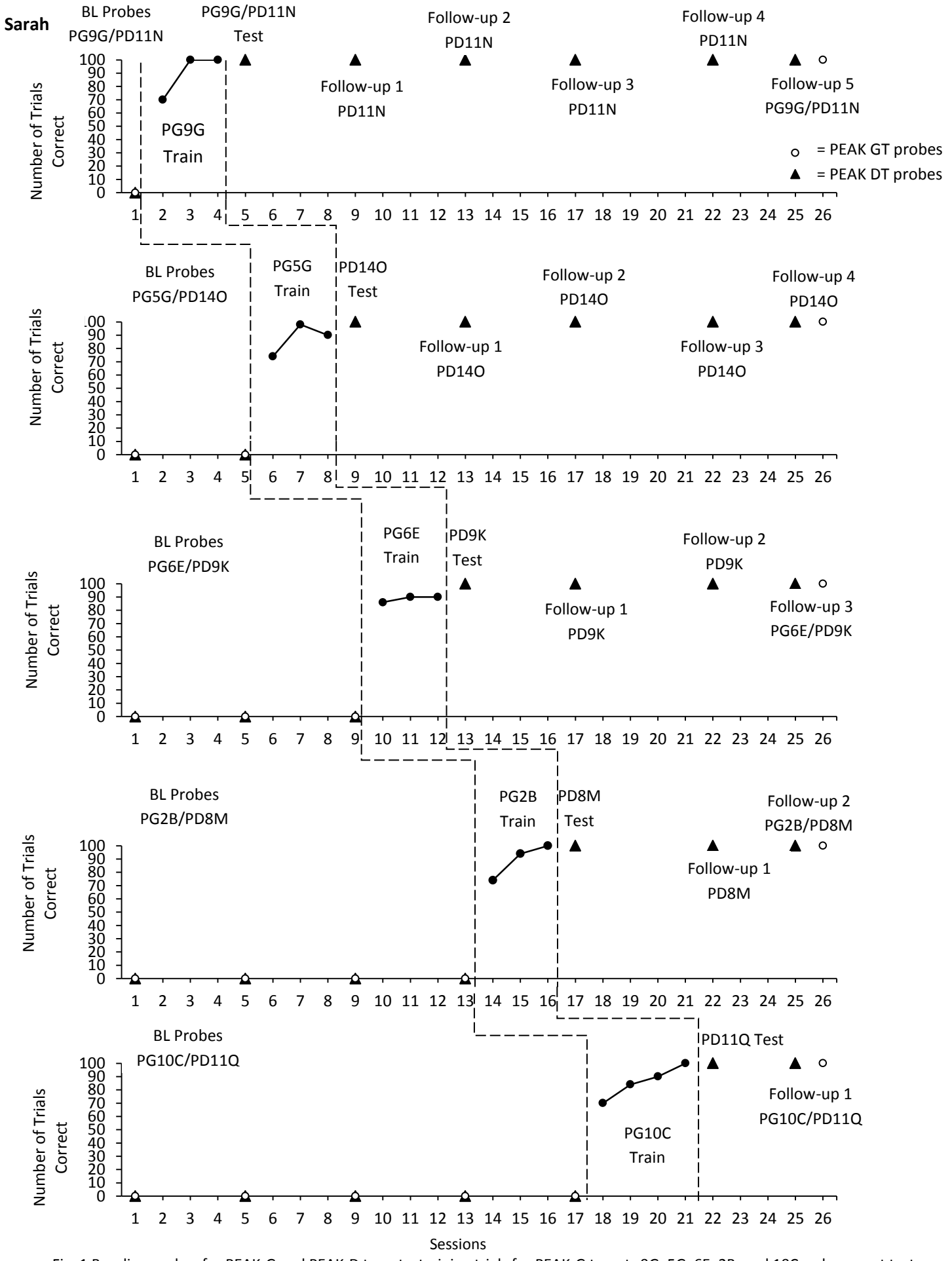

Fig. 1 Baseline probes for PEAK-G and PEAK-D targets; training trials for PEAK-G targets 9G, 5G, 6E, 2B, and 10C; subsequent test for collateral acquisition of PEAK-D targets. BL: Baseline G9G: PEAK Generalisation Module, Programme #9G; D11N: PEAK Direct Module, Programme #11N

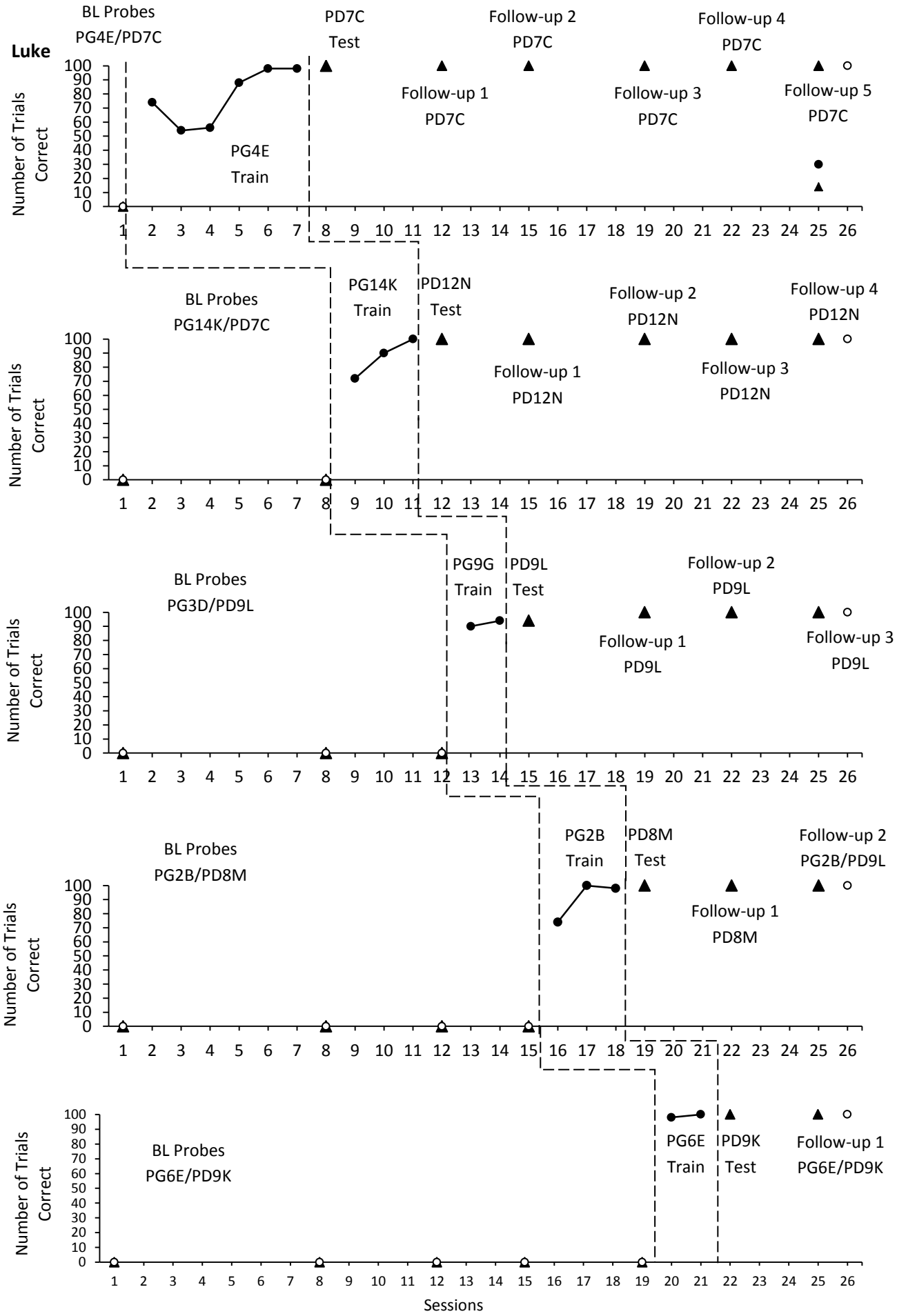

Fig. 1 Baseline probes for PEAK-G and PEAK-D targets; training trials for PEAK-G targets 4E, 14K, 3D, 2B, and 6E; subsequent test for collateral acquisition of PEAK-D targets. BL: Baseline G9G: PEAK Generalisation Module, Programme #9G; D11N: PEAK Direct Module, Programme #11N

Discussion

The findings from Study Two demonstrated that when PEAK-G assessment targets were mastered via training, their related PEAK-D targets were concurrently mastered, which was verified by post-intervention PEAK-D assessment probes, without exposure to the PEAK-D training protocol. These findings partially replicated those reported in Study One, where it was demonstrated that implementation of PEAK-G training resulted in a statistically significantly increased PEAK-G and PEAK-D assessment scores without exposure to the PEAK-D protocol, i.e. that PEAK-G and PEAK-D targets were being mastered simultaneously without participant exposure to the PEAK-D training protocol.

The results of Study Two also refined these findings by using a multiple probe design which offered greater experimental control and graphically demonstrated that a functional relationship exists between related PEAK-G and PEAK-D targets. These findings were complimented by the results that the PEAK-G targets that were trained throughout the implementation of the MBD procedure were also maintained at full mastery regardless of what order they were trained to mastery, i.e. the PEAK-G target which was targeted as the first learning target in the MBD was still maintained as mastered despite the difference in time between the session it was mastered in and the session in which it was re-assessed for maintenance. Re-assessment of staff implementation fidelity also demonstrated that staff had not only maintained the levels of implementation measured during training, but had improved these to 100% levels across the board.

One limitation of this study was that it only assessed PEAK-D skills that were determined to be related to the PEAK-G skills that were found absent from participant's skill repertoires. This was justified as Study Two aimed to explore whether a functional relationship existed between interrelated PEAK-G and PEAK-D targets under stricter experimental conditions than

those employed in Study One. Another limitation was the small sample size which precluded valid within-participant comparisons and limited the generalisability of findings to a clinical level. For further in-depth discussion of Study 2 findings, limitations and recommendations for future research please see Chapter Four.

Overall this study demonstrated that under high experimental control the functional relationship reported in Study One was replicated and thus less likely to have been an effect of extraneous variables such as maturation. It also demonstrated that staff implementation levels were maintained despite a 6-month break between the two assessments of implementation fidelity, and that mastered PEAK-G skills were maintained at mastery regardless of what order they were mastered in during the MBD. These results are important additions to the existing PEAK literature as they contribute to the current base of knowledge surrounding the practical implications of implementing the PEAK protocols in a practical setting.

Chapter Four

General Discussion

This research set out to address several outstanding questions from the existing PEAK research regarding the practical implementation of PEAK. As previously mentioned, the results from this research indicate that the PEAK Relational Training System is an effective, socially valid language assessment and training tool which is easily trained and teaches targets in a manner which promotes generalisation. This research was conducted as a preliminary exploration of many outstanding questions which surround the implementation of PEAK in practical environments.

Study One consisted of three interventions that sought to explore several aspects of implementing the PEAK Generalization and Direct Training modules in a practical pre-school environment. The results of Intervention One demonstrated that PEAK-G training was both an effective method for training PEAK-G targets and that it resulted in statistically significant PEAK-G and PEAK-D assessment score gains for all participants. It was also found that there was a statistically significant difference between the amount of trials required to reach criterion by the ASD participants and those for the TD participants. Intervention Two found that over 90% of PEAK-G and PEAK-D mastered skills had generalised to their appropriate natural environment. Finally, Intervention Three, found that staff that were naïve to implementing PEAK or ABA-based interventions could implement PEAK with a very high level of implementation fidelity after receiving Beahvioural Skills Training. This result was further supported by the high levels of inter-observer agreement during PEAK assessments, PEAK training and NET assessments.

Study Two consisted of Two interventions. Intervention One sought to further explore the relationship found in Intervention One of Study One between the implementation of PEAK-G and the subsequent mastery of PEAK-D targets without exposure to PEAK-D training.

Maintenance probes were also conducted for each of the PEAK-G targets after participants had mastered all 5. Intervention Two sought to re-assess the levels of staff implementation fidelity to determine if staff had maintained the high levels of implementation fidelity that were recorded during their Behavioural Skills training. The results of Intervention One indicated that, despite the use of an experimental design that provided strict experimental control between the dependent and independent variables, the mastery of targets via PEAK-G training still resulted in the subsequent mastery of related targets on the PEAK-D assessment without any exposure to PEAK-D training. It was also found that mastered PEAK-G skills were maintained regardless of which order they were mastered throughout the MBD. The results of Intervention Two were that staff had maintained implementation fidelity scores of 100% across the board dspite a 6-month gap in assessment prior to the follow-up assessment.

A major aim of both Study One and Study Two was to explore the relationship between the PEAK-G module and the PEAK-D module when both are applied in a practical setting, with PEAK-G being implemented first. What was found in Study One was that when PEAK-G training was implemented for 5 PEAK-G targets to mastery that both ASD and TD participants made statistically significant gains on both their PEAK-G and PEAK-D post-intervention assessments. The implications of this could be very important for practitioners conducting PEAK training in practical settings. First, these findings suggest that PEAK-G training is an effective intervention for the PEAK-G targets contained in the PEAK-G assessment. To date this kind of intervention-outcome evaluation of a PEAK module has only been conducted with PEAK-D, (McKeel, Dixon, Daar, Rowsey, and Szekely, 2015), and so these results represent the first intervention-outcome evaluation for the PEAK-G training module. Another important aspect of the results from Intervention One was that Although participant scores were expected to have

increased by a number loosely related to the number of PEAK-G targets that participants had trained to mastery, it was found that for 10 out of 12 participants a PEAK-G assessment score increase was achieved that was above the number the number of PEAK-G programs taught to mastery. Similar results were found when the pre- and post- PEAK-G intervention PEAK Direct Assessment scores were analysed. It was found that despite having no exposure to the PEAK-D protocol, all 12 participants displayed an increase in PEAK-D assessment scores within a range of 13 to 51. This rise in PEAK-G assessment scores pre- and post- PEAK G training would indicate that PEAK-G training is an effective intervention for not only teaching participants the PEAK-G targets but also for promoting the generalisation of other skills. In Study Two, this effect was replicated on a more controlled scale with PEAK-G targets which have interrelated PEAK-D targets in order to explore if training specific PEAK-G targets would result in their related PEAK-D targets also being mastered. A multiple baseline design was also used as it provides stricter experimental control between the variables. The results demonstrated the same effect as reported in Study One, with all participants mastering all PEAK-G and the inter-related PEAK-D targets without any exposure to the PEAK-D curriculum. This result gave strong support for the previously reported effect and clearly demonstrated that the results from Study One were not due to Type I or Type II error. From a visual analysis of the graphs from Study Two it's evident that participant performance on PEAK-G targets only improved once PEAK-G training was implemented and that the related PEAK-D targets were mastered once post-intervention assessments were conducted. These two results both give strong support to the conclusion that the implementation of PEAK-G training not only trains the targets identified by the PEAK-G assessment but also trains participants to generalise previously learned skills in a way that improves their PEAK-D assessment scores post PEAK-G training.

Another interesting effect that was reported in Study One was the results of the Natural Environment Teaching assessments. The results from the NET assessments have several implications. First, as mentioned in Chapter 1, NET is based on the field of Naturalistic Teaching Approaches and the PEAK protocols are based in the field of Applied Verbal Behaviour. Therefore, in order for the results to show that previously mastered PEAK skills can be elicited and assessed via an NET-based assessment further strengthens the findings found in LeBlanc, et al., (2006) where researchers found an existing compatibility between the AVB and NTA approaches. Second, the fact that such a high percentage of mastered PEAK skills from both modules were found to have generalised to their appropriate natural environments is also significant. In TD participants 100% of mastered skills from both modules were exhibited in their appropriate natural environment, whereas in the ASD participants 94.05% of mastered PEAK-G skills and 90% of PEAK-D scores were exhibited by participants in their appropriate environments.

Throughout the NET assessment, it was to be expected that TD participants would exhibit a high level of generalisation of the skills taught to mastery through PEAK to their appropriate natural environments, as TD individuals have fully functional relation frames and do not experience the same difficulties with generalisation of mastered skills as ASD individuals do. However, the results reporting a similar effect in the ASD participant's ability to generalise their mastered PEAK skills to the natural environment are more significant as the development of skill generalization is often impaired in ASD populations. These findings were further supported by the maintenance probes in Study Two which indicated that mastered PEAK-G skills were maintained at mastery levels regardless of the period between the mastery probes and the final training session of the skills. What this would suggest is that once skills are mastered via PEAK-

G training they generalise to, and are maintained in, participant's natural environments. This also gives support to the conclusion that PEAK-G trains participants to generalise their skills beyond the structured setting within which they are implemented.

Both Study One and Study Two also had a secondary aim of exploring various elements of staff implementation of PEAK-G and PEAK-D. The importance of an exploration such as this has been highlighted in both PEAK publications and other publications concerned with staff implementation of ABA interventions, (Reed, Luiselli, 2016; Foster, & Mash, 2000; Belisle, Rowsey, & Dixon, 2016; McKeel, Rowsey, Dixon, & Daar, 2015; Dixon, Stanley, Belisle, & Rowsey, 2016; Dixon, et al., 2014; Dixon, Whiting, Rowsey, & Belisly, 2014). Using just the instructions published in the PEAK-G and PEAK-D manuals also provided a further exploration of how effective the manuals are at training those with very minimal knowledge of the tactics used in ABA such as the parents or care staff for which PEAK was also developed to be implemented by. The results from Study One demonstrated that by using BST staff, that were previously naïve to the PEAK protocols and the tactics employed in ABA, were trained to demonstrate 100% implementation fidelity and between 91.03%-100% inter-observer agreement for all PEAK assessments, PEAK training and NET assessments over an extended period of time (6 months at least). The implications of this is that by following the BST training model and using the implementation instructions published within the PEAK-G and PEAK-D manuals, PEAK protocols could be delivered with very high levels of implementation fidelity and IOA for an extended period of time by staff in practical settings. Further implications of these results are that facilities with time and resource constraints then don't have to invest extended periods of time or resources training their staff in implementing PEAK for a high-level implementation fidelity to be achieved or maintained.

Another finding from Study One that may have large implications for the PEAK protocols was the results from the Social Validity assessment completed by staff. To date, there has been no exploration of the level of social validity of the PEAK protocols. Although PEAK has been demonstrated as an effective tool for language assessment and as a curriculum tool, Schwartz, and Baer (1991) reported that besides treatment efficacy and validity, measures of social validity are also important as social validity assessments supplement efficacy outcomes. Thus, the results from the Social Validity and Treatment Acceptability assessment represent the first instance of the PEAK protocols being rated as socially valid and acceptable methods of language assessment and training by implementers. Staff participants rated both modules of PEAK a separately for their perceived level of social validity with both TD and ASD participants. What was reported is that the PEAK-G training and assessment protocol received a mean rating of 71.8% and the PEAK-D training and assessment protocol received a mean rating of 77.5%. Both these scores fell above the score of 57.7% as found by Von Brock, and Elliott (1987) to be the score above which interventions can be considered socially valid and acceptable. Therefore, as both the PEAK modules fall above this threshold they can be considered as being rated as socially valid by the staff as training protocols and assessments.

This research also had several limitations and these will be discussed alongside recommendations for future research. The first limitation was the use of a small *n* design throughout precluded valid within-participant comparisons and limited the generalisability of findings to a clinical level. This limitation can be managed in future research through the recruitment of larger numbers of TD, ASD, and staff participants. The addition of more participants would greatly increase the generalisability of findings and would allow further exploration of the effects reported in this research. The second limitation was that in Study One

an AB design was implemented for Intervention One. Using this type of design was justified as an important element of exploring the implementation of PEAK in a practical setting involves implementing it as it typically would be implemented in this type of setting; however, using an AB design prevents any strong conclusions be made from the data as a lack of experimental control means that results are open to confounding and extrinsic variables such as maturation. An example of how this limitation may have interfered with the results of intervention one is that all TD and ASD participants in this research were undergoing teaching via the Montessori method as their main form of teaching throughout the course of this research. This style of teaching could possibly have had an influence on the post-PEAK-G intervention PEAK-D and PEAK-G assessment scores in Study One as the Montessori method targets many of the relational skills that PEAK-G and PEAK-D target. However, as the Montessori method does not involve recording any individual or group data during teaching sessions, it cannot be concluded that this teaching method had any effect on the results of this study.

Future research can address the above possible confound by either taking data on Montessori teaching of targets contained within PEAK, or by implementing an experimental design that a higher level of experimental control, such as a multiple baseline across behaviours design, to demonstrate that a functional relationship does in fact exist between the training of PEAK-G targets and increased scores on the PEAK-D assessment. The third, due to the timeframe within which this research was conducted being limited, suspended PEAK programs in Study One had to be abandoned as there was not sufficient time to train up the necessary pre-requisite skills in order to reinstate the programmes. Thus, future research should address this by conducting an extended study of PEAK-G outcomes where enough time is available that if PEAK programs have to be suspended that there is an opportunity to train up the necessary pre-

requisite skills and re-instate the suspended programs. Fourth, there were only two assessments of implementation fidelity conducted throughout the course of this research in order to explore whether implementation fidelity levels were maintained after an extended period of time, however, future research may wish to conduct more frequent assessments of implementation fidelity to protect against the effects of implementation fidelity possibly decreasing over time. With more frequent assessments and decrease in fidelity levels will be detected at a faster rate and will allow researchers to intervene sooner if necessary. As has already been conducted with PEAK-D (Dixon, Belisle, Whiting, & Rowsey, 2014), all the remaining PEAK modules would require future research to develop PEAK training outcome data with a typically developing population. This would aid in furthering the exploration of how individuals with autism deviate from their typically developing peers in terms of the cognitive and language deficits that these individuals experience through the results of the more complex skills which PEAK has available for researchers to implement. Staff ratings of PEAK's social validity also need be established in settings other than early intervention and young children so that those situations where PEAK training is not socially valid can be established and direct future research and implementation of PEAK.

In conclusion, The PEAK Relational Training System and the PEAK Generalization and Direct training manuals have emerged as conceptually sound, psychometrically robust, and an innovative advancement of conventional ABA tactics for teaching children and youth who have autism and other developmental disabilities (Reed, Luiselli, 2016). However, previous PEAK research had identified many suggestions for future research that would expand the current base of knowledge surrounding the implementation of PEAK in practical settings. Thus, the current research aimed to address several of the key limitations that had been identified and

to further explore the processes involved in implementing PEAK in a practical setting. Overall through this research, PEAK has been demonstrated as an assessment and training protocol that can be trained to be implemented with a high level of fidelity over a long period of time, that is rated as socially valid by implementers, and that is an effective assessment and training tool in a practical setting. Skills taught to mastery through the PEAK protocols have also been demonstrated to be maintained by participants over time, generalise to their appropriate natural environment in over 90% of cases, and yields results that can be replicated with ease even under experimental designs that employ a strict level of experimental control. However, once all four PEAK manuals are published future research, replication, and expansion upon the results found throughout this research for all the remaining PEAK manuals will also be necessary in order to build upon the existing evidence base and to aid future implementers in making effective use of these protocols.

References

Anderson, S. R., Avery, D. L., DiPietro, E. K., Edwards, G. L., & Christian, W. P. (1987). Intensive Home-Based Early Intervention with Autistic Children. *Education and Treatment of Children, 10*(4), 352-366.

Baer, D. M. (1993). A Brief, Selective History of the Department of Human Development and Family Life At The University Of Kansas: The early years. *Journal of Applied Behavior Analysis, 26*(4), 569.

Baer, D. M., Wolf, M. M. & Risley, T. R. (1968). Some Current Dimensions of Applied Behavior Analysis. *Journal of Applied Behavior Analysis, 1*, 91-97.

Baer, Donald M. (1993). A brief, selective history of the Department of Human Development and Family Life at the University of Kansas: The early years. *Journal of Applied Behavior Analysis* 26 (4): 569–570.

Bailey, J. S., & Wallander, R. J. (1999). Verbal Behavior. In B. A. Thyer (Ed.), *The philosophical legacy of behaviorism* (pp.117– 152). Great Britain: Kluwer Academic Publishers.

Barnes, D. (1994). Stimulus equivalence and relational frame theory. *The Psychological Record, 50*, 61–81.

Barnes-Holmes, D., Barnes-Holmes, Y., & Cullinan, V. (2000). Relational Frame Theory and Skinner's Verbal Behavior: A Possible Synthesis. *Behavior Analyst, 23*(1), 69-84.

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.

Belisle, J., Rowsey, K. E., & Dixon, M. R. (2016). The use of in situ behavioral skills training to improve staff implementation of the PEAK relational training system. *Journal of Organizational Behavior Management, 36*(1), 71-79.

Blackledge, J. T. (2003). An introduction to relational frame theory: Basics and applications. *The Behavior Analyst Today*, *3*(4), 421-433. doi:10.1037/h0099997

Braam, S. J., & Poling, A. (1983). Development of intraverbal behavior in mentally retarded individuals through transfer of stimulus control procedures: Classification of verbal responses. *Applied Research in Mental Retardation, 4*, 279–302.

Brown-Chidsey, R., & Steege, M. W. (2004). Discrete Trial Teaching. *Encyclopaedia of School Psychology*, 96-97.

Carter, S. L. (2010). *The Social Validity Manual: A Guide to Subjective Evaluation of Behavior Interventions*. US: Academic Press.

Charania, S. M., LeBlanc, L. A., Sabanathan, N., Ktaech, I. A., Carr, J. E., & Gunby, K. (2010). Teaching Effective Hand Raising to Children with Autism during Group Instruction. *Journal of Applied Behavior Analysis, 43*(3), 493–497.

Chase, P. N., & Danforth, J. S. (1991). The Role of Rules in Concept Learning. In L. J. Hayes & P. N. Chase (Eds.), *Dialogues on verbal behavior* (pp. 205–225). Reno, NV: Context Press.

Cooper, J. O., Heron, T. E., & Heward, W. L. (2007).*Applied Behavior Analysis* (4th ed.). Upper Saddle River, N.J: Pearson/Merrill-Prentice Hall.

Dardig, J. C., Heward, W. L., Heron, T. E., Neef , N. A., Peterson, S., Sainato, D. M., Cartledge, G., Gardner, R., Peterson, L. R., Hersh, S. B. (2005). *Focus on Behavior Analysis in Education: Achievements, Challenges, and Opportunities*. Upper Saddle River, NJ: Pearson/Merrill/Prentice Hall.

DeVries, J.E., Burnette, M.M., & Redmon, W.K. (1991). AIDS Prevention: Improving Nurses' Compliance with glove wearing Through Performance Feedback. *Journal of Applied Behavior Analysis* 24 (4): 705–11.

Dixon, D. R., Tarbox, J., Najdowski, A. C., Wilke, A. E., & Granpeesheh, D. (2011). A Comprehensive Evaluation of language for Early Behavioral Intervention Programs: The Reliability of the SKILLS Language Index. *Research in Autism Spectrum Disorders, 5*(1), 506-511.

Dixon, M. R. (2014a). *The PEAK Relational Training System: Direct Training Module.* Carbondale, IL: Shawnee Scientific Press

Dixon, M. R. (2014b). *The PEAK Relational Training System: Generalization Module*. Carbondale: Shawnee Scientific Press.

Dixon, M. R. (2015). *PEAK Relational Training System: Equivalence*. USA: Shawnee Scientific Press.

Dixon, M. R., Belisle, J., Whiting, S.W., & Rowsey, K. E. (2014). Typically developing Sample of the PEAK Relational Training System: Direct Training Module and Subsequent

Comparisons to Individuals with Autism. *Research in Autism Spectrum Disorders, 8*(11), 1597–1606.

Dixon, M. R., Carman, J., Tyler, P. A., Whiting, S. W., Enoch, M. R., & Daar, J. H. (2014). 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*(5), 603-614.

Dixon, M. R., Small, S. L., & Rosales, R. (2007). Extended Analysis of Empirical Citations with Skinner's Verbal Behavior: 1984-2004. *Behavior Analyst, 30*(2), 197-209.

Dixon, M. R., Small, S. L., & Rosales, R. (2007). Extended Analysis of Empirical Citations with Skinner’s Verbal Behavior: 1984–2004. *The Behavior Analyst, 30*, 197–209

Dixon, M. R., Speelman, R. C., Rowsey, K. E., & Belisle, J. (2015;2016;). Derived rule-following and transformations of stimulus function in a children's game: An application of PEAK-E with children with developmental disabilities. *Journal of Contextual Behavioral Science, 5*(3), 186-192.

Dixon, M. R., Whiting, S. W., Rowsey, K., & Belisle, J. (2014). Assessing the Relationship between Intelligence and the PEAK Relational Training System. *Research in Autism Spectrum Disorders, 8*(9), 1208–1213.

Dunn, L. M., & Dunn, D. M. (2007). *Peabody Picture Vocabulary Test* (4th ed.). San Antonio: Pearson.

Dymond, S., O'Hora, D., Whelan, R., & O'Donovan, A. (2006). Citation Analysis of Skinner's Verbal Behavior: 1984-2004. *Behavior Analyst, 29*(1), 75-88.

Foster, S. L., & Mash, E. J. (1999). Assessing social Validity in Clinical Treatment Research: Issues and Procedures. *Journal of Consulting and Clinical Psychology, 67*(3), 308-319.

Fox, D.K., Hopkins, B.L., & Anger, W.K. (1987). The long-Term Effects of a token Economy on safety Performance in Open-Pit Mining. *Journal of Applied Behavior Analysis 20(3): 215–24.*

Foxx, R. M. (2008). Applied Behavior Analysis Treatment of Autism: The State of the Art. *Child and Adolescent Psychiatric Clinics of North America, 17*(4), 821-834.

Frampton, S. E., Robinson, H. C., Conine, D. E., & Delfs, C. H. (2017). An abbreviated evaluation of the efficiency of listener and tact instruction for children with autism. *Behavior Analysis in Practice,* 1-14.

Grannan, L., & Rehfeldt, R. (2012). Emergent Intraverbal Responses via Tact and Match-To Sample Instruction. *Journal of Applied Behavior Analysis*, 45(3), 601–605.

Green, V. A., Pituch, K. A., Itchon, J., Choi, A., O'Reilly, M., & Sigafoos, J. (2006). Internet Survey of Treatments Used by Parents of Children with Autism. Research in Developmental Disabilities, 27, 70–84.

Greer, R. D., Yaun, L., & Gautreaux, G. (2005). Novel Dictation and Intraverbal Responses as a Function of a Multiple Exemplar Instructional History. *Analysis of Verbal Behavior, 21*(1), 99-116.

Gresham, F. M., & Lopez, M. F. (1996). Social Validation: A Unifying Concept for School-Based Consultation Research and Practice. *School Psychology Quarterly, 11*(3), 204-227.

Grey, I. M., & Hastings, R. P. (2005). Evidence-Based Practices in Intellectual Disability and Behaviour Disorders. *Current Opinion in Psychiatry, 18*(5), 469-475.

Gross, A. C., & Fox, E. J. (2009). Relational Frame Theory: An Overview of the Controversy. *Analysis of Verbal Behavior, 25*(1), 87-98.

Gunby, K. V., Carr, J. E., & Leblanc, L. A. (2010). Teaching Abduction-Prevention Skills to Children with Autism. *Journal of Applied Behavior Analysis*, 43(1), 107–112.

Hagopian, L.P., & Thompson, R.H. (1999). Reinforcement of Compliance with Respiratory Treatment in a Child with Cystic Fibrosis. *Journal of Applied Behavior Analysis* 32 (2): 233–36.

Hayes, S. C., & Wilson, K. G., (1994). Acceptance and Commitment Therapy: Altering the Verbal Support for Experiential Avoidance. *The Behavior Analyst.* 17:289-303.

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.

Hess, K. L., Morrier, M. J., Heflin, L. J., & Ivey, M. (2008). Autism Treatment Survey: Services Received by Children with Autism Spectrum Disorders in Public School Classrooms. *Journal of Autism and Developmental Disorders*, 38, 961–971.

Hill, A. P., Zuckerman, K., & Fombonne, E. (2015). Epidemiology of Autism Spectrum Disorders. In *Translational Approaches to Autism Spectrum Disorder* (pp. 13-38). Springer International Publishing.

Houten, R. V. (1979). Social validation: The Evolution of Standards of Competency for Target Behaviors. *Journal of Applied Behavior Analysis, 12*(4), 581.

Howlin, P. (1997). Interventions for people with autism: recent advances. *Advances in Psychiatric Treatment,* 3, 94–102.

Illinois State Board of Education. (2013). *Illinois Early Learning and Development Standards*. Retrieved January 10, 2016 from http://www.isbe.state.il.us/earlychi/pdf/early_learning_standards.pdf.

Johnson, G., Kohler, K., & Ross, D. (2017;2016;). Contributions of skinner's theory of verbal behaviour to language interventions for children with autism spectrum disorders. *Early Child Development and Care, 187*(3-4), 436-11.

Kazdin, A. E. (1977). Assessing the clinical or applied importance of behavior change through social validation. *Behavior Modification, 1*(4), 427-452.

Koegel, R. L., Russo, D. C., & Rincover, A. (1977). Assessing and training teachers in the generalized use of behavior modification with autistic children. *Journal of Applied Behavior Analysis, 10*(2), 197-205.

Leblanc, L. A., Esch, J., Sidener, T. M., & Firth, A. M. (2006). Behavioral language interventions for children with autism: Comparing applied verbal behavior and naturalistic teaching approaches. *The Analysis of Verbal Behavior, 22*(1), 49-60.

Lovaas, O. I. (1987). Behavioral treatment and normal educational and intellectual functioning in young autistic children. *Journal of Consulting and Clinical Psychology, 55*(1), 3–9.

Lovaas, O. I. (2003). *Teaching individuals with developmental delays: Basic intervention techniques.* PRO-ED, Inc.

Luciano, M. C., Herruzo, J., & Barnes-Holmes, D. (2001). Generalization of say-do correspondence. *Psychological Record, 51*(1), 111-130.

Mayer, G. R., Sulzer-Azaroff, B., & Wallace, M. (2011).*Behavior analysis for lasting change* (2nd ed.). Cornwall-on-Hudson, NY: Sloan Pub.

McKeel, A. N., Dixon, M. R., Daar, J. H., Rowsey, K. E., & Szekely, S. (2015). Evaluating the efficacy of the PEAK relational training system using a randomized controlled trial of children with autism. *Journal of Behavioral Education*, 24, 230–241.

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.

Michael, J. (1984). Verbal Behavior. *Journal of the Experimental Analysis of Behavior, 42*, 363–376.

National Autism Center (2009). National Standards Report. Retrieved from http://www.nationalautismcenter.org/pdf/NAC%20Standards%20Report.pdf.

Osborne, J. W., & Costello, A. B. (2004). Sample size and subject to item ratio in principal components analysis. *Practical Assessment, Research and Evaluation, 9*(11)

Partington, J. W. (2010). *The assessment of basic language and learning skills-revised (the ABLLS-R)* (Version 3.2. ed.). Pleasant Hill, CA: 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*(1), 60-69.

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*. Hoboken, USA: Wiley Subscription Services, Inc.

Reed, D. D., DiGennaro Reed, F. D., Jenkins, S. R., & Hirst, J. M. (2014).The zeitgeist of behavior analytic research in the 21st century: A keyword analysis. *The Behavior Analyst Today, 14*(1-2), 17-25.

Rehfeldt, R. A. (2011). Toward a technology of derived stimulus relations: An analysis of articles published in the journal of applied behavior analysis, 1992–2009. *Journal of Applied Behavior Analysis, 44*(1), 109-119.

Remington, B., Hastings, R. P., Kovshoff, H., Espinosa, F. D., Jahr, E., Brown, T., Ward, N. (2007). Early intensive behavioral intervention: Outcomes for children with autism and their parents after two years. *American Journal on Mental Retardation, 112*(6), 418-438.

Sautter, R. A., & Leblanc, L. A. (2006). Empirical applications of skinner's analysis of verbal behavior with humans. *The Analysis of Verbal Behavior, 22,* 35.

Schechtman, M. A. (2007). Scientifically unsupported therapies in the treatment of young children with Autism Spectrum Disorders. *Pediatric Annals,* 36, 497–505.

Schwartz, I. S., & Baer, D. M. (1991). Social validity assessments: Is current practice state of the art? *Journal of Applied Behavior Analysis, 24*(2), 189-204.

Sidman, M. (1971). Reading and auditory-visual equivalences. *Journal of Speech and Hearing Research, 14*(1), 5–13.

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

Sidman, M., Rauzin, R., Lazar, R., Cunningham, S., Taliby, W., and Carrigan, P. (1982). A search for symmetry in the conditional discriminations of rhesus monkeys, baboons and children. *Journal of Experimental Analysis of Behaviour, 37,* 23-44.

Skinner, B. F. (1938). *The Behavior of organisms: An experimental analysis*. New York: Appleton-Century.

Skinner, B. F. (1957). *Verbal Behavior*. Acton, MA: Copley Publishing Group.

Smith, T. (1999). Outcome of early intervention for children with autism. *Clinical Psychology: Science and Practice*, 6, 33–49.

Stephen Gallagher & Mickey Keenan (2006) Gerontology and applied social technology, *European Journal of Behavior Analysis*, 7:1, 77-86.

Stokes, T. F., & Baer, D. M. (1977). An implicit technology of generalization. *Journal of Applied Behavior Analysis*, *10*(2), 349–367.

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.

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

Sweeney, M. R,, Staines, A. (2016). *Autism Counts: A Report on Autism Spectrum Disorder prevalence estimation in the Republic of Ireland.* Retrieved from Alphagalileo website: http://www.alphagalileo.org/AssetViewer.aspx?AssetId=109368&CultureCode=en

Szatmari, P., Bryson, S., Boyle, M., Streiner, D., & Duku, E. (2003). Predictors of outcome among high functioning children with autism and asperger syndrome. *Journal of Child Psychology and Psychiatry, 44*(4), 520–528.

Tiger, J. H., Hanley, G. P., & Bruzek, J. (2008). Functional communication training: A review and practical guide. *Behavior Analysis in Practice, 1*(1), 16-23.

Van Houten, R., Malenfant, J.E.L., Austin, J., Lebbon, A. (2005). Vollmer, Timothy, ed. The effects of a seatbelt-gearshift delay prompt on the seatbelt use of motorists who do not regularly wear seatbelts. *Journal of Applied Behavior Analysis* 38 (2): 195–203.

Venter, A., Lord, C., & Schopler, E. (1992). A follow-up study of high-functioning autistic children. *Journal of Child Psychology and Psychiatry, 33*, 489–597.

Virués-Ortega, J. (2010). Applied behavior analytic intervention for autism in early childhood: Meta-analysis, meta-regression and dose–response meta-analysis of multiple outcomes. *Clinical Psychology Review, 30*(4), 387-399.

Wolf, M. M. (1978). Social validity: The case for subjective measurement or how applied behavior analysis is finding its heart. *Journal of Applied Behavior Analysis, 11*(2), 203.

Yamamoto, J., & Mochizuki, A. (1988). Acquisition and functional analysis of manding with autistic students. *Journal of Applied Behavior Analysis, 21*(1), 57-64.

Appendix 1

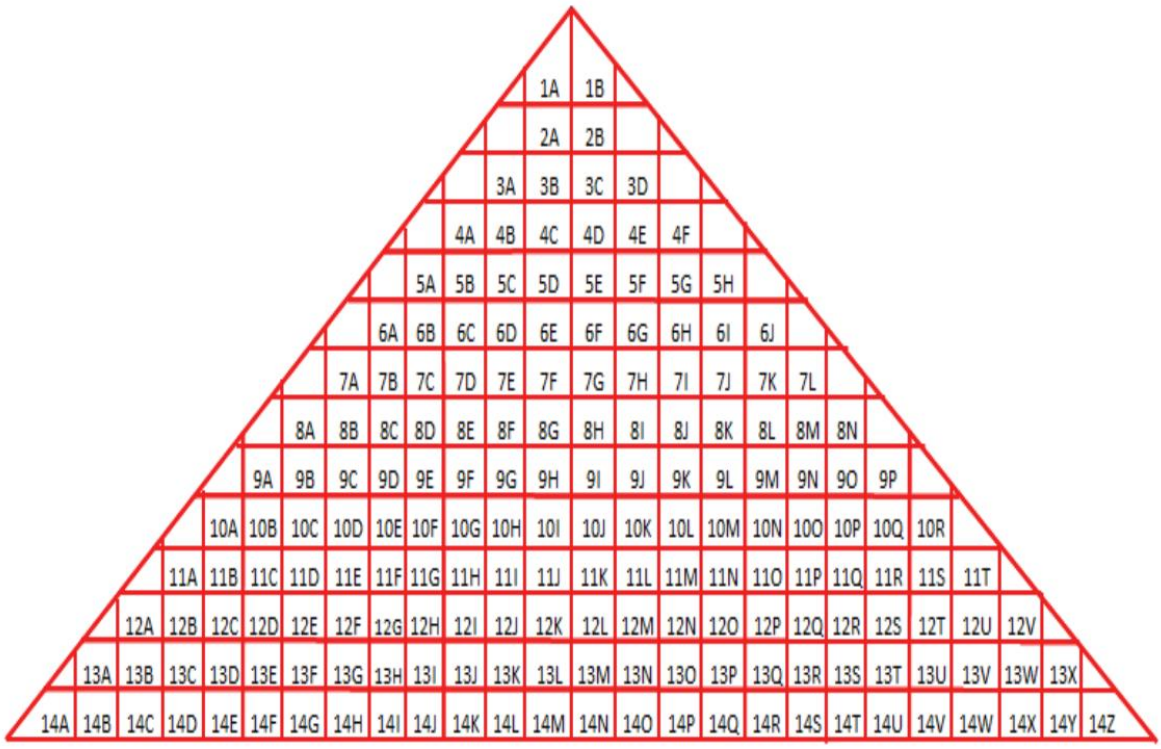

Fig. 1

Above image displays the PEAK generalization training performance matrix. This *triangle* represents the 184 skills of the Generalizatio Training Module and is organized such that 1A is the least complicated skill and 14Z represents the most difficult. Practitioners are advised to begin assessing and teaching skills at 1A and then to progress along the alphanumeric sequence.

Taken from:

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.

Appendix 2

Task	Trail Block 1 Program _____ Correct	Incorrect	Trail Block 2 Program _____ Correct	Incorrect	Trail Block 3 Program _____ Correct	Incorrect	Trail Block 4 Program _____ Correct	Incorrect	Trail Block 5 Program _____ Correct	Incorrect
Begins session with all relevant materials										
Uses appropriate method to determine preferences										
Arranges stimuli correctly for trial										
Gains the attention of the learner										
Establishes motivation prior to trial										
Presents the SD/Question clearly										
Allows 3 seconds for response										
If Discreet Trial Training										
Provides reinforcement if correct										
If incorrect, represents the SD and provides appropriate prompt										
Records trial correctly on data sheet										
Allows appropriate amount of reinforcement										
Minimizes time between trials										

Note. PEAK = Promoting the Emergence of Advanced Knowledge Relational Training Syste