A systematic review of pivotal response treatment (PRT) in children with autism spectrum disorder (ASD)

Kathryn Farr, Allaina McMurrer, Kaitlan Sommerfeldt, Rachel Thomas

Supervisor: Dr. Joanne Volden Reader: Dr. Karen Pollock

A review of PRT in children with ASD

ABSTRACT

Purpose

To review the evidence regarding the effectiveness of Pivotal Response Treatment (PRT) as a treatment for Autism Spectrum Disorder (ASD).

Methods

This review employed the systematic review process developed by the American Academy for Cerebral Palsy and Developmental Medicine (AACPDM). A comprehensive search of eight databases was performed. Studies were included if: 1) participants had a primary diagnosis of ASD; 2) PRT was the primary intervention; and 3) participants (receiving treatment) were under 18 years of age. Articles which focused solely on training techniques and did not report on relevant behavioural changes were excluded.

Results/Discussion

Twenty studies were included in this review**.** Many studies demonstrated a positive trend in outcomes, clearly defined the independent and dependent variables, reported inter- and/or intra-rater reliability, and provided sufficient details for clinical replicability. However, several weaknesses were identified in the majority of studies reviewed, including but not limited to: 1) highly variable intervention conditions and independent variables**,** making comparison across studies difficult**;** 2) lack of sufficient statistical analysis; 3) lack of standardized assessments; and 4) possible bias with regards to participant selection.

Conclusions

Further research is needed on the effectiveness of PRT. Studies with larger sample sizes, greater adherence to rigorous methodology, and use of standardized assessments are required.

INTRODUCTION

Pivotal Response Treatment (i.e., PRT) is a behavioural intervention developed for use with children with autism spectrum disorder (ASD). Developed by Drs. Robert and Lynn Koegel, PRT was developed from the principles of applied behavioural analysis (ABA). It resembles ABA interventions in that it analyzes behaviour in terms of stimuli that precede (i.e. antecedents) and follow (i.e. consequences) the behaviour of interest. Antecedent stimuli are conceptualized as eliciting the behaviour of interest, which is then strengthened and maintained by administering reinforcing consequences. In PRT, the antecedent stimulus is the teaching instruction or action; the behaviour is the target response from the child; and the consequence is the direct natural reinforcer. Unlike traditional ABA, PRT is a more naturalistic approach. Instead of repetitive, drill like activities, intervention is done during naturally occurring teaching opportunities throughout the day in the child's typical environments (e.g., home, school, community, etc.) The fundamental difference between ABA and PRT is that PRT targets "pivotal" behaviours, which are areas of functioning that when fostered result in changes in more widespread behaviours, including language and social communication, symbolic play, and academic skills, as well as a decrease in disruptive behaviours (http://education.ucsb.edu/autism/prt.html; Koegel, Carter, & Koegel, 2003; Stahmer, 1995; Stahmer, 2006).

"Pivotal" behaviours include motivation, social initiation, self-management, and responsivity to multiple cues (Koegel & Frea, 1993; Koegel, Carter, & Koegel, 2003). For example, increasing a child's motivation to respond is regarded as pivotal because it leads to increased exposure to complex stimuli and as a result, increases learning opportunities (Koegel

et al., 2003). Social initiations are also an essential pivotal area. Koegel et al. (1999) showed that an increased frequency of initiations was an important prognostic indicator of positive outcomes in language, cognitive, and social development. Therefore, targeting either of these two areas would have an impact on multiple behaviours.

Strategies used in PRT to improve motivation to participate include: 1) providing the child with choices; 2) providing clear instructions (after ensuring that the child is paying attention); 3) reinforcing approximations/attempts of the target behaviour; 4) using direct and natural reinforcement; 5) varying tasks; and 6) interspersing tasks for maintaining previously learned behaviours into the session. By providing children with choices regarding the activities and toys used during interactions, the child directs the interaction so his/her motivation to respond is increased. Approximations of the target are reinforced in order to shape behaviours until they are the desired response. Maintenance tasks, in which the child is already proficient, are interspersed with new tasks so that the child experiences success frequently, thus increasing the child's motivation to attempt new tasks. Reinforcement is direct and natural, meaning that the reinforcement is directly linked to the target behaviour (e.g. if the child asks for bubbles, the reinforcement is having access to the bubbles/giving the child the bubbles). Reinforcement should be immediate so that the child learns the connection between the behaviour and the consequence (i.e. the reinforcement).

Over two decades have passed since Drs. R. L. Koegel, M. C. O'Dell, and L. K. Koegel (1987) outlined the early principles of PRT. During that time, PRT has grown as an intervention, with research on the topic revolving around the success of PRT as an intervention and the success of modifications to the original intervention format (e.g., expanding the scope of who

may be trained to provide the intervention and what "pivotal" behaviours the intervention may target). However, at the time of this review's completion, no systematic review of the literature had been published. This is unfortunate because it is then difficult to evaluate the utility of PRT as an intervention approach. The purpose of this systematic review is to provide those working in the field of ASD research and/or intervention with a summary of the literature to date regarding the effectiveness of PRT, in the hopes of benefitting future research and intervention.

METHODS

This systematic review employed the methodology laid out by Darrah, Hickman, O'Donnell, Vogtle and Wiart (2008), in their American Academy for Cerebral Palsy and Developmental Medicine *(*AACPDM) Methodology to develop systematic reviews of treatment interventions (revision 1.2).

Using this method, outcome measures were sorted into three categories under the World Health Organization - International Classification of Functioning, Disability and Health (WHO-ICF). The categories were: 1) Body structures/functions; 2) Activities and participation; and 3) Contextual factors. To code the outcome measures for the aforementioned categories, the reviewers followed guidelines from *The indicators for intervention coding manual* developed by the National Allied Health Classification Committee (2007).

Criteria for considering studies for this review

Types of studies. The following types of studies were considered: 1) Cohort studies with concurrent or historical control groups; 2) Randomised single subject research designs with multiple baseline design; and 3) Non-randomised single subject research designs with multiple baseline design.

***Types of participants*.** Studies with participants that met the following criteria were considered: 1) Children and adolescents aged 2;0 to 16;0; and 2) children diagnosed with ASD (i.e., Autism, Asperger Syndrome, Autism Spectrum Disorder, Pervasive Developmental Disorder (PDD), PDD Not Otherwise Specified (PDD-NOS), Kanner Syndrome, or Childhood Disintegrative Disorder) as the primary diagnosis.

***Types of intervention*.** Studies with interventions utilising Pivotal Response Training (PRT) as the primary intervention were included if: 1) if they stated that PRT, or PRT techniques were used; or 2) the study described an intervention consistent with the principles of PRT and referenced either the original Koegel, O'Dell, and Koegel (1987) work or the PRT training manual developed by Koegel et. al. (1989). Studies that measured behavioural changes in the participants as well as studies measuring the success of the provider in delivering PRT were included. Studies that included alternative treatment groups were also included, provided that the children receiving PRT intervention did not receive any other form of treatment at the same time, and that appropriate outcome measures were reported.

***Types of outcome measures*.** Studies which included measures for specific behaviours such as sentence length, number of communication initiations, or reciprocal play, or that reported change in terms of standardized assessment scales were included.

***Primary outcomes*.** The primary outcomes of participants' specific behaviours were related to social communication, language, and behaviour. These included, but were not limited to, measures of initiating conversation, maintaining reciprocal interaction, topic maintenance; a complete list can be found in Table 2. Primary outcomes were measured pre- and post-intervention for all studies. Single subject research design (SSRD) studies included multiple

measures during each phase.

***Secondary outcomes*.** Secondary outcomes included, but were not limited to, parent or teacher reports focusing on emotion regulation and disruptive behaviours.

Search methods for identification of studies

***Electronic searches*.** Published studies in peer reviewed journals were considered for inclusion. Searches were run in March 2012 and again in June 2012, restricted to English-only results with no date limit to ensure the earliest articles referencing PRT were included.

***Search strategy design*.** The following search criteria were entered into all relevant databases: Pivotal response AND (Autism OR ASD OR Asperger* OR PDD OR PDD-NOS OR Pervasive developmental disorder OR Childhood disintegrative disorder OR Kanner* OR Speech or communication disorder).

The following electronic databases were last searched June 4, 2012 using the above search terms for each:

- CINAHL (Cumulative Index to Nursing and Allied Health Literature – 1981 to present)
- ComDisDome (Communication Sciences and Disorders Information Service – 1950 to present)
- Linguistics and Language Behaviour Abstracts (1973 to present)
- MEDLINE via EBSCOhost (1997 to present)
- MEDLINE via Ovid (1997 to present)
- PsychINFO via Ovid (1806 to present)
- SciVerse Scopus (1966 to present)
- Web of Science (1900 to present)

Searching other resources

***Reference lists*.** Reference lists of included studies were reviewed in order to identify additional studies not obtained by the electronic search.

***Forward searching*.** Forward searching, using Web of Science, (1987 to present) of the initial Koegel et al. (1987) and Koegel et al. (1989) works was done to find additional works citing these papers to identify relevant articles. This was last searched June 4, 2012.

Data collection and analysis

***Selection of studies*.** The search strategy yielded 248 records. The titles and abstracts of the citations generated by the search strategy were screened against the inclusion criteria listed above. Full text was obtained and screened when the abstract and title did not provide sufficient information to determine the study's inclusion or exclusion status. A single reviewer was responsible for screening each article. Of the 248 articles obtained in the search, 91 were considered relevant. Duplicate articles, articles which focussed on the training techniques but did not report behavioural changes in the child participants, or articles which were not research studies were eliminated. After these exclusions, 18 single subject research design studies and two group studies qualified for inclusion. The reviewers were not blind to the study authors, institutions, or to the publication journals.

***Data extraction and management*.** Two reviewers independently extracted data from each of the 20 included studies using "Study data extraction summary forms" included in Darrah et al. (2008) which document the level of evidence, participants, intervention, quality, measures, outcomes, and adverse events. Disagreements were resolved through discussion with all reviewers.

Table 1a lists the guidelines used to assign a level of evidence to each group study that was reviewed. Table 1b describes the corresponding guidelines for the single subject design studies that were reviewed.

Table 1a
Levels of Evidence (Group Studies)

Level	Intervention (Group) Studies
I	Systematic review of randomized controlled trials (RCTs) Large RCT (with narrow confidence intervals) (n>100)
II	Smaller RCTs (with wider confidence intervals) (n<100) Systematic reviews of cohort studies "Outcomes research" (very large ecologic studies)
III	Cohort studies (must have concurrent control groups) Systematic reviews of case control studies
IV	Case series Cohort study without concurrent control group (e.g. with historical control group) Case-control study
V	Expert opinion; Case study or report Bench Research Expert opinion based on theory or physiologic research Common sense/anecdotes

Table 1b
Levels of Evidence (Single Subject Design Studies)

Level	Single Subject Design Studies
I	Randomized controlled N-of-1 (RCT), alternating treatment design (ATD), and concurrent or non-concurrent multiple baseline design (MBDs); generalizability if the ATD is replicated across three or more subjects and the MBD consists of three of a minimum of three subjects, behaviours, or settings. These designs can provide causal inferences.
II	Non-randomized, controlled, concurrent MBD; generalizability if design consists of a minimum of three subjects, behaviours or settings. Limited causal inferences.
III	Non-randomized, non-concurrent, controlled MBD; generalizability if design consists of a minimum of three subjects, behaviours or settings. Limited causal inferences.
IV	Non-randomized, controlled SSRDs with at least three phases (ABA, ABAB, BAB, etc.); generalizability if replicated across three or more different subjects. Only hints at causal inferences.
V	Non-randomized, controlled AB SSRD; generalizability if replicated across three or more different subjects. Suggests causal inferences allowing for testing of ideas

(Darrah et al., 2008)

Table 2 summarizes each group study and single subject study (SSS) that were included in this review, regardless of the level of evidence assigned to it. The table is arranged in chronological order indicating the primary author's name, the year of publication, the level of

evidence assigned to that study and the corresponding conduct rating. Conduct ratings are a measure of the strength of a study's research design, in terms of the parameters followed and the information reported in the publication. These were determined by answering the conduct rating questions found in Table 3b, outlined by Darrah et al. (2008). The results of the conduct rating questions provided each article with a ranking of weak (W), moderate (M), or high (H). The table also describes the parameters of the population, the total number of participants and their ages, and the type of intervention.

Table 3a and 3b show the answers to each conduct rating question for each study that was assigned a level of evidence I, II, or III. The conduct rating questions are listed above Table 3a and 3b. If "yes" is entered in the table, it indicates that the criterion/criteria referred to in the corresponding conduct question was/were present. Conversely, a "no" indicates the corresponding criterion/criteria was/were not present. For group studies, a conduct rating of 6-7 warranted a ranking of strong, 4-5 indicated a ranking of moderate, and 3 or less indicated a ranking of weak. For single subject designs, a conduct rating of 11-14 indicated a ranking of strong, 7-10 indicated a ranking of moderate, and score of less than 7 indicated a ranking of weak.

Table 4 showcases the outcomes measured in each study, the measurement tool used, and the results of each study. This table also displays the distribution of outcomes under the aforementioned WHO-ICF categories. Once again, only studies with a level of evidence of III or lower are included in this table, group studies and single subject studies are listed in separate sections, and studies are arranged in chronological order. The Components of Health column lists the statistical results found for each outcome of interest, a report of "nr" indicates that

statistical results were not reported in the study.

RESULTS

The results of this systematic review are organized in accordance with the guidelines proposed by Darrah, et al. (2008). For a description of tables, refer to the methods section.

Table 2
Summary of Studies

Group Studies	LOE[1] & CR[2]	Participants	n	Ages	Intervention	Control intervention
2006 Stahmer	III – M (5/7)	Males diagnosed with autism presenting with Expressive One-Word Picture Vocabulary Test (EOWPVT) AES[3] ranging from 2;7 – 4;10, NVIQ[4] ranging from 64 – 111, and deficits in symbolic play.	6	4;3 – 7;2 y	PRT used to teach symbolic play. Each child was video-taped playing with same adult for 7 minutes pre- and post-treatment.	Controls participated in the same video-taped play session but did not receive treatment.
2010 Nedft	II – M (5/7)	Primary caretakers of children with autism who had less than 20 functional words.	27	< 60 m	Parents learned PRT from a self-study DVD and interactive tasks. 5 aspects of PRT were scored during 10 minute videos of parents and their children.	Waitlist condition: 5 aspects of PRT were scored during 10 minute videos of parents and their children.

SSS[5]	LOE &	Participants	n	Ages	Intervention

[1] Level of evidence
[2] Conduct rating
[3] Age-equivalent scores
[4] Non-verbal IQ
[5] Single subject studies

	CR				
1993 Koegel	II – M (8/14)	Males with an early diagnosis of autism. IQ above 70, no atypical physical characteristics or handicaps, part of a normal education classroom for at least part of the day, and functioning successfully in school.	2	13 & 16 y	Baseline Measures obtained during normal conversation. PRT used to train participants to differentiate appropriate vs. inappropriate instances of target behaviours.
1995 Pierce	II – M (7/14)	Males diagnosed with autism, expressive language AES 3;2 and 3;5, receptive language AES of 3;1 and 5;5. Typical: age matched pairs.	4	10 y	Baseline data taken over 2 months. Peer training took place over 2 weeks. Intervention consisted of 10 minute play sessions. Post-treatment data taken immediately after intervention period. Generalization data taken after 2 months.
1995 Stahmer	III – M (10/14)	Children diagnosed with autism, expressive language AES from 2;7 to 4;10, receptive language AES from 2;5 to 3;7, NVIQ from 48 to 82. Typical peers.	14	4 – 7 y	Baseline measures were obtained prior to symbolic play training (SPT) or language training (LT) occurred. 5 children received 3 hours of SPT first. Children with autism also underwent 3 hours of LT.
1995 Thorp	III – W (6/14)	Males diagnosed with autism, expressive language AES of 3;7 and 5;2, NVIQ of 47, 58, and 78.	3	5 – 9 y	Baseline measures included standardized tests, play-history interview, and video. 16 hours of intervention in total.
1997a Pierce	II – W (3/14)	Males diagnosed with autism, NVIQ of 50 and 76. Typical peers.	10	7 – 9 y	Baseline data taken over several weeks to 2 months. Intervention provided 1-2 times/day.
1997b	II – M	Males diagnosed	10	7-8 y	Baseline data obtained in up to

Pierce	(7/14)	with autism, expressive language AES of 3;6 and 4;6,receptive language AES 2;11 and 3;3, NVIQ of 50 and 76. Typical peers.			2 months. Peers trained with 4 x 30 min sessions for 2 weeks. Post-treatment assessment taken over 3 months. Generalization & follow-up probes taken.
2002 Koegel	III – M	Five families of middle to upper-middle socioeconomic status with children diagnosed with autism.	5	3;10 – 5;7y	An informal telephone interview. Intervention consisted of 5 hours/day over 5 consecutive days. Data was collected through home videos during the baseline, intervention, and follow-up phases.
2003 Koegel	II – W (6/14)	Males diagnosed with autism, expressive language AES 3;9 an 7;2, receptive language AES 3;8 and 5;3, language development AES 2;2 and 5;3, NVIQ 96 and 107.	2	4;4 and 6;3	Baseline included home and in-clinic language samples & probes. Intervention included children being taught to self-initiate a question to evoke a target temporal morpheme. Unstructured language samples used to judge generalization.
2005 Sherer	II – M (10/14)	Children diagnosed with autism. Divided into responders and non-responders.	6	3 – 5;10	Baseline sessions occurred 3x/day; 4-5x/week. Intervention consisted one-on-one PRT 4-5x/week. Responders received treatment for 6 months; non-responders were referred to a different treatment program after 5 weeks.
2007 Baker	III – M (10/14)	Families enrolled in a parent education program at a children's hospital in Southern California from 1999-2003 with a child with either a diagnosis of ASD or PDD.	158	2;0 – 9;4 y	12 week parent education program; families met with a therapist 1h/week. Families read a training manual, completed teaching activities, and discussed strategies. Child assessments were completed on the first and last day of treatment.

2008 Harper	II – W (5/14)	Males diagnosed with autism and typical peers.	8	8 -9 y	Baseline data taken for 13-18 days. Peer training took place over 7, 20 minute sessions. 7 consecutive days of intervention in triads of 2 typical peers and 1 child with ASD. 4-5 generalization probes were taken.
2008 Kuhn	II – W (2/14)	Males diagnosed with autism, deficits in both receptive and expressive language. Peers with various types of delays.	7	7 – 8 y	Baseline data taken over 10 minute sessions. Peer training took place in 20 minute sessions 2-3x/week for 10 sessions until accuracy of strategy use reached 80%.
2009 Koegel	IV	Males diagnosed with autism, deficits in language, eye contact, and affect, adaptive behaviour composite ages ranging from 0;11 – 1;9.	3	3;2 – 3;5 y	Phase A - PRT in a non-embedded reinforcer condition; 3-5, 2 hour sessions 1x/week. Phase B – PRT in an embedded social condition; 4-6, 2 hour sessions 1x/week.
2010 Coolican	III – M (10/14)	Males and females diagnosed with autism with Preschool Language Scale (PLS) auditory comprehension AES from 0;7 – 3;10 and PLS expressive communication AES from 1;3 – 2;11.	8	2;4 – 4;8 y	Parents received 3, 2 hour PRT training sessions over 2 weeks (6 hours PRT training). Parents were introduced to PRT and techniques were modeled before parents implemented PRT with their children and received feedback.
2010 Smith	IV	Children with autism, diagnosis based on Autism Diagnostic Observation Schedule (ADOS), Autism Diagnostic Interview-Revised (ADI-R), & DSM –	53	2.08 – 6 y	Baseline assessments taken in the clinic or child's home. 1st cohort received PRT training for parents in a hands-on fashion while working with their own children. 2nd cohort received in-home parent training and one-to-one home, preschool, and/or daycare intervention.

		IV-TR[6].			Assessments conducted 6 & 12 months post-treatment.
2011 Randolph	II – W (6/14)	Children diagnosed with autism. Caregivers included parents, home care providers, and grandparents.	3	3 -7 y	Baseline probes taken 2x/week for 30 minute sessions. 4-6 sessions were completed until a stable baseline was obtained. PRT intervention implemented 2x/week for 10 weeks. Sessions lasted 30-55 minutes. Follow-up measures were taken 2 weeks post intervention.
2011 Robinson	II – M (9/14)	Paraprofessionals, focal students, generalization students, and trainers. All focal students were males who had received a diagnosis of autism.	13	3-8 y	2-9 baseline probes taken over 1-6 weeks. 10-15 minute sessions and occurred 1-3x/week. A total of 45 minutes of PRT was conducted by trainer while paraprofessionals observed. Video feedback provided until paraprofessionals achieved at least 80% fidelity across 2 consecutive probes. Generalization and follow-up probes taken 4-8 weeks post-treatment.
2011 Minjarez	IV	Children diagnosed with ASD and their parents. Evidence of language delay with the ability to make contingent vocalizations.	17	2 – 6;11 y	1-3, 10 minute baseline videos taken by parents of themselves interacting with their children. Parent training consisted of 10, 90 minute sessions and 1, 50 minute individual session.

Table 3a

Conduct Questions of Group Design Studies

Study	***Level/Quality***	1	2	3	4	5	6	7
2006 Stahmer	III – M (5/7)	Yes (Y)	No (N)	Y	Y	Y	Y	N
2010 Nedft	II – M (5/7)	Y	Y	Y	Y	N	N	Y

[6] Diagnostic and Statistical Manual of Mental Disorders (fourth edition, text revised)

Conduct Questions

1. Were inclusion and exclusion criteria of the study population well described and followed?
2. Was the intervention well described and was there adherence to the intervention assignment? (for 2-group designs, was the control exposure also well described?) Both parts of the question need to be met to score 'yes'.
3. Were the measures used clearly described, valid and reliable for measuring the outcomes of interest?
4. Was the outcome assessor unaware of the intervention status of the participants (i.e., were the assessors masked)?
5. Did the authors conduct and report appropriate statistical evaluation including power calculations? Both part of the question need to be met to score 'yes'.
6. Were dropout/loss to follow-up reported and less than 20%? For 2-group designs, was dropout balanced?
7. Considering the potential within the study design, were appropriate methods for controlling confounding variables and limiting potential biases used? (Darrah et al., 2008)

Table 3b

Conduct Questions for Single Subject Design Studies

Study	***Level/ Quality***	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1993 Koegel	II – M (8/14)	Y	Y	Y	Y	Y	N	N	Y	Y	N	N	Y	N	N
1995 Pierce	II – M (7/14)	Y	Y	N	Y	Y	N	Y	Y	N	N	N	Y	N	N
1995 Stahmer	III – M (10/14)	Y	Y	Y	Y	N	N	N	Y	N	Y	Y	Y	Y	Y
1995 Thorp	III – W (6/14)	Y	Y	Y	N	N	N	N	Y	N	Y	N	Y	N	N
1997a Pierce	II – W (3/14)	N	N	N	Y	N	N	N	Y	N	N	N	Y	N	N
1997b Pierce	II – M (7/14)	Y	Y	Y	Y	Y	N	N	Y	N	N	Y	N	N	N
2002 Koegel	III – M (9/14)	Y	Y	Y	Y	Y	Y	N	Y	N	Y	N	Y	N	N
2003 Koegel	II – W (6/14)	Y	N	Y	Y	Y	N	Y	Y	N	N	N	N	N	N
2005 Sherer	II – M (10/14)	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	N	Y	N	N
2007 Baker	III – M (10/14)	Y	Y	Y	Y	N	N	Y	N	N	Y	Y	Y	Y	Y
2008 Harper	III – W (5/14)	N	Y	Y	Y	Y	N	N	Y	N	N	N	N	N	N
2008	II – W	N	N	N	Y	N	N	N	Y	N	N	N	N	N	N

Kuhn	(2/14)														
2010 Coolican	III – M (10/14)	Y	Y	N	Y	Y	Y	N	Y	N	Y	Y	N	Y	Y
2011 Randolph	II – W (6/14)	N	Y	Y	Y	Y	N	N	Y	N	N	N	Y	N	N
2011 Robinson	II – M (9/14)	Y	Y	Y	Y	Y	Y	N	Y	N	Y	N	Y	N	N

Conduct Questions

1. Was/were the participant(s) sufficiently well described to allow comparison with other studies or with the readers own patient population?
2. Were the independent variables operationally defined to allow replication?
3. Were the intervention conditions operationally defined to allow replication?
4. Were the dependent variables operationally defined as dependent measures?
5. Was inter-rater or intra-rater reliability of the dependent measures assessed before and during each phase of the study?
6. Was the outcome assessor unaware of the phase of the study (intervention versus control) in which the participant was involved?
7. Was stability of the data demonstrated in baseline, namely lack of variability or a trend opposite to the direction one would expect after application of the intervention?
8. Was the type of SSRD clearly and correctly stated, for example, A-B, multiple baseline across subjects?
9. Were there an adequate number of data points in each phase (minimum of five) for each participant?
10. Were the effects of the intervention replicated across three or more subjects?
11. Did the authors conduct and report appropriate visual analysis, for example, level, trend and variability?
12. Did the graphs used for visual analysis follow standard conventions, for example x- and y-axes labeled clearly and logically, phases clearly labeled (A, B, etc.) and delineated with vertical lines, data paths separated between phases, consistency of scales?
13. Did the authors report tests of statistical analysis, for example celeration line approach, two-standard deviation band method, C-statistic, or other?
14. Were all criteria met for the statistical analyses used? (Darrah et al., 2008)

Table 4

Summary of Studies: Outcomes, Measures, and Results

Group Studies	Outcome of interest	Measure	Components of Health		
			Body Structures /Functions	Activities and Participation	Context-ual Factors
2006 Stahmer	Overall play ability	VO[7] – 6 pt Likert scale		$p < 0.05$ PT[8]	

[7] Video observation

	Creativity	VO – 6 pt Likert scale		p < 0.05 PT	
	Enjoyment	VO – 6 pt Likert scale		p < 0.05 PT	
	Social interaction	VO – 6 pt Likert scale		p < 0.05 PT	
	Play complexity	VO – 6 pt Likert scale		p < 0.05 PT	
2010 Nedft	Functional verbal utterances	VO – Scoring of TI[9]		p = 0.001 TG[10]	

Single Subject Design Studies	Outcome of interest	Measure	Components of Health		
			Body Structures /Functions	Activities and Participation	Context-ual Factors
1993 Koegel	Eye gaze	VO – Scoring of TI		nr	
	NV[11] mannerisms	VO – Scoring of TI		nr	
	Voice volume	VO – Scoring of TI	nr		
	Perseveration of topic	VO – Scoring of TI		nr	
	Facial expressions/ affect	VO – Scoring of TI		nr	
1995 Pierce	MI[12]	VO – Scoring of TI		nr	
	Play initiation	VO – Scoring of TI		nr	
	CI[13]	VO – Scoring of TI		nr	
	Object engagement	VO – Scoring of TI		nr	
	Supported JA[14]	VO – Scoring of TI		nr	
	Coordinated JA	VO – Scoring of TI		nr	
	Non-engagement	VO – Scoring of TI		nr	
	On-looking	VO – Scoring of TI	nr		
	NAWS[15]	VO – Scoring of TI	nr		
	Sentence length	VO – Scoring of TI	nr		
1995	Play behaviour	VO – Scoring of TI		p < 0.01	

[8] Post-training
[9] Timed intervals
[10] Treatment group
[11] Non-verbal
[12] Maintaining interactions
[13] Conversation initiation
[14] Joint-attention
[15] Number of appropriate words spoken

Stahmer	Play complexity Type of response Number of initiations Functioning level Language use (home)	VO – Scoring of TI VO – Scoring of TI VO – Scoring of TI SB4[16] & LIPS[17] CDI[18]	 nr	nr nr nr nr	
1995 Thorp	Play behaviour Speech/language Social behaviour	VO – Scoring of TI PPVT-R[19] & EOWPVT-R[20] VO – Scoring of TI	 nr	nr nr	
1997a Pierce	MI Play initiation CI	VO – Scoring of TI VO – Scoring of TI VO – Scoring of TI		nr nr nr	
1997b Pierce	MI Initiations NAWS Sentence length Toys played with (#) Duration of toy play	VO – Scoring of TI VO – Scoring of TI VO – Scoring of TI VO – Scoring of TI VO – Scoring of TI VO – Scoring of TI	 nr 	nr nr nr nr nr	
2002 Koegel	Functional verbal responses	VO		nr	
2003 Koegel	Prod.[21] target tense MLU[22] Prod. of target query Res.[23] target tense Generalized query Total verbs produced	Observation – V/O[24] Observation – V/O Observation – V/O Observation – V/O Observation – V/O Observation – V/O Observation – V/O	nr nr nr nr nr nr	 nr 	

[16] Stanford-Binet Performance Scales – Fourth Edition
[17] Leiter International Performance Scale
[18] MacArthur Communicative Development Inventory
[19] Peabody Picture Vocabulary Test-Revised
[20] Expressive One-Word Picture Vocabulary Test-Revised
[21] Production
[22] Mean length of utterance
[23] Response with
[24] Video or online

	Diversity of verbs				
2005 Sherer	IQ	DAS[26]	nr		
	Cognitive functioning	BSID-2[27]	nr		
	Non-verbal IQ	LIPS	nr		
	Language	CDI, PPVT-R, & EOWPVT	nr		
	Adaptive functioning	VABS[28]		nr	
	Autism severity	CARS[29]	nr		
	Functional play	VO – Scoring of TI		nr	
	Symbolic play	VO – Scoring of TI		nr	
	Varied play	VO – Scoring of TI		nr	
	Echolalia	VO – Scoring of TI	nr		
	Verbally cued speech	VO – Scoring of TI	nr		
	NV cued speech	VO – Scoring of TI	nr		
	Spont.[25] utterances	VO – Scoring of TI		nr	
	MI	VO – Scoring of TI		nr	
	Social initiations	VO – Scoring of TI		nr	
2007 Baker	Communication	VABS		$p < 0.001$	
	Daily living skills	VABS		$p < 0.001$	
	Socialization	VABS		$p < 0$ .001	
	Motor skills	VABS	$p < 0.001$		
	ABC[30]	VABS		$p < 0.001$	
2008 Harper	Gaining attention	Observation – online		nr	
	Turn-taking	Observation – online		nr	
	Play initiations	Observation – online		nr	
2008 Kuhn	RTP[31]	Video		nr	
	Rate of RTP	Video		nr	
	Initiations	Video		nr	
2009 Koegel	Non-verbal dyadic orienting	Video - scoring of TI		nr	

[25] Spontaneous
[26] Differential Abilities Scales
[27] Bayley Scales of Infanct Development – Second Edition
[28] Vineland Adaptive Behaviour Scales
[29] Childhood Autism Rating Scale
[30] Adaptive behaviour composite
[31] Responses to prompts

	(eye contact) Self-Initiated social engagement during communication General Child Affect	 Video – scoring of TI Video – 6 pt. Likert Scale		 nr nr	
2010 Coolican	Language Functional verbal utterances Appropriate utterances Disruptive behaviour Model prompt Indirectly prompted Inappropriate responses No responses Initiations	PLS-4 PPVT-III VO – Scoring of TI VO – coding of FS[32] VO – Scoring of TI VO – Scoring of TI VO – Scoring of TI VO – coding of FS VO – coding of FS VO – coding of FS	p > 0.05 PT p = 0.11 FOL[33] p = 0.11 FOL	 p < 0.05 PT p > 0.05 FOL p > 0.05 PT p > 0.05 FOL p > 0.05 PT p > 0.05 FOL p > 0.05 PT - FOL p < 0.05 PT p > 0.05 FOL p > 0.05 PT - FOL p < 0.05 PT p > 0.05 FOL p < 0.05	
2011 Rand-olph	Com.[34] responses NV responses CI Appropriate play Inappropriate play Varied play Communication Daily living skills Socialization Motor skills	Online coding Online coding Online coding Online coding Online coding Online coding VABS-2 VABS-2 VABS-2 VABS-2	 nr 	nr nr nr nr nr nr nr nr nr	

[32] Full sample
[33] Follow-up
[34] Communicative

	Social validity	Questionnaire - 5 pt. Likert Scale		nr	
2011 Robinson	RVI[35] Word combinations (#) Verbal requests (#) PDV[36] Student affect	Video observation Video observation Video observation Video observation Video observation – 6 pt. Likert Scale	nr	nr nr nr nr	

(Darrah et al., 2008)

DISCUSSION

The results of this systematic review revealed that the majority of the evidence in PRT research had poor levels of evidence and conduct ratings; however, the reviewers noted that the strength of evidence of the included articles was on par with the field of speech-language pathology as a whole (Dodd, 2007). For the 18 SSRD studies, the mean level of evidence was 2.6 with a range of two to four. Of these, 15 were eligible for analysis using the conduct questions, which was applicable for studies with level of evidence rating of one to three. For the two group studies, the mean level of evidence was 2.5 with a range of two to three.

Overall, results of PRT intervention were reported to be positive. All 18 SSRD studies and both group studies reported that subjects demonstrated functional improvements of the dependant variables.

Strength of existing evidence

The following themes were found to be strengths of the studies:

[35] Number of reciprocal verbal interactions

[36] Number of spontaneous peer-directed verbalizations

Description of variables. Fourteen out of 15 SSRD studies satisfactorily identified the dependant variables, and 12 out of 15 SSRD studies satisfactorily identified the independent variables.

Reliability. Ten out of 15 SSRD studies reported that inter-rater and/or intra-rater reliability of dependant measures was assessed before and during each phase of the study.

Clinical relevance. Strengths in this area were found in 1) recency of publication, 2) clinical replicability, and 3) successful implementation on intervention by frequent communication partners.

Twelve out of the 18 total SSRD studies and both group studies were published since 1999. Additionally, the methodology in several studies (e.g., participant description, variables, and intervention conditions) was described well enough to be clinically replicated. Lastly, almost all studies incorporated parents, caregivers, typically developing peers, and/or peers with developmental delays or learning disorders. These participants were trained to implement PRT with the child with ASD, and as mentioned, most studies indicated this was successful. This is promising as it increases the number of people capable of implementing PRT, and thereby, the amount of time that the child receives intervention.

WHO-ICF. After coding each of the 108 outcome measures by ICF component, the reviewers found that 27 dependent variables fell under the body structures and functions category, 81 fell under the activity/participation category, and none fell under the contextual factors category. This helped to frame the dependent measures as outcomes that have an impact on quality of life, and not simply measurements of statistical significance. Also, several

studies indicated that intervention resulted in the desired change of the dependent variables, which suggests improved quality of life when interpreted using the ICF model.

Weakness of existing evidence

Despite a reported trend of improvement in dependent measures, the strength of these findings is limited due to the variability in study design and overall weak methodology of the studies considered. Highly variable intervention conditions makes comparison across studies difficult and has resulted in a small body of research for any one type of PRT intervention. The quality of group design studies was better overall than that of SSRD studies; however, very few group studies have been conducted on this subject. This is reflected by the fact that only two group design studies met criteria for this analysis.

Inconsistent intervention targets. The SSRD studies included in this review tended to identify child specific goals and intervention targets. While this allowed for functional and patient-centred intervention, it resulted in a wide variety of variables being studied. Few studies shared any significant overlap of intervention targets. Those studies which did consider similar kinds of behaviour used unique terminology and differing definitions of behaviour which were specific to each participant instead of conforming to a vocabulary shared by a larger body of evidence. This made comparing outcome measures between studies very difficult as each study investigated functionally different phenomena. Therefore, it is difficult to determine the overall strength and quality of the evidence for PRT as an intervention for general behaviours associated with ASD.

Insufficient statistical measures. The amount of time a participant spent in the baseline, treatment, or follow-up phases differed both between studies and between participants within

studies. Overall, SSRD studies did not contain sufficient data during each phase to demonstrate a statistically relevant intervention effect. Just two of 15 SSRD studies included in this review contained sufficient data in each treatment phase and only three of 15 SSRD studies demonstrated baseline stability of target behaviours before intervention. Only three of 15 SSRD studies reported any statistical analyses at all. This can be contrasted with the group design studies included in this review which both included analyses of statistical significance.

Non-standardised assessment of the desired outcome variable. Several studies reported the use of standardised testing, however, this was typically done to establish a diagnosis of ASD as opposed to measuring the treatment effects for the relevant intervention targets. Measurements of intervention targets were typically made by rating behaviour during observation. Inter- and/or intra-rater reliability was generally reported; however, 11 of 15 SSRD studies did not report sufficient blinding of outcome assessors to participant diagnosis or stage of intervention in order to eliminate possible rater bias.

Sample size and participant inclusion. Due to the nature of ASD and due to practical intervention considerations, most SSRD studies evaluating PRT involved small sample sizes ranging from two to eight subjects. Seven out of 15 SSRD studies included fewer than three subjects which limited the strength with which any conclusions may have been drawn. Participation of individuals in relevant research studies requires a diagnosis of ASD, proximity to the treatment centre, and the means and availability to undergo a specified intervention program. These criteria introduce a population pool constraint regarding the type of participants who may be considered for inclusion initially, as well as a self-selection bias regarding the type of participants who ultimately choose to participate in PRT intervention. Few

studies sufficiently addressed the degree, if any, to which study participants differed from the larger ASD population. Therefore, individual child variability, non-representative samples, and self-selection biases must all be considered possible confounding variables in these study outcomes.

Possible author bias. Most studies evaluating the effectiveness of PRT were done by those who first developed and popularised the approach, and the clinicians trained by or working with them. Although all articles included in this review are from peer-reviewed journals; five of the included articles were authored or co-authored by individuals who directly profit from the sale of PRT material (now a registered trademark).

Agreement with other studies or reviews. No other formal reviews of PRT, with which to compare this review's findings, had been performed at the time of this review's writing.

Overall completeness and possible bias of review process. The inclusion criteria of this review were broad; all studies claiming to use PRT and which directly measured child outcomes were included. However, this review did not include studies whose intervention conditions may have been similar but which did not refer to their intervention specifically as PRT, or a variation thereof, or used other interventions in addition to PRT. Furthermore, this review was not able to control for the studies' authors' interpretation and implementation of PRT guidelines and practices. Therefore, it is possible that certain non-standard implementations were included and that some PRT consistent interventions not labelled as such, were excluded.

Conclusions

Implications for practice. PRT has the potential for achieving functional gains in children with ASD. However, the existing evidence for PRT as an intervention is weak. Therefore, care must be taken by the clinician to determine the appropriateness of the client and target behaviour. Strong baseline data and on-going treatment data should be collected as highly individual variation may be expected.

Implications for research. More research on PRT intervention is necessary to positively determine its treatment effect. Currently, very few group design studies, and no randomized control trials, have been conducted on PRT. Any further contributions of this type would greatly strengthen the evidence for PRT. Nevertheless, the current body of research could be meaningfully improved by future SSRD studies adhering to strong methodology and by implementing uniform intervention procedures across multiple studies.

Specifically, new research should include replicating previous studies using larger sample sizes in order to strengthen the current findings. Researchers in the field should endeavour to use consistent vocabulary and definitions in order to help establish a cohesive body of research. Care should be taken to provide complete descriptions of participants and methodology, and to use standardised measures of behaviour assessment in order to reliably compare interventions and participants across studies. Finally, rigorous data collection and statistical analyses should be included to determine real treatment significance.

REFERENCES

Allied Health Classification Committee (NAHCC). (2007). *The indicator for intervention (IFI) coding manual*. Australia: The Australian Psychological Society Ltd..

Baker-Ericzen, M., Stahmer, A., & Burns, A. (2007). Child demographics associated with outcomes in a community-based pivotal response training program. *Journal of Positive Behaviour Interventions*, *9*(1), 52-60.

Coolican, J., Smith, I., & Bryson, S. (2010). Brief parent training in pivotal response treatment for preschoolers with autism. *Journal Of Child Psychology & Psychiatry*, *51*(12), 1321-1330.

Darrah, J., Hickman, R., O'Donnell, M., Vogtle, L., & Wiart, L. (2012). *American Academy for Cerebral Palsy and Developmental Medicine* (AACPDM) methodology to develop systematic reviews of treatment interventions (Revision 1.2). 2008.

Dodd, B. (2007). Evidence-based practice and speech-language pathology: Strengths, weaknesses, opportunities and threats. *Folia Phoniatrica et Logopaedica*, *59*, 118-129.

Harper, C., Symon, J., Frea, W. (2008). Recess is time-in: Using peers to improve social skills of children with autism. *Journal of Autism and Developmental Disorders, 38,* 815-826.

Koegel, L. K., Carter, C. M., & Koegel, R. L. (2003). Teaching children with autism self-initiations as a pivotal response. *Topics in Language Disorders*, *23*(2), 134-145.

Koegel, R., & Frea, W. (1993). Treatment of social behaviour in autism through the modification of pivotal social skills. *Journal of Applied behaviour Analysis*, *26*(3), 369-77.

Koegel, R. L., O'Dell, M. C., & Koegel, L. K. (1987). A natural language teaching paradign for nonverbal autistic children. *Journal of Autism and Developmental Disorders*, *17*(2), 187-200.

Koegel, R. L., Schreibman, L., Good, A. B., Cerniglia, L., Murphy, C., & Koegel, L. K. (1989). *How to teach pivotal behaviors to autistic children*. Santa Barbara: University of California.

Koegel, R. L., Symon, J. B., & Koegel, L. K. (2002). Parent education for families of children

with autism living in geographically distant areas. *Journal of Positive Behavior Interventions*, *4*(2), 88-103.

Koegel, R. L., Vernon, T. W., & Koegel, L. K. (2009). Improving social initiations in young children with autism using reinforcers with embedded social interactions. *Journal of Autism and Development Disorders*, *39*(9), 1240–1251.

Kuhn, L., Bodkin, A., Devlin, S., Doggett, A. (2008). Using pivotal response training with peers in special education to facilitate play in two children with autism. *Education and Training in Developmental Disabilities, 43*(1), 37-45.

Minjarez, M., Williams, S., Mercier, E., Hardan, A. (2011). Pivotal response group treatment program for parents of children with autism. *Journal of autism and developmental disorders, 41*(1), 92-101.

Nefdt, N., Koegel, R., Singer, G., & Gerber, M. (2010). The use of a self-directed learning program to provide introductory training in pivotal response treatment to parents of children with autism. *Journal of Positive Behavior Interventions*, *12*(1), 23-32.

Pierce, K. & Schreibman, L. (1995). Increasing complex social behaviours in children with autism: Effects of peer-implemented pivotal response training. *Journal of Applied Behaviour Analysis, 28*(3), 285-295.

Pierce, K. & Schreibman, L. (1997). Multiple peer use of pivotal response training to increase social behaviours of classmates with autism: Results from trained and untrained peers. *Journal of Applied Behaviour Analysis, 30*(1), 157-160.

Pierce, K., Schreibman, L. (1997). Using peer trainers to promote social behaviour in autism: are

they effective at enhancing multiple social modalities?. *Focus on Autism and Other Developmental Disabilities, 12*(4), 207-218.

Randolph, J., Stichter, J., Schmidt, C., O'Connor, K. (2011). Fidelity and effectiveness of PRT implemented by caregivers without college degrees. *Focus on Autism and Other Developmental Disabilities, 26(4*), 230-238.

Robinson, S. (2011). Teaching Paraprofessionals of Students With Autism to Implement Pivotal Response Treatment in Inclusive School Settings Using a Brief Video Feedback Training Package. *Focus On Autism & Other Developmental Disabilities*, *26*(2), 105-118.

Sherer, M. R., & Schreibman, L. S. (2005). Individual behavioural profiles and predictors of treatment effectiveness for children with autism. *Journal of Consulting & Clinical Psychology*, *73*(3), 525-538.

Smith, I. M., Koegel, R. L., Koegel, L. K., Openden, D. A., Fossum, K. L., & Bryson, S. E. (2010). Effectiveness of a novel community-based early intervention model for children with autistic spectrum disorder. *American Journal on Intellectual & Developmental Disabilities*, *115*(6), 504-523.

Stahmer, A. C. (1995). Teaching symbolic play skills to children with autism using pivotal response training. *Journal of Autism & Developmental Disorders*, *25*(2), 123-141.

Stahmer, A., Schreibman, L., Powell, N. (2006). Social validation of symbolic play training for children with autism. *Journal of Early and Intensive Behaviour Intervention, 3*(2), 196-210.

Thorp, D. M., Stahmer, A. C., & Schreibman, L. (1995). Effects of sociodramatic play training on children with autism. *Journal of Autism & Developmental Disorders*, *25*(3), 265-282.