Differential Observing Responses: A Systematic Review

A Capstone Project

presented to

the Faculty of the Graduate School

at the University of Missouri-Columbia

---

by

LANDON KIRLIN

Dr. Lorraine A. Becerra, Thesis Supervisor

MAY 2020

The undersigned, approved by the dean of the Graduate School, have examined the thesis entitled

DIFFERENTIAL OBSERVING RESPONSES: A SYSTEMATIC REVIEW

presented by Landon Kirlin,

a candidate for the degree of Master of Science, and hereby certify that, in their opinion, it is worthy of acceptance.

_____

Dr. Lorraine A. Becerra

_____

Dr. Casey J. Clay

_____

Dr. Kerri P. Nowell

TABLE OF CONTENTS

ABSTRACT……iv

Introduction……1

Methods……7

Results……10

DOR Procedures and Effectiveness……10

Target Skills……11

Participant Characteristics……12

Prerequisite Skills……13

Quality Assessment……14

Strength of Evidence……14

Discussion……15

Figures and Tables ……19

References……26

Appendix A: Coding Sheet……30

ABSTRACT

Some individuals with autism spectrum disorder (ASD) respond inaccurately on conditional discrimination tasks involving matching to sample procedures (MTS). Differential observing responses (DOR) have been effective at increasing discriminated responding of relevant features of stimuli on these tasks (Farber, Dickson, & Dube, 2017). However, there is conflicting evidence of successful conditional discrimination when this procedure has been removed and the longer-term implications of this procedure are unknown (Walpole, Roscoe, & Dube, 2007). The purpose of the current review was to evaluate previous research investigating the effects of DOR procedures in learners with ASD and other developmental disabilities. Fifteen articles were included, and results indicated expressive DOR procedures more commonly maintained accurate responding relative to receptive DORs. Furthermore, only five of the articles selected met high quality standards.

Keywords: differential observing response, autism spectrum disorder, intellectual disability, conditional discrimination

Introduction

Individuals diagnosed with autism spectrum disorder (ASD) can exhibit a range of communication, social, and restrictive behavioral characteristics (American Psychiatric Association [APA], 2013). These deficits may lead to difficulty in the acquisition of new skills (Farber, Dickson, & Dube, 2017). One topography of restrictive responding which some people with ASD demonstrate is stimulus overselectivity. Stimulus overselectivity refers to atypically limited observing behavior (i.e., orienting toward or looking at the teaching materials) with respect to range, breadth or number of stimuli, or stimulus features (Lovaas, Koegel, & Schreibman, 1979). For example, overselectivity may account for inaccuracy in matching to sample (MTS) tasks in educational settings. Inaccuracy in responding can involve selecting stimuli with similar features in common (Walpole, Roscoe, & Dube, 2007).

Overselectivity is a considerable barrier to demonstrating conditional discrimination, which facilitates stimulus control and skill acquisition. Stimulus control refers to how the rate, latency, duration, or amplitude of a response is altered in the presence of an antecedent stimulus (Cooper, Heron, & Heward, 2007). Conditional discrimination occurs when behavior is under the operant control of one stimulus when it is in the presence or context of another stimulus feature (Ploog, 2010). Conditional discrimination is a prerequisite skill that serves as a signal for reinforcement that manifests itself in a variety of situations. One example is demonstrated by early play skills which involve the conditional awareness that pressing a button on a toy results in a pleasant song being played. The presence of a light switch signals to a learner with proper conditional discrimination skills that this switch can illuminate a dark room (Cooper et al., 2007).

Overselectivity can be assessed using a number of tasks including MTS. MTS tasks often include stimuli printed on sheets of paper and tests if the participant demonstrates MTS accuracy within a field of different printed stimuli to a corresponding identical stimulus (Walpole et al., 2007). Walpole et al. (2007) assessed overselectivity in a learner who demonstrated MTS accuracy using printed words with no letters in common but demonstrated difficulty with MTS using three letter words with two letters in common.

Overselectivity remediation techniques identified in the literature include training with sufficient number of exemplars (Stokes & Baer, 1977), stimulus fading (McIlvane & Dube, 1992), conditional discrimination tasks (Schreibman, 1975), overtraining (Schreibman et al., 1977), and increasing observing behavior (Walpole et al., 2007). Training with a sufficient number of exemplars is demonstrated when the participant attends to a series of previously unattended stimuli. This can be developed by presenting these stimuli in isolation while reinforcing correct responses to them (Stokes & Baer, 1977). In contrast to a sufficient number of exemplars used, stimulus fading is a procedure where a relevant feature of a stimulus is presented in an exaggerated format so that a specific single stimulus dimension such as size and color is more visible to the learner and consistently faded over time (McIlvane & Dube 1992). Procedures which do not require the systematic removal of particular stimuli are conditional discrimination tasks. These involve reinforcement of a response when stimuli are in the presence or context of another stimulus (Catania, 1998). An example of a conditional discrimination task used in classrooms are MTS procedures. Overselectivity can interfere with accurate responding on MTS and other tasks within discrete trial training procedures as measured by inaccurate responding.

Overtraining procedures are described as multiple presentations of a single stimulus and requiring accurate discrimination of the stimulus across multiple sessions (Schreibman et al., 1977). This procedure has the potential to expose participants to extended opportunities to conditionally discriminate relevant features, however, a significant amount of time may be spent on overtraining procedures.

Other techniques that increase observing behavior involve procedures that direct the individual's attention towards relevant stimuli. Redirecting attention in this way increases stimulus control, such as in conditional discrimination tasks. An observing response is a behavior which produces discriminative stimuli correlated with differential reinforcement contingencies (Ploog, 2010). Overt observing behaviors can be verified by receptively selecting or expressively naming a stimulus during conditional discrimination tasks. Prerequisite skills for increasing observing behavior can be determined by assessment of conditional discrimination using MTS tasks within fields of stimuli of varying number and complexity (Kisamore, Karsten, Mann, & Conde, 2013). Doughty and Hopkins (2011) increased an observing response and demonstrated that accuracy increased in a delayed two sample MTS task. The observing response required two participants with ASD to "click" on a shape displayed on a computer screen either 10 times (FR 10) or once (FR 1) before progressing to the target MTS task.

One procedure that increases observing behavior is a differential observing response (DORs; Dube & McIlvane, 1999). DORs prompt the behavior of attending to and discriminating relevant stimulus features across an array of stimuli (Ploog, 2010). In contrast to observing responses, differential observing responses require a different response for each sample stimulus prior to the target response (Dube & McIlvane, 1999). An example of this procedure is

expressively naming a sample stimulus “red” before a receptive listening task requiring the participant to select a red color card among an array of different colored cards. Saying “red” functions as a DOR due to verifying observing and attending to relevant features of a target stimulus but is not the target response. The conceptual features of a DOR are intended to function as a within stimulus prompt which directs the participant’s observing behavior to target features previously unobserved, therefore increasing discrimination (Ploog, 2010).

A range of responses have been increased using a DOR response, including increased accurate responding to sample stimuli in MTS tasks (Dube & McIlvane, 1999; Walpole et al., 2007), increased independent responding to receptive listening tasks (Charlop, 1983; Fisher, Retzlaff, & Akers 2019;), and increased independent responding in intraverbal performance tasks (Kisamore et al., 2013). One advantage of a DOR procedure is the ability to modify the procedures to meet the needs of the learner. The DOR prompt can be an expressive response such as emitting an echoic or prompted via MTS which can be beneficial for non-vocal learners (Dube & McIlvane, 1999; Jones & Zarcone, 2014; Leung & Wu, 1997).

Researchers have used DORs to reduce stimulus overselectivity in participants with ASD (Farber et al., 2017) as well as with participants of neurotypical development (Kisamore et al., 2013). There is some evidence that has demonstrated DORs to be an effective short-term approach to reducing overselectivity. Although a Walpole and colleagues (2007) demonstrated increased response accuracy during an MTS procedure that maintained during a return to baseline, the generalizability of this result is questionable. Specifically, the limited number of generalization conditions, the use of only one participant, and no measure assessing whether the effect maintained over a sustained period of time generates many questions about the utility of

this procedure in applied settings. It is unclear if the effects of a DOR on increasing response accuracy to a number of different tasks maintain when the procedure is removed or require additional training to maintain accurate performance. Dube and McIlvane (1999) reported that responding returned to chance levels upon removal of the DOR and a return to baseline. More recently, Farber et al. (2017) included a titration procedure which systematically removed the DOR dependent on increases in the participant's correct performance. This demonstrated that the accuracy of the target response maintained on a delayed MTS task and the DOR requirement was successfully decreased based on past accurate performance. One limitation is that researchers did not examine whether this effect maintained over extended periods of time. Therefore, the long-term implications of DORs remain unclear.

Although DORs have been effective at decreasing overselectivity, it is unclear the range of skills acquired using this procedure. Similarly, the necessary prerequisite skills for individuals to successfully use the DOR procedure have not been described comprehensively in the literature. Finally, an evaluation of the extent to which the DOR procedures can be removed while maintaining the effects is still to be determined. Therefore, a review is needed to investigate the range of skills taught using a DOR, the prerequisite skills necessary to benefit from a DOR procedure, and the extent to which accurate responding maintains after the removal of a DOR. Hence, the purpose of this systematic review is to examine previous studies utilizing a DOR as an intervention method with children diagnosed with developmental disabilities. Furthermore, we assessed the previous limitations across the literature and specified recommendations for future research. Specifically, research questions included the following:

1. What target skills have been taught utilizing DOR procedures?

2. What type of expressive and receptive DOR procedures demonstrated the most effective increases in responding to the target skill?

3. What prerequisite skills do participants demonstrate prior to inclusion of DOR studies in order to benefit from this procedure?

4. To what extent will accurate responding maintain and generalize after removal of the DOR?

5. To what extent do the included studies meet the What Works Clearinghouse Standards 3.0 (2014) without reservations?

6. What recommendations can be made regarding future DOR research?

Methods

Search Procedure

To investigate the utility of previous DOR procedures, researchers conducted a systematic literature review which began first with conducting a search of academic electronic databases including: Science Direct, ERIC, Academic Search Premier, and PsychINFO. The search term combinations included: (a) “differential observing response” (b) “differential observing response” AND “autism spectrum disorder”, and (c) “differential observing response” AND disabilities”. These keywords were limited to the abstracts and titles of the articles. Furthermore, only publications that were printed in English and that were peer-reviewed were included in this search. Following this, an ancestral search was conducted which consisted of reviewing the reference section of each article located during the systematic keyword search. A forward search was then conducted using Google Scholar to locate additional articles that may have been published more recently. Lastly, a hand search of journals was then conducted to search for relevant articles that may have been missed in the search index. The hand search included relevant authors (a) William Dube, (b) Wayne Fisher and (c) Rachel Farber and were searched utilizing databases described above.

Inclusion Criteria

Researchers used the following criteria to select studies that were relevant to this literature review: (a) a differential observing response was utilized in the study as the independent variable, (b) the dependent variable was an overt response demonstrating skill acquisition, and (c) all participants were diagnosed with ASD or intellectual disability. No restrictions with respect to participant age were included in this study. All duplicate articles were

Commented [NKP1]: Participants? Ages, diagnoses? It’s mentioned in the coding sheet section and needs to be included here

excluded following screening of relevant studies. Figure 1 depicts the search process and the number of articles located during each step.

Coding Procedures

Coding sheet.

A coding sheet (Appendix A) was devised to assist in ensuring that appropriate studies were selected and was designed with the inclusion criteria shown above. Articles that did not meet eligibility criteria were not included in the review. The coding sheet assessed (a) participant characteristics, (b) dependent measurement procedures, (c) experimental design, (d) interobserver agreement, (e) procedural fidelity, (f) independent variable procedures, (g) limitations, and (h) strength of evidence as methodological quality that meets stands with or without reservations.

Quality of Evidence.

Eligible studies that met inclusion criteria were further assessed using the quality of the evidence as indicated by the methodological study design and demonstrated effects. Relevant participant characteristics, operational definitions of target responses and measurement procedures, experimental design characteristics, IOA, procedural fidelity, and limitations were all assessed across each included study. Following this, a quality assessment based on the guidelines stated by the What Works Clearinghouse (WWC) Procedures and Standards 3.0 (What Works Clearinghouse, 2014) was used to evaluate the quality of the experimental design. WWC standards provides a systematic method to analyze the methodological quality of studies to more effectively determine whether changes in the dependent variables were due to the independent variable being studied. WWC standards (What Works Clearinghouse, 2014)

Commented [NKP2]: Would be nice to include a brief statement about the rationale for using the WWC procedures and standards

describes assesses quality at three levels: Meets standards without reservations, meets with reservations, and not enough information. These levels were determined based on study design. Specifically the number of conditions, the data points across conditions, and the extent of the effects across conditions procedures.

Commented [NKP3]: Remove extra space after this paragraph

Inter-observer Agreement (IOA)

The primary author coded each of the articles included in this review. A second researcher was trained to independently code 33%, or five of the included fifteen articles, using the coding sheet described above (See Appendix A). Questions within each section of the coding sheet were collected for each study and categorized as one response. Point by point IOA was calculated by dividing the number of agreements across both coders by agreement plus disagreements and then multiplied by 100 to obtain a percentage. An agreement was defined as both observers selecting the identical response within the code sheet or writing the same description. Responses across five code sheets indicating 94% reliability (range, 88% - 100%), therefore additional training was not conducted.

Results

The results are depicted in Figure 1 with the initial search yielding a total of 45 articles. The primary author screened the title and abstract and removed duplicate articles to initially identify nine relevant articles. A forward search of the nine articles identified four additional articles for a total of 13 articles. The primary author obtained one additional article through an ancestral search of the reference list across each of the 13 articles. Following this a hand search consisting of relevant authors was conducted which revealed one additional article totaling 15 studies.

DOR Procedures and Effectiveness

Expressive DORs. Echoic responses taught as a DOR comprise 60% of the literature (Carp et al., 2015; Charlop et al., 1983; Fisher et al., 2019; Gutowski & Stromer, 2003; Kisamore et al., 2016; Leung & Wu, 1997; Tanji et al., 2012; Vedora & Barry, 2016; Vedora et al., 2017). These studies utilized a DOR in the form of prompted expressive labeling to increase response accuracy on intraverbal performance and auditory-visual discrimination target tasks (Carp, Peterson, Petursdottir, & Ingvarson, 2015; Charlop, 1983; Fisher et al. 2019; Gutowski, & Stromer, 2003; Kisamore et al., 2016; Leung & Woo, 1997; Tanji, Takahashi, & Noro, 2012; Vedora & Barry, 2016; Vedora, Barry, & Ward-Horner; 2017).

Commented [NKP4]: Should this data be included under the "expressive DOR" paragraph?

Commented [KL5R4]: Moved and modifed

Seventy-five percent of these articles report reaching mastery criteria while the treatment was in effect. Thirty-seven percent of articles featuring an echoic DOR component reported that high accuracy maintained following removal of the treatment (Fisher et al., 2019; Vedora et al., 2017, Tanji et al., 2012). This finding is significant due to repeated demonstrations that expressive DORs could maintain accurate responding after discontinuation in learners with ASD

or ID in the literature. How long these responses maintain were not assessed. All of these studies utilized echoics of the sample stimulus as a DOR before a receptive ID task. Reported mastery level responding was acquired more quickly in the DOR conditions when being compared to other prompt procedures (Carp et al., 2015; Fisher et al., 2019; Vedora et al., 2017, Tanji et al., 2012). Thirty-seven percent of expressive studies tested for generalization and reported increased accuracy in responding for at least two participants (Charlop 1983; Leung and Wu, 1997; Tanji et al., 2012).

Receptive DORs. All articles featuring non-verbal or receptive DORs report increased accuracy above 80% for at least one participant while the intervention is still in effect or systematically reduced via titration procedure (Dube & McIlvane, 1999; Farber et al., 2017; Fisher et al., 2007; Jones & Zarcone 2014; Moore et al., 2018; Vedora & Barry, 2016; Walpole et al., 2007), however, six of these studies (85%) report a return to baseline responding if procedures were removed (Dube & McIlvane, 1999). Farber and colleagues (2017) demonstrated that accurate responding can be maintained while requiring a DOR requirement which increased or decreased based on past performance, the only significant DOR study involving a titration procedure. As discussed in the introduction section, Walpole et al. (2007) published the only receptive DOR study demonstrating increased responding after removal. This article also included a generalization condition that was left unfinished and demonstrated maintenance was not tested further.

Target Skills

Teach targets are those for which overselectivity is demonstrated and for which a DOR procedure increases accuracy and acquisition. Table 1 illustrates the variety of target skill areas

identified in this review that have been demonstrated while a DOR is implemented during treatment resulting in increased accurate responding on MTS, receptive listening, and intraverbal performance. Six of the articles (40%) utilized MTS as a DOR intervention to increase accurate responding on receptive ID matching tasks (Dube & McIlvane, 1999; Farber et al., 2017; Fisher, Kodak, & Moore, 2007; Jones & Zarcone, 2014; Moore, Russo, Gilfeather, Whipple, & Stanford, 2018; Walpole et al., 2007). Three (33%) of these MTS studies used non-representative forms to ensure increased accuracy and attending skills were not confounded by previous history effects (Dube & McIlvane, 1999; Farber et al., 2017). The remaining studies utilized educationally relevant stimuli such a printed word sets (Walpole et al., 2007) and common objects such as shapes, colors, animals and famous people printed on laminated cards (Fisher, et al., 2007; Jones & Zarcone, 2014; Kisamore et al., 2016; Moore et al., 2018).

Participant Characteristics

Table 2 features the characteristics of participants within DOR studies. As many as 22 individuals participated in the DOR studies, with an average of four (*range:* 1-22). A variety of age groups also participated with an average of 10 years and median of 11 years (*range:* 3-53). Ten studies (66%) included participants with ASD (Carp et al., 2015; Charlop et al.,1983; Fisher et al., 2019; Gutowski & Stromer, 2003; Kisamore et al., 2016; Leung & Wu, 1997; Moore et al., 2018; Tanji et al., 2012; Vedora & Barry, 2016; Walpole et al., 2007), while individuals with an intellectual disability (ID) participated in two studies (13%) (Dube & McIlvane, 1999; Jones & Zarcone, 2014). The remaining 21% included a combination of participants diagnosed with ASD and/or ID (Farber et al., 2017; Fisher et al., 2007; Vedora et al., 2017).

Prerequisite Skills

Identification of appropriate prerequisite skills is necessary when deciding on appropriate intervention procedures (Cooper et al., 2007). Assessment of prerequisite skills were addressed within eight articles (53%), which included a brief description of the participant's exposure to past discrete trial training target skills and engagement in responses such as echoics, identity matching with pictures and one word mands as reported by the educator. Seven articles (47%) reported participants could accurately respond to generalized identity matching in some form but demonstrated difficulty when the task became more complex during assessment (Dube & McIlvane, 1999; Farber et al., 2017; Fisher et al., 2007; Fisher et al., 2019; Moore et al., 2018; Tanji et al., 2013; Walpole et al., 2007). These initial identity matching responses may serve as a behavioral cusp for increased attending and discrimination skills. Compliance is also a sufficient consideration when evaluating the appropriateness of DOR procedures. All 15 studies included participants who could follow one-step instructions and had some receptive listening repertoire based on caregiver or therapist report prior to inclusion of study as well as reviewing participant's early intervention records (Carp et al., 2015; Charlop et al., 1983; Dube & McIlvane, 1999: Farber et al., 2017; Fisher et al., 2007; Fisher et al., 2019; Gutowski & Stromer, 2003; Jones & Zarcone, 2014; Kisamore et al., 2016; Leung & Wu, 1997; Moore et al., 2018; Tanji et al. 2012; Vedora & Barry, 2016; Vedora et al., 2017; Walpole et al., 2007). The prerequisite skill most associated with increases in accurate responding with potential to maintain and generalize during DOR interventions were echoic behaviors. Eight studies (53%) found increases in accurate responding. (Carp et al., 2015; Charlop et al., 1983; Fisher et al.,

2019; Gutowski & Stromer, 2003; Kisamore et al., 2016; Leung & Wu, 1997; Tanji et al., 2012; Vedora et al., 2017).

Quality Assessment

Authors evaluated the quality of evidence within each article utilizing the What Works Clearinghouse Standards 3.0 (What Works Clearinghouse, 2014) that provides criteria for quality assessment based on research design, interobserver agreement, and procedural fidelity. WWC Standards revealed 33% of studies met standards without reservations (Fisher et al., 2019; Gutowski & Stromer, 2003; Jones & Zarcone, 2014; Kisamore et al., 2016; Moore et al., 2018). Twenty percent met standards with reservations (Vedora & Barry, 2016; Vedora et al., 2017; Walpole et al., 2007) and the remaining 47% did not meet standards due to insufficient reporting of interobserver agreement and procedural integrity while meeting all other design standard criteria including sufficient datapoints and phases needed to demonstrate an effect (Carp et al., 2015; Charlop et al., 1983; Dube & McIlvane, 1999; Farber et al., 2017; Fisher et al., 2007; Leung & Wu, 1997; Tanji et al; 2012).

Strength of Evidence

Further examination of articles using WWC criteria demonstrated that all fifteen studies presented an effect while the procedure was in place. Visual analysis revealed an increased change in level and trend compared to baseline. Removal of treatment demonstrated the short-lived nature of the effect.

Discussion

The purpose of this literature review was to evaluate available research on DORs and identified 15 studies. Two general DOR forms have been identified (i.e. receptive vs expressive) across target responses when individuals with ASD and/or ID demonstrate difficulty in receptive identity matching, intraverbal performance, and MTS. This review revealed the utility of a DOR procedure when learners demonstrate past inaccurate responding on skill acquisition targets and the DOR procedure be a viable alternative when past prompting procedures have not been effective. This review also revealed that nearly half (47%) of included studies failed to meet WWC standards on the basis of lack of treatment integrity and procedural fidelity measures. Recommendations are provided below for best practices in providing quality DOR research.

Expressive DORs appear to have a higher likelihood of maintaining following cessation of the intervention given the number of studies evaluating reversal procedures and follow up probes for maintenance (Fisher et al., 2019; Tanji et al., 2012; Vedora et al., 2017). This finding is confounded by more rigorous testing procedures employed during expressive DOR procedures compared to receptive DOR studies. One consideration for this observed effect is that expressive DORs elicit verbal behavior with respect to attending to environmental discriminative stimuli which readily signals reinforcement and can better generalize to natural environments and acquire stimulus control. Motor responses elicited via receptive MTS tasks may have limited generalizability to attending in natural environments in part due to the lack of socially mediated generalized conditioned reinforcement that may maintain increased attending. Learner sensitivity to different frequencies of reinforcement is also a significant consideration (Dube & McIlvane, 1999). While receptive DOR tasks may present a novel way to increase conditional

discrimination in non-verbal learners in the short term, further research is needed to determine whether accurate responding during receptive DOR procedures can maintain or generalize. Systematic fading or titration of stimuli within intervention procedures may be an effective way to demonstrate sustained accurate performance (Farber et al., 2017). Participant's past experience with stimuli is a factor that appears to be positively associated with maintained effects. Studies involving non-representative forms had marked decreases following removal while studies involving stimuli that learners have past history of exposure were more likely to maintain (Dube & McIlvane, 1999; Farber et al., 2017; Fisher et al., 2019; Tanji et al., 2012; Vedora et al., 2017).

This literature review demonstrates the utility of DORs to teaching a variety of target skills, including receptive identification and expressive labeling. Furthermore, the presentation of the DOR can represent the form of expressive or receptive actions thereby increasing the accessibility to all learner abilities. Importantly, the practical implications of this review described prerequisite skills that increases the efficacy of DOR implementation. Given the effectiveness of DOR procedures on increasing the fundamental skill of conditional discrimination and reducing overselectivity, this review provided a critical analysis to support educators who seek to use these specialized procedures. DORs have often been utilized when learners with ASD have demonstrated prompt dependence on previous target skills (Tanji et al., 2012; Fisher et al., 2007) as well as increased accurate responding on intraverbal performance tasks especially when other prompting procedures have been insufficient (Kisamore et al., 2016; Fisher et al, 2019). Therefore, DORs have been demonstrated to be an effective method to evoke increased accurate responding on tasks when other more traditional prompting procedures have

failed to produce an effect. Further research and discussion is needed to determine the environmental variables that possibly maintain this.

WWC standards revealed significant limitations within the current literature especially relating to treatment integrity and interobserver agreement. Researchers interested in increasing the utility of the DOR should include stringent criteria for implementing treatment integrity in future studies to control for bias and observer drift effects. Another limitation of this study involves the inability for researchers to obtain articles that were electronically or physically unavailable due to limited accessibility of databases associated with the University of Missouri. One seminal article evaluating expressive naming of objects on matching to sample tasks was unavailable to researchers due to this limitation (Constantine & Sidman, 1975). Other limitations involve the restrictive nature of search keywords included in this study that may have obscured other findings. It is possible that other older research articles may have utilized a DOR with the population of interest without using this term thus may not have been identified though additional forward or hand searches. A broader search of the academic literature outside of educational and behavioral analytic studies may also yield novel findings.

Despite the limitations of this review a number of different implications and future recommendations can be made when considering a DOR as a treatment for future topic of study. Future research investigating the utility of DOR procedures should focus on methods which increase generalization and maintenance of accurate responses. This research could involve DORs featuring augmented and assistive communication (AAC) devices or Picture Exchange communication systems to better generate stimulus control in nonverbal learners. Research on intraverbal performance tasks could increase the stimuli taught using a DOR such as

incorporating responding to features of expressive language such as prepositions and pronouns while comparing other prompting procedures utilizing components of alternating treatment designs. MTS tasks using three dimensional objects along with picture cards is another consideration which could facilitate accurate responding as well as generalization skills with nonverbal learners. For example, individuals who demonstrate inaccuracy in a receptive listening task using samples with multiple features in common such as a block and Lego could match the three-dimensional objects contingent upon incorrect responding as a DOR before requiring responding to a picture of objects. Consideration should also be taken regarding systematic fading of DORs as a prompt and may account for why DORs do not maintain increased responding after removal. MTS tasks using fading could increase the exposure the learner has to the DOR treatment while systematically decreasing the stimulus opacity and maintaining accurate responding. Additional titration procedures in which the DOR is systematically removed across time using ratios of DOR presentation relative to target presentation. Furthermore, opacity and the shade of stimuli color can be gradually removed within the DOR until the DOR procedure is no longer present, but the effects still remain.

Figures

Figure 1

Article selection process flow chart

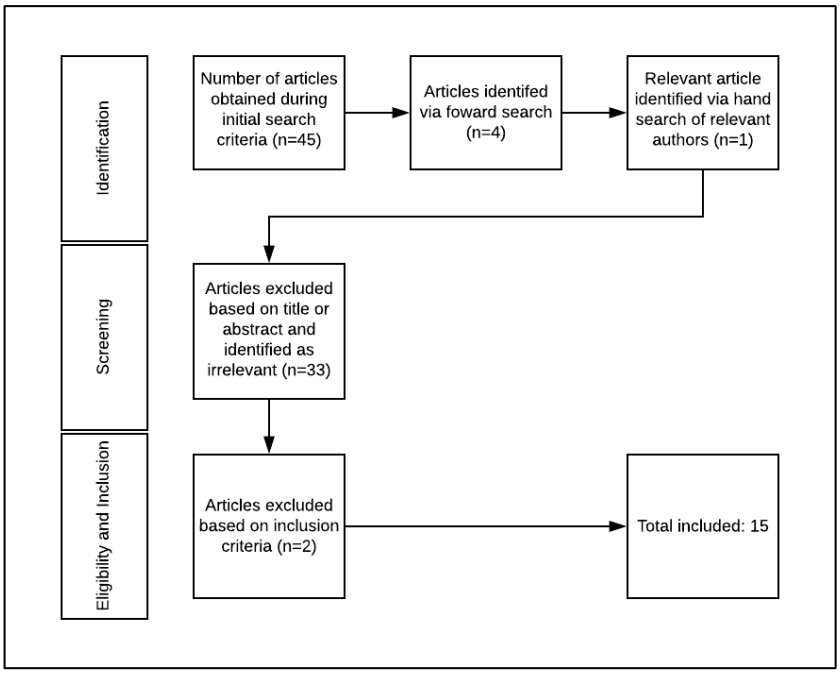


Figure 2

WWC Assessment

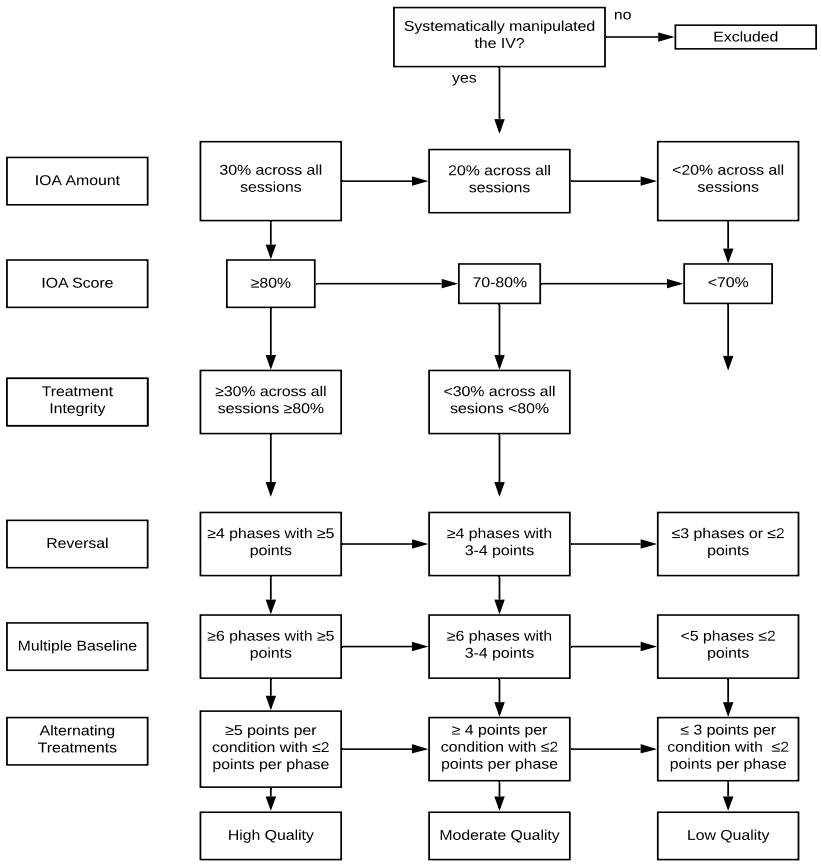
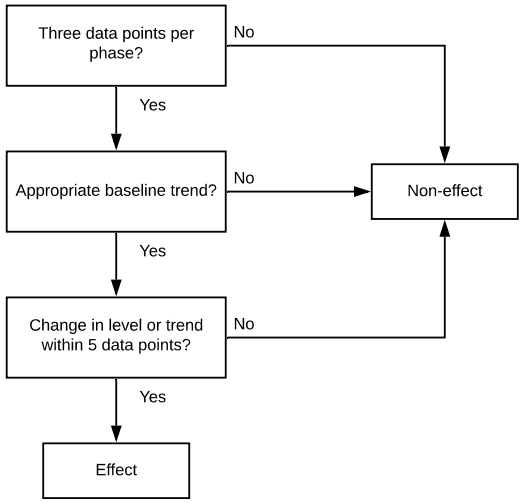


Figure 3

Assessment of effect flow chart



Tables

Table 1

Summary of DOR characteristics

Study	Design	Target Skill	DOR Type	Performance Criterion; % Met	Fidelity				
					G	M	SV	IOA	PF
Carp et al. (2015)	Alternating Treatments	Receptive ID	Echoic	87.5 % within 16 sessions; 0%	N	N	N	Y	N
Charlop (1983)	MBL+ Reversal	Intraverbal Performance	Echoic	90% within 20 sessions; 100%	Y	N	N	Y	N
Dube & McIlvane (1999)	Reversal	MTS	MTS	N/A	N	N	N	N	N
Farber et al. (2017)	Reversal	ID Matching	MTS	89% within 12 sessions; 57%	N	N	N	Y	Y
Fisher et al. (2007)	MBL	ID matching	MTS	80% within last 5 sessions; 100%	N	N	Y	Y	N
Fisher et al. (2019)	MBL	Receptive ID	Echoic	80% across 3 sessions; 100%	N	N	N	Y	Y
Gutowski & Stromer (2003)	Reversal	Receptive ID	Echoic	N/A	N	N	N	Y	Y
Jones & Zarcone (2014)	MBL	MTS, Receptive ID	MTS	N/A	N	N	N	Y	N
Kisamore et al. (2016)	MBL	Intraverbal performance	Echoic	80% across 5 sessions; 100%	N	N	N	Y	N

Note. MBL Multiple Baseline, *ASD* Autism Spectrum Disorder, *ID* Intellectual Disability, *G* Generalization, *M* Maintenance, *SV* Social Validity, *IOA* Interobserver Agreement, *PF* Procedural Fidelity/Treatment Integrity

Table 1 (Continued)

Study	Design	Target Skill	DOR Type	Performance Criterion; % Met	Fidelity				
					G	M	SV	IOA	PF
Leung & Wu (1997)	MBL	Receptive ID	Echoic	90% within 20 trials; 100%	Y	N	N	Y	N
Moore et al. (2018)	MBL+ Alternating Treatment	Receptive ID	MTS	80% across 3 sessions; 100%	N	N	N	Y	Y
Tanji et al. (2012)	Multiple Probe	Receptive ID	MTS+Echoic	3 sessions above 90%; 100%	Y	N	Y	Y	N
Vedora & Barry (2016)	Multiple Probe	Receptive ID	MTS	88% across 3 sessions; 100%	N	Y	N	Y	Y
Vedora et al. (2017)	Alternating Treatments	Receptive ID	Echoic	92% across 2 sessions; 100%	N	Y	N	Y	N
Walpole et al. (2007)	MBL	MTS	MTS	89% across 6 sessions; 100%	N	Y	N	Y	N

Note. *MBL* Multiple Baseline, *G* Generalization, *M* Maintenance, *SV* Social Validity, *IOA* Interobserver Agreement, *PF* Procedural Fidelity/Treatment Integrity

Table 2

Participant characteristics

Study	Characteristics				
	N	CA	Stimuli	Diagnosis	Setting
Carp et al. (2015)	2	3, 5	CO	ASD	School
Charlop et al. (1983)	5	5-11	CO	ASD	School
Dube & McIlvane (1999)	3	13-19	NR	ID	School
Farber et al. (2017)	22	8-21	NR	ASD+ID	Clinic
Fisher et al. (2007)	2	10, 12	Shapes, NR	ASD+ID	Clinic
Fisher et al. (2019)	4	3-5	CO	ASD	Clinic
Gutowski & Stromer (2003)	2	32, 53	CO, Animals	ASD	School, Residence
Jones & Zarcone (2014)	3	13-17	CO, Animals	ID	Hospital
Kisamore et al. (2016)	7	4-18	FP, CO, Colors	ASD	School, Home, Clinic
Leung & Wu (1997)	4	8-10	M	ASD	School
Moore et al. (2018)	2	11, 11	Colors	ASD	School
Tanji et al. (2012)	3	9-11	JA	ASD	School
Vedora & Barry (2016)	2	19, 15	Animals	ASD	School
Vedora et al. (2017)	6	15, 21	Flags	ASD+ID	School
Walpole et al. (2007)	1	16	English letters	ASD	School

Table 3

Quality assessment and strength of evidence

Study	WWC Standards	Effect/Non-effect
Carp et al. (2015)	Does not meet standards	Effect
Charlop et al. (1983)	Does not meet standards	Effect
Dube & McIlvane (1999)	Does not meet standards	Effect
Farber et al. (2017)	Does not meet standards	Effect
Fisher et al. (2007)	Does not meet standards	Effect
Fisher et al. (2019)	Meets w/o reservations	Effect
Gutowski & Stromer (2003)	Meets w/o reservations	Effect
Jones & Zarcone (2014)	Meets w/o reservations	Effect
Kisamore et al. (2016)	Meets w/o reservations	Effect
Leung & Wu (1997)	Does not meet standards	Effect
Moore et al. (2018)	Meets w/o reservations	Effect
Tanji et al. (2012)	Does not meet standards	Effect
Vedora & Barry (2016)	Meets with reservations	Effect
Vedora et al. (2017)	Meets with reservations	Effect
Walpole et al. (2007)	Meets with reservations	Effect

References

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

Catania A. C. (1998). *Learning*. 4th edition. Upper Saddle River, NJ: Prentice Hall;

Carp, C. L, Peterson, S. P, & Petursdottir, A. I. (2015). Preliminary evaluation of an echoic response requirement following errors during auditory-visual conditional discrimination training. *International Journal of Behavior Analysis & Autism Spectrum Disorders, 1*, 1-12.

Charlop, M. H., (1983). The effects of echolalia on acquisition and generalization of receptive labeling in children, *Journal of Applied Behavior Analysis, 16*(1), 111-126. https://doi.org/10.1901/jaba.1983.16-111

Cooper, J. O., Heron, T. E., & Heward, W. L. (2007). *Applied Behavior Analysis* (2nd edition). Pearson.

Constantine, B., & Sidman, M. (1975). Role of naming in delayed matching to sample. *American Journal of Mental Deficiency*, *79*(6), 680-689.

Doughty, A. H., & Hopkins, M., N. (2011). Reducing stimulus overselectivity through an increased observing response requirement. *Journal of Applied Behavior Analysis, 44(3),* 653-657. https://doi.org/10.1901/jaba.2011.44-653

Dube, W., & McIlvane, W. (1999). Reduction of stimulus overselectivity with nonverbal differential observing responses. *Journal of Applied Behavior Analysis, 32*(1), 25-33. https://doi.org/10.1901/jaba.1999.32-25

Farber, R. S., Dickson, C. A., & Dube W. V., (2017). Reducing overselective stimulus control with differential observing responses. *Journal of Applied Behavior Analysis, 50*(1), 87-105. https://doi.org/10.1002/jaba.363

Fisher, W. W., Kodak, T., & Moore, J. W. (2007). Embedding an identity matching task within a prompting hierarchy to facilitate acquisition of conditional discriminations in children with autism. *Journal of Applied Behavior Analysis, 40*(3), 489-499. https://doi.org/10.1901/jaba.2007.40-489

Fisher, W. W., Retzlaff, B. J., Akers, J. S., DeSouza A. A., Kaminski, A., J., & Machado, M. A. Establishing initial auditory-visual conditional discriminations and emergence of initial tacts in young children with autism spectrum disorder. *Journal of Applied Behavior Analysis, 52*(4), 1089-1106. https://doi.org/10.1002/jaba.586

Gutowski, S. J., & Stromer, R. S. (2003) Delayed matching to two picture samples by individuals with and without disabilities: analysis of role naming. *Journal of Applied Behavior Analysis, 36*(4), 487-505. https://doi.org/10.1901/jaba.2003.36-487

Jones, A. S., & Zarcone, J. R. (2014). Comparison of prompting strategies on two types of tasks with children diagnosed with autism spectrum disorders. *Behavior Analysis in Practice, 7*(2), 51-60. https://doi.org/10.1007/s40617-014-0010-7

Kisamore, A. N., Karsten, A. M., & Mann, C. C. (2016) Teaching multiply controlled intraverbals to children and adolescents with autism spectrum disorders. *Journal of Applied Behavior Analysis, 49*(4), 826-847. https://doi.org/10.1002/jaba.344

Kisamore, A. N., Karsten, A. M., Mann, C. C., & Conde, K. (2013). Effects of a differential observing response on intraverbal performance of preschool children: A preliminary investigation. *The Analysis of Verbal Behavior, 29,* 101-108. https://doi.org/10.1007/bf03393127

Leung, J. P., & Wu, K. I. (1997) Teaching receptive naming of chinese characters to children with autism by incorporating echolalia. *Journal of Applied Behavior Analysis, 30*(1), 30-59. https://doi.org/10.1901/jaba.1997.30-59

Lovaas, O. I., Koegel, R. L., & Schreibman, L. (1979). Stimulus overselectivity in autism: A review of research. *Psychological Bulletin, 86*(6), 1236-1254. https://doi.org/10.1037/0033-2909.86.6.1236

McIlvane, W. J., & Dube, W. V. (1992). Stimulus control shaping and stimulus control topographies. *The Behavior Analyst, 15*, 89–94. https://doi.org/10.1007/bf03392591

Moore, J. W., Russo, K., Gilfeather, A., Whipple, H. M., & Stanford, G. (2018) a comparison of error correction procedures within audio visual conditional discrimination training. *Journal of Behavioral Education, 27,* 532-543. https://doi.org/10.1007/s10864-018-9305-z

Ploog, B. (2010). Stimulus overselectivity four decades later: A review of the literature and its implications for current research in autism spectrum disorder. *Journal of Autism & Developmental Disorders, 40,* 1332-1349. https://doi.org/10.1007/s10803-010-0990-2

Schreibman, L. (1975). Effects of within stimulus and extra stimulus prompting on discrimination learning in autistic children. *Journal of Applied Behavior Analysis, 8,* 91-112. https://doi.org/10.1007/bf00915090

Schreibman, L., Koegel, R. L., & Craig, M. S. (1977). Reducing stimulus overselectivity in autistic children. *Journal of Abnormal Child Psychology*, *5*, 425–436. https://doi.org/10.1007/bf00915090

Sundberg, Mark L. (2008). *VB-MAPP Verbal Behavior Milestones Assessment and Placement Program: a language and social skills assessment program for children with autism or other developmental disabilities*. AVB Press.

Stokes, T. F., & Baer, D. M. (1977). An implicit technology of generalization. *Journal of Applied Behavior Analysis*, *10(2)*, 349–367. https://doi.org/10.1901/jaba.1977.10-349

Tanji, T., Takahashi, K., Noro, F. (2012) Teaching generalized reading and spelling to children with autism. *Research in Autism Spectrum Disorders, 7*, 276-287. https://doi.org/10.1016/j.rasd.2012.09.005

Vedora, J., & Barry, T. (2016). The use of picture prompts and prompt delay to teach receptive labeling. *Journal of Applied Behavior Analysis, 49*(4), 960-964. https://doi.org/10.1002/jaba.336

Vedora, J., Barry, T., & Ward-Horner, J. (2017). An evaluation of differential observing responses during receptive label training. *Behavior Analysis in Practice, 10*, 290-295. https://doi.org/10.1007/s40617-017-0188-6

Walpole, C., Roscoe, E, & Dube, W. (2007). Use of a differential observing response to expand restricted stimulus control. *Journal of Applied Behavior Analysis, 40*(4), 707-712. https://doi.org/10.1901/jaba.2007.707-712

Appendix A: Code Sheet

Coder Information	
Name:	
Date:	
Primary or Reliability:	

Report Characteristics	
Authors:	
Year:	
Title:	
Journal:	
Purpose statement	
Other:	

Eligibility Check		
Single subject design	Yes	No
Peer-reviewed	Yes	No
Published in English	Yes	No
Learners with ASD or a Developmental Disability	Yes	No
Used instructional format or mastery criteria as the IV	Yes	No
Measured impact on student's performance as the DV	Yes	No
Notes:		

**If you score "No" on any of these, do not complete the rest of the coding sheet

Participant Characteristics			
Age	Setting	Diagnosis	Notes
Notes:			

Operational definition of target response (DV)	Measurement procedures

Single Subject Design
What was the experimental design?
Did the study demonstrate experimental control of the IV on the DV? How strong was the evidence? Research question #1: Research question #2: Research question #3:

Inter observer Agreement			
Measured:	Yes No	% of Sessions	
Mean score		Range	
Other notes:			

Procedural Fidelity/Treatment Integrity			
Measured:	Yes No	% of Sessions	
Mean score		Range	
Other notes:			

IV Procedures	
Stimuli used (card pictures of shapes, worksheet, flags, etc.)	
Target style	Expressive Receptive Other: _____
Mastery Criteria	
Prerequisite skills required	
Notes:	

Limitations
Discuss the limitations the authors mentioned in the study:

What Works Clearinghouse Standards			
IV manipulated and measured	Yes No		
IOA in each phase for at least 20% of data points	Yes No		
Attempt to demonstrate effect over time and data points per phase (fill in sections that are applicable)			
	Meets	Meets w/ Reservations	Does not meet
Reversal/Withdrawal	4 phases with 5 points	4 phases with 3-4 points	3 phases or 2 points
Multiple Baseline	6 phases with 5 points	6 phases with 3-4 points	5 phases or 2 points
Alternating Treatment	5 points per condition with points per phase	4 points per condition with 2 points per phase	3 points per condition with 2 points per phase

Quality of effect: (circle one) Effect Non-effect
=