META-ANALYSIS OF VIDEO BASED MODELING INTERVENTIONS

FOR INDIVIDUALS WITH DISABILITIES: PROCEDURE, PARTICIPANT,

AND SKILL SPECIFICITY

A Dissertation

by

ROSE ANN KATHERINE MASON

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

May 2012

Major Subject: Educational Psychology

Meta-analysis of Video Based Modeling Interventions

for Individuals with Disabilities: Procedure, Participant, and Skill Specificity

META-ANALYSIS OF VIDEO BASED MODELING INTERVENTIONS

FOR INDIVIDUALS WITH DISABILITIES: PROCEDURE, PARTICIPANT,

AND SKILL SPECIFICITY

A Dissertation

by

ROSE ANN KATHERINE MASON

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Approved by:

Co-Chairs of Committee,	Jennifer B. Ganz
	Mack D. Burke
Committee Members,	Richard P. Parker
	Mandy J. Rispoli
	Kathryn McKenzie
Head of Department,	Victor Willson

May 2012

Major Subject: Educational Psychology

ABSTRACT

Meta-analysis of Video Based Modeling Interventions

for Individuals with Disabilities: Procedure, Participant, and Skill Specificity.

(May 2012)

Rose Ann Katherine Mason, B.S., Ambassador University; M.S., The University of

Texas at Tyler

Co-Chairs of Advisory Committee: Dr. Jennifer B. Ganz
                                 Dr. Mack D. Burke

The purpose of the present research was to address gaps in the video based

modeling (VBM) literature through the use of meta-analytic techniques to provide clarity

and specificity regarding the practical utility of VBM for participants with disabilities.

Two meta-analyses of published single-case VBM research were conducted.

Improvement rate difference, an effect size measure, was utilized to analyze the fifty-six

single-case studies. The purpose of study one was to determine if differential effects

occurred based on the type of model utilized and variations in procedural

implementation. In addition, the quality of research was evaluated. The purpose of study

two was to determine if participant characteristics, intervention components by

participant characteristics, and targeted outcome moderated the effectiveness of video

modeling with other as model (VMO). Results of Study One indicated moderate to

strong effects for both VMO and video self-modeling, however, when further

disaggregated based on type of model utilized, VMO with adult as model demonstrated

statistically significant superiority in terms of outcome effects. Results also indicated VBM with reinforcement demonstrated greater effects than when delivered alone or as part of a package. Additionally, the evaluation of quality of research indicated a tendency of the previously published VBM research not to evaluate treatment integrity. Study Two found that age and diagnosis moderate the effectiveness of VMO, although strong effects were found across levels for both moderators. VMO was found to be more effective for elementary age participants and participants with autism spectrum disorders. Additionally, VMO with reinforcement demonstrated statistically significant stronger effects for participants with ASD than when it is delivered alone or as part of a package. However, VMO delivered as part of a package was more potent for participants with developmental disabilities. Considering targeted outcomes, the results indicated strong effects across skill areas, however, VMO was found to be most impactful when utilized to improve play skills versus other measured skills. Implications related to the practical application of VBM for individuals with disabilities particularly in regards to treatment decision making were discussed. Additionally, implications for future research were addressed.

DEDICATION

To my precious family – Benjamin, Reese, Sloane, and Cohen. I have been frequently asked, "How do you do this with a family?" Cohen, you help keep me quick on my toes. Sloane, you remind me to be patient and experience life with passion. Reese, you never let me forget that there is always more to be known. Benjamin, you are my strength and sense of humor. My answer – "I couldn't do it without them!"

Also, to my dad – you made me believe it was possible to fulfill your wish to have a "doctor" in the family. I am always grateful for your love, support, and smiling eyes!

ACKNOWLEDGEMENTS

This would not be possible without the support of my committee co-chairs, Dr. Ganz and Dr. Burke. Both of you have spent countless hours, not only helping me complete this project, but guiding my program and fostering my development. Dr. Ganz, you have stretched me beyond my comfort but always been a voice of reason. Dr. Burke, your unique perspective has guided me to find the unanswered questions.

Also, a special thanks to my committee members, Dr. Parker, Dr. Rispoli, and Dr. McKenzie, for their guidance and support throughout my program and this research. Dr. Parker, your patience with my endless statistical questions is greatly appreciated. Dr. Rispoli, thanks for finding the gaps and adding the finishing touches. Dr. McKenzie, despite your busy schedule you managed to fit me in and provide a fresh perspective. Additionally, I acknowledge the Department of Educational Psychology staff, particularly Kristie Stramaski. Without her organization, reminders, and encouragement this would likely not have happened – at least not on time.

Thanks also go to my cohort members - Eric, Melissa, Leigh Ann, Stacey, Amy, and Jaime. This doctoral program has been a group project in many ways. Without the shared laughter, tears, and stats tutoring this would have been a much longer, more tedious, and less enjoyable journey.

As always, I save the best for last. A colossal thanks to my husband (i.e., my cheerleader, comedian, love of my life) who encouraged me to begin this journey in the

first place. Thanks to my beautiful children for giving me no other choice but to finish.

And thanks also to my dear parents, brothers, and sister.

TABLE OF CONTENTS

Page

LIST OF FIGURES

Page

LIST OF TABLES

CHAPTER I

INTRODUCTION

Identifying and implementing appropriate intervention for individuals with disabilities can be a formidable task given the heterogeneity within and between categories of disabilities (Odom et al., 2005). Further, the need for modeling multiple exemplars and repetition of instruction is often rigorous and demands generous allocations of resources, particularly teacher time (McLeskey & Waldron, 2011). Modeling is one such intervention that, although effective for enhancing skill acquisition for individuals with disabilities, requires a significant amount of time commitment from practitioners (Biedermann & Freedman, 2007).

Modeling stems from social learning theory which posits that individuals are able to learn by observing others and then mimicking the observed behavior (Bandura, 1969). From a behavioral perspective, this occurs when the observer indirectly experiences the consequences of the observed behavior and, as a result, learns to mimic that behavior (Crain, 1992) without direct application of consequences and across settings (Bellini & Akullian, 2007). Bandura's theory utilizes a "passive" approach of modeling that does not involve the engagement of another individual to facilitate accurate implementation, which Biedermann and Freedman (2007) suggest is a more effective means of learning.

This dissertation follows the style of *Exceptional Children*.

Modeling is frequently incorporated into instruction, (Biederman & Freedman, 2007). However, the limitations of live, or in vivo, modeling is constricted in effectiveness and efficiency, particularly when implemented with students with disabilities, as it requires the model to conduct the demonstration repeatedly, does not allow for previewing and correction of model demonstrations, and limits the flexibility of the procedure (Ayres, Maguire, & McClimon, 2009; McCoy & Hermansen, 2007).

Video-based modeling (VBM) is an alternative means for incorporating modeling into instructional strategies in a more efficient and feasible manner, while avoiding many of the challenges frequently associated with in vivo modeling (Biederman & Freedman, 2007; McCoy & Hermansen, 2007). Unlike in vivo modeling, the technology of recording allows for a one time creation of one or multiple exemplars of a targeted skill with the ability to then edit the video to ensure the demonstration is correct and precise (Corbett & Abdullah, 2005). Following the creation of the video model, the skill can then be modeled repeatedly to one or more individuals with or without the presence of an instructor (Shukla-Mehta, Miller, & Callahan, 2010). VBM is an efficient and effective strategy that is both time and labor saving.

A proliferation of educational research has recently identified VBM as an effective intervention across disability categories and skills (Bellini & Akullian, 2007; Nikopoulos & Nikopoulos-Smyrni, 2008; Rayner et al., 2009). VBM appears effective for participants with several different types of disabilities (Mechling, 2005), including intellectual disabilities (Cannella-Malone et al., 2006; Norman, Collins, & Schuster, 2001; Sigafoos et al., 2005), speech delays, developmental delays (Hitchcock, Prater, &

Dowrick, 2004), spina bifida, cerebral palsy, ADHD, and emotional-behavioral disorders (Hitchcock, Dowrick, and Prater, 2003). Additionally, VBM has been identified as an evidence-based educational practice for individuals with autism spectrum disorders (ASD: Bellini et al., 2007; Bellini & McConnell, 2010; Shukla-Mehta et al., 2010; Wang & Spillane, 2009).

In addition to implementation across participants with diverse disabilities, VBM has been implemented to address a wide variety of skills. VBM has been utilized for skills including social initiation (Deitchman, Reeve, Reeve, & Progar, 2010; Scattone, 2008), play skills (Kroeger, Schultz, & Newsom, 2007; Palechka & MacDonald, 2010), self-help skills (Horn et al., 2008; Lasater & Brady, 1995; Norman et al., 2001; Ruth, Dahman, Young, Cherry, & Davis, 2003; Sigafoos et al., 2007), compliance (Conyers et al., 2004), and social skills (Charlop, Dennis, Carpenter, & Greenberg, 2010; Tetreault & Lerman, 2010). Additionally VBM has been effectively implemented to enhance vocational (Allen, Wallace, Greene, Bowen, & Burke, 2010) and academic skills (Hitchcock et al., 2004). VBM has also been utilized to increase appropriate classroom behavior (Lonnecker, 1994) and encourage communication in a participant with selective mutism (Pigott & Gonzales, 1987).

VBM can be an educational accessory that allows the instructor to produce a precise enactment of the targeted skill that can be modeled to a number of students at different times and locations (Biederman & Freedman, 2007; Corbett & Abdullah, 2005); however the evidence base lacks implementation clarity, inhibiting the transfer of research to practice. "Video modeling" is an umbrella term used to describe a wide variety of

VBM interventions with and without other components, is employed with participants with a variety of disabilities, and is used to enhance a range of skills (Bellini & Akullian, 2007; Shukla-Mehta et al., 2010). It is plausible that numerous moderators impact the effectiveness of VBM yet sufficient assimilation and statistical analysis of the VBM research has not occurred to identify the variables that facilitate or hinder the degree to which socially valid changes occur. Despite the promising impact of VBM, many gaps regarding contextual specificity, such as participant characteristics and targeted outcomes, and procedural implementation remain (Shukla-Mehta et al., 2010).

Variations in Procedural Implementation

Many variations in procedural implementation of VBM exist. First, the type of model utilized varies in the literature. Video modeling with other as model (VMO) involves another person, either an adult or peer, modeling the target skill (Bellini & Akullian, 2007; Shukla-Mehta et al., 2010). Video self-modeling (VSM) on the other hand includes the person of interest accurately performing the targeted skill (Dowrick, 1999). VSM can be produced utilizing the feedforward method, which involves recording the participant accurately producing the targeted skill with assistance such as prompts and error correction (Dowrick, Kim-Rupnow, & Power, 2006; Hitchcock et al., 2003). The alternative for VSM is positive self-review. Positive self-review involves recording the individual for long periods of time until the video footage contains accurate implementation of each aspect of the targeted skill (Hitchcock et al., 2004). Both types

require extensive editing prior to viewing, removing any extraneous footage, inaccurate exemplars, and assistance from others.

In addition to different types of model utilized, other variations in procedural implementation exist. First, the video can be produced from a third-person perspective, in which the entire scene is visible, or from a first-person perspective, referred to as point-of-view modeling. In a point-of-view model only the model's view is displayed (Hines & Wolery, 2006; Tetreault & Lerman, 2010). Additionally, VBM has been implemented either alone or as part of a treatment package that includes other interventions such as prompting, reinforcement, performance feedback, or error correction procedures.

Only a few studies have investigated differential effectiveness of VBM based on type of model utilized, with mixed results. Cihak and Schrader (2008) implemented an alternative treatment design to determine which was more effective in promoting acquisition of independent living skills for four adolescents with ASD -- VMO with an unknown adult model or VSM. Three of the participants were able to perform each step of the VSM task independently before they were able to perform the VMO task independently, whereas the other participant did not demonstrate a difference between the two model types. Yet, in another study comparing the effect of model type in the acquisition of independent living skills for three adolescents with developmental disabilities, targeted skills were acquired more quickly when VMO was implemented for two of the participants (Van Laarhoven, Zurita, Johnson, Grider, & Grider, 2009). Both studies utilized least-to-most prompting to correct for errors in skill performance. In addition to analysis of results based on type of model utilized, Van Laarhoven et al.

(2009) also analyzed time investment for creation of each type of video and concluded that production of the VSM took significantly longer than required for the VMO video. Given efficiency and resource-use considerations, clarification regarding the presence or absence of differential effects based on the type of VBM implemented have practical implications. If there are only minimal or no differences in effectiveness between video self-modeling and VMO, VMO would be the more socially valid intervention in terms of resource allocation (Bellini & Akullian, 2007).

These variable results tentatively indicate that model type and inclusion of additional components impact the magnitude of change that occurs. Two of the key components for identification as an evidence-based practice are a clearly defined intervention protocol and a desired change in the dependent variable that can be attributed to the precise implementation of the intervention (Horner et al, 2005). However, without a clearly defined protocol including controlling for additional variables, the causal relationship between VBM and the targeted skills cannot be established. Analysis of VBM with clear delineation regarding the effectiveness of each type of VBM, as well as the impact of the inclusion or exclusion of additional intervention components, does not currently exist, preventing the establishment of replication across studies and the development of unambiguous implementation protocols (Shukla-Mehta et al., 2010).

Variation of Participant Characteristics

In addition to providing information regarding implementation protocol, precise descriptions of participant traits for which the intervention yielded the greatest response

are necessary to practically convey research results (Horner et al., 2005; McDonald et al., 2006; Odom, 2009). There is a dearth of information regarding age-specific best practices, as well as those strategies that might yield differential effects when comorbid diagnoses are present (Ganz et al., 2011). The differential effects of VBM based on particular participant characteristics, such as age and primary disability have not previously been analyzed.

VBM has been employed with participants ranging from toddlers to adults and with a wide variety of disabilities (Bellini and Akullian, 2007; McCoy & Hermansen, 2007; Shukla-Mehta et al., 2010). For instance, Sherer et al. (2001) compared the use of VMO and VSM to increase social-communication skills for 5 boys, ages 4-11, with ASD. Both model types resulted in increased skills for all participants except for the youngest one, for which neither model was effective in improving the targeted outcome. Further, Hepting and Goldstein (1996) experienced minimal improvement in socio-communicative skills with the implementation of VSM for a 4 year old male with an intellectual disability. VSM has also been successful at improving adaptive behavior for 4 males ages 10-13 with EBD (Kehle, Clark, Jensen, & Wampold, 1986) however, when VSM was implemented for 2 males ages 9 and 7 with learning disabilities, mixed results were obtained across participants (Lonnecker, Brady, McPherson, & Hawkins, 1994).

The inconsistent results across participants (Rayner et al., 2009) suggest that participant characteristics need to be further explored as potential moderators. Additionally, clarification regarding what implementation variables moderate effectiveness for which participant characteristics is warranted as it could lend further

guidance to practitioners in regards to choosing the appropriate intervention (Odom et al., 2005). However, this area has not yet been addressed in the VBM research.

Variability Based on Targeted Outcomes

In addition to clarifying for whom an intervention is effective and the procedural requirements for that intervention, distinguishing for which targeted skills the practice can be expected to have the desired impact is also necessary (Horner et al., 2005). To assist with practical implementation, information regarding the efficacy of VBM for particular skill sets is necessary. As demonstrated previously, VBM has been implemented to teach a broad assortment of skills yet the type of VBM that is most effective for which skills with which participants continues to be ambiguous.

Variable results on targeted outcomes are evident when reviewing the current VBM evidence base. For instance, implementation of VBM resulted in increased use of pretend play for a 4 and 7 year old with ASD (MacDonald, Clark, Garrigan, & Vangala, 2005). Additionally, acquisition of perspective taking occurred for 3 elementary aged participants with ASD with the implementation of VBM and reinforcement (LeBlanc et al., 2003). However, when VBM was introduced in an effort to teach toilet training to 3 elementary age participants with ASD, results were inconsistent across participants (Keen, Brannigan, and Cuskelly, 2007). Again, the variation in effects on targeted outcomes across studies implies VBM may be more effective for certain outcomes than others. Research that provides clearer guidance on the type of skill for which VBM would be an appropriate intervention is needed.

Purpose and Research Questions

The lack of statistical analysis quantifying the magnitude of change that allows one to make comparison between studies is a major limitation of the current VBM literature base. Definitive statements regarding the differential effectiveness of VBM across implementation variables, participant characteristics, and targeted outcomes have not been established. When such quantities of research exist, employing methical procedures to aggregate the individual studies is possible and necessary to derive specific and practical information (Scruggs & Mastropieri, 1998) regarding conditions under which the treatment protocol will provide the most beneficial results in real-world settings. Evaluating single-case studies utilizing the same metric allows one to combine the results across studies with similar features (e.g., all studies using VMO alone) and compare the effect to studies with different features to provide further clarification regarding participant and contextual factors that will and will not yield socially significant results (Scruggs & Matropieri, 1998).

Previous meta-analyses of VBM have been conducted (Bellini & Akullian, 2007; Baker, Lang, & O'Reilly, 2009); however, the authors calculated PND for the studies reviewed. PND has been questioned regarding its statistical soundness (Parker, Hagan-Burke, & Vannest, 2007). Further, the inability to calculate confidence intervals with PND inhibits the ability to meaningfully compare results across studies and participants based on overlap (Kavale, 2010; Parker, Vannest, & Brown, 2009; Vannest, Davis, Davis, Mason, & Burke, 2010) and also limits one of the benefits of meta-analysis, the ability to assess for variability based on study features (Kavale, 2010; Scruggs & Mastropieri,

1998). Newer statistical indices, such as improvement rate difference (IRD: Parker, Vannest, and Brown, 2009), enable the calculation of confidence intervals and analysis of potential moderators. This dissertation addresses gaps in the VBM literature, utilizing IRD to provide clarity and specificity regarding the practical utility of VBM.

The first study, Chapter 2, focuses on assessing differential effects that occur based on implementation factors such as the type of model implemented and procedural implementation for VBM with participants with disabilities via a meta-analysis of single-case research. Differential effects based on the type of VBM implemented, the model used (i.e., self, peer, or adult), and whether or not VBM was delivered alone, with reinforcement, or as part of an intervention package are evaluated. Additionally, the quality of research is evaluated.

Given the ease of delivery and more efficient nature of VMO, the second study, Chapter 3, focuses solely on the efficacy of VMO. VMO single-case research was evaluated utilizing meta-analytical techniques to determine differential effects that are related to participant characteristics and targeted outcomes. Specific participant characteristics that are evaluated include age, gender, and diagnostic category. Additionally, differential effects that occur when implementation variables are considered in conjunction with participant diagnosis are analyzed. Further, the study analyzes targeted outcomes as a potential moderator for the magnitude of change that occurs with the implementation of VMO.

CHAPTER II

VIDEO-BASED MODELING: DIFFERENTIAL EFFECTS DUE TO

TREATMENT PROTOCOL

Educational reform and legislation such as No Child Left Behind (NCLB; (U.S. Department of Education, Office of the Deputy Secretary, 2001; U.S. Department of Education, Office of the Deputy Secretary, 2001) and revisions to the Individuals with Disabilities Education Act (Individuals with Disabilities Education Improvement Act, 2004) obligate schools to utilize research-based teaching practices. Given the emphasis on isolating evidence based practices for individuals with disabilities, the literature base has been proliferated with seemingly efficacious and promising interventions (Odom, 2009). Video-based modeling (VBM) has received much attention as a practice with considerable potential for enhancing multiple skills for individuals with disabilities (Rayner, Denholm, & Sigafoos, 2009; Shukla-Mehta, Miller, & Callahan, 2010). However, questions remain regarding the impact of variations in procedural implementation on the efficacy of the intervention.

VBM induces simulation of observed skills or behaviors by exposing the target individual to a model correctly demonstrating the target skill or behavior via a video-recording (Delano, 2007; Hitchcock, Dowrick, & Prater, 2003). Theoretically, VBM influences skill acquisition via the observational learning and imitative components of social learning theory (Delano, 2007; Rayner et al., 2009).   Social learning theory posits that a reciprocal process occurs involving a constant interplay between cognitions, behavioral responses and the environment (Shukla-Mehta, Miller, & Callahan, 2010). By

observing the behaviors of others, individuals are able to learn new skills through vicarious observational learning (Bandura, 1969). From an applied behavior analysis standpoint, the behavior of the model serves as an antecedent, as the model's actions prompt the observer to mimic those actions (Nikopoulos & Nikopoulos-Smyrni, 2008).

The effectiveness of video modeling is dependent on an observer's ability to understand the information to the exclusion of other stimuli, cognitively store the information, and then reproduce the observed skill (Nikopoulos & Nikopoulos-Smyrni, 2008). Furthermore, modeling is thought to be most effective when the model has qualities comparable to the observers (Bandura, 1969) and when the observer regards the outcome as desirable (Corbett & Abdullah, 2005). VBM provides a feasible and efficient means for capitalizing on the educational benefits of imitation and modeling as it allows for unlimited exposure to a variety of examples (Ayres, Maguire, & McClimon, 2009). Additionally, VBM is an educational tool that allows for a precise enactment of the targeted skill that can be modeled to a number of students at numerous times and locations (Biederman & Freedman, 2007; Corbett & Abdullah, 2005). Even if empirical evidence indicated statistically identical outcomes between in-vivo and video-based modeling, VBM would be more valuable due to the additional advantages related to time and cost, as well as continuity across settings (Biederman & Freedman, 2007; Delano, 2007; Mechling, 2005; Rayner et al., 2009).

VBM can be categorized based on the type of model utilized. The categories include video modeling with other as model (VMO), video self-modeling (VSM), and point-of-view modeling (Bellini & Akullian, 2007). Point-of-view is filmed from the perspective of the viewer and utilizes either self or another person as the model (Shukla-Mehta et al., 2010) and as the observer does not actually see the model, this format is not included in the current analysis. Additionally, VBM interventions have been implemented either alone, as a primary component of an intervention, or as part of an intervention package (Shukla-Mehta et al., 2010).

VMO involves the individual watching a video of another adult or peer, either known or unknown, demonstrating the desired skill with the expectation that the behavior will be imitated (Bellini & Akullian, 2007; Shukla-Mehta et al., 2010). VSM involves the target individual performing as the model for the creation of the video (Dowrick, 1999; Shukla-Mehta et al., 2010). VSM is comprised of two distinct subcategories including positive self-review and feedforward (Dowrick, 1999; McCoy & Hermansen, 2007). Feedforward involves video-recording the individual while he or she is being assisted in performing a skill to the best of his or her ability, with the implementation of additional interventions including coaching, instruction, and reinforcement, which ensure that each step or component of the skill is correctly demonstrated by the target individual, with or without prompts or cues (Dowrick, Kim-Rupnow, & Power, 2006; .Dowrick, 1999; Hitchcock et al., 2003). The individual then views an edited video which has had the assistance and incorrect performances removed. Positive self-review, on the other hand, involves recording the individual engaged in natural activities and then editing out

examples of undesired behavior or inaccurate performance (Dowrick et al., 2006; Hitchcock, Prater, & Dowrick, 2004). The final video model is then a sample of the individual engaged in only positive exemplars of the targeted outcome (Dowrick, 1999; Hitchcock et al., 2003).

Previous literature reviews (Ayres & Langone, 2005; Delano, 2007; Hitchcock et al., 2003; McCoy & Hermansen, 2007; Shukla-Mehta et al., 2010) and meta-analyses (Baker, Lang, & O'Reilly, 2009; Bellini & Akullian, 2007) have examined the use of VBM as an intervention for individuals with disabilities. The qualitative literature reviews together cover a span from 1978 to 2008. However, the qualitative reviews (Ayres & Langone, 2005; Delano, 2007; Hitchcock et al., 2003; McCoy & Hermansen, 2007; Shukla-Mehta et al., 2010) lack careful and explicit descriptions of the independent variable. Consideration as an evidence-based intervention that is transferable into practice requires explicit description of all components of the independent variable in order to promote replication both in future research and applied settings (Horner et al., 2005). The published reviews for VBM fail to differentiate variations in procedural implementation.

Of the three literature reviews that included studies utilizing all three types of VBM (Ayres & Langone, 2005; McCoy & Hermansen, 2007; Shukla-Mehta et al., 2010), only McCoy and Hermansen (2007) and Shukla-Mehta et al (2010) summarized the studies based on the type of VBM utilized. Three reviews limited their analyses to VMO and VSM (Baker et al., 2009; Bellini & Akullian, 2007; Delano, 2007); however, only Bellini & Akullian (2007) considered VMO and VSM as separate interventions. Of the reviews (Bellini & Akullian, 2007; Hitchcock, 2003; McCoy & Hermansen, 2007; and

Shukla-Mehta et al., 2010) specifically addressing VSM as a separate intervention, none identified whether or not the VSM intervention utilized positive self-review or feedforward. Additionally, of the studies that examined the use of VMO, only one (McCoy & Hermansen, 2007) sorted the studies based on the type of model utilized (i.e., adult or peer). Without an examination of the variations in procedural implementation, identification of differential effects related to intervention variables is not possible.

Further, only two reviews have examined VBM delivered alone or as part of an intervention package (Delano, 2007; Shukla-Mehta et al., 2010). Delano's review of VBM interventions that were not part of a treatment package indicated VBM to be effective for participants with ASD across a range of targeted skills. However, the study did not consider inclusion of prompting and reinforcement as additional intervention components (Delano, 2007). Shukla-Mehta et al. (2010), also reviewing VBM interventions for participants with ASD, disaggregated the studies based on procedural implementation, considering reinforcement and prompting as additional component. Based on their qualitative review, Shukla-Mehta et al. (2010) concluded that the type of VBM did not impact the effectiveness, however inclusion of additional components appeared to enhance results. Analysis of VBM with clear delineation between each type of VBM, and the variables for each type, as well as differentiation regarding effects based on whether the intervention was delivered alone or as part of a package continues to be a gap in the research.

In addition to the qualitative literature reviews, two meta-analyses (Baker et al, 2009; Bellini & Akullian, 2007) have evaluated the impact of VBM on targeted skills.

Both meta-analyses utilized mean percent of non-overlapping data (PND) as the primary indicator of effectiveness across studies. Baker et al. (2009) conducted a meta-analysis on use of VBM with participants with emotional disabilities as a primary diagnosis. Baker et al. (2009) separated the studies based on targeted skill categories regardless of the type of VBM implemented. The reported PND for each skill set addressed (increasing peer interaction, increasing appropriate interaction, increasing on-task behavior, and decreasing inappropriate behavior) had a range of 0-100% across all skills with mean PNDs across the studies ranging from 14-100% (Baker et al, 2009).

Bellini and Akullian (2007) also utilized PND analyses to assess the impact of VMO and VSM,, for increasing a variety of skills for individuals with ASD. Information about individual participant outcomes was not noted and the range of the reported mean PNDs, 29-100%, indicated variable results. Utilizing Kruskall-Wallis, the Bellini and Akullian (2007) indicated no statistically significant differences in results between VMO and VSM, types of targeted skills, and age groups. In addition, Bellini and Akullian found many of the studies lacked measures of treatment fidelity and social validity. Furthermore, Bellini and Akullian (2007) acknowledged the grouping of VBM with other interventions as a limitation in terms of evaluating the benefits of VBM alone. The variation in magnitude of change was not explained in either meta-analysis, particularly in regards to differential results based on the type of VBM utilized or variations in delivery. This is important as the range of results seems to indicate implementation factors may moderate the effectiveness of the intervention.

The previous qualitative literature reviews (Ayres & Langone, 2005; Delano, 2007; Hitchcock et al., 2003; McCoy & Hermansen, 2007; Shukla-Mehta et al., 2010) and quantitative synthesis (Baker, Lang, & O'Reilly, 2009; Bellini & Akullian, 2007) on VBM have several limitations. First, the inability to calculate p-values or confidence intervals with PND prevents meaningful comparison of results across studies and participants based on overlap (Kavale, 2010; Parker, Vannest, & Brown, 2009; Vannest, Davis, Davis, Mason, & Burke, 2010) and limits one of the benefits of meta-analysis, which is the ability to assess for variability based on study features (Kavale, 2010; Scruggs & Mastropieri, 1998).

Second, the reviews failed to differentiate types of VBM and variations in procedural implementation. Identifying "video modeling" as an effective intervention, is imprecise, as the term is used liberally to embody a wide variety of interventions employing the theory of modeling via video recordings (Delano, 2007). When VBMs are evaluated without segregating based on the type of VBM employed as well as the additional components included in the intervention, the comparisons cannot be considered equivalent (Shukla-Mehta et al., 2010). The previous reviews do not evaluate differences in implementation protocols and the impact those protocols are likely to have on treatment outcomes. Partitioning the current VBM literature according to variations in delivery of the intervention by model, presentation, and additional components will allow for analyses that identify differential effects due to procedural implementation variables.

Finally, a thorough evaluation of the quality of these studies is warranted. Although Bellini and Akullian (2007) evaluated the quality of the VBM studies in their

meta-analysis, clearer and more precise criteria for the assessment of single-case research attributes have been established (Kratochwill & Levine, 2010). With the emphasis on implementing evidence based practices, an assessment of the quality of the literature being reviewed to determine the reliability of the studies is necessary (Jitendra, Burgess, & Gajria, 2011; Lane, Kalberg, & Shepcaro, 2009).

The purpose of this meta-analysis is to quantitatively analyze studies utilizing VBM to provide specificity regarding the implementation factors that yield the greatest magnitude of change on targeted outcomes. The following questions are addressed: (a) Are there differential effects on participant outcomes based on the type of VBM? (b) Does the model influence the magnitude of effect on outcomes? (c) Are there differential effects based on whether VBM is used alone, with reinforcement, or as part of an intervention package? In addition to answering these questions the quality of research is evaluated.

Method

Study Identification

Search method. Studies were identified through an electronic search utilizing the ERIC and PsycInfo database through the Cambridge Scientific Abstracts (CSA) database. Additionally, the Education Full Text database through the Wilson search engine was also utilized. The search was limited to peer-reviewed studies and the following Boolean string searches were conducted: *modeling* or *"observational learning"* and *disability* or *autism* or *ADHD* or *"attention deficit hyperactivity disorder"* or *behavior disorder* or

"developmental delay" or *"mental retardation"* and *video* or *videotape* or *video self-modeling*. The search resulted in 182 articles.

Inclusion criteria. Following the literature search, each article was evaluated to determine if the criteria for inclusion in the comprehensive review and meta-analysis were met. The criteria included: (a) the independent variable was a video based intervention using another person or self as model shown from a third- person perspective; (b) the article was published in English; (c) the article was published in a peer –reviewed journal; (d) the outcome variable(s) included communication, social skills, academic skills, challenging behavior, or adaptive behaviors; (e) at least one of the participants was identified as having a disability; (f) the study utilized a single-case research design establishing experimental control as evidenced by three or more phases (i.e., multiple baseline, reversal, changing criterion); and (g) raw data were provided in some format (i.e., line graph, table) identifying scores with time sequence. Studies implementing VBM from a first-person perspective, point-of-view, were excluded.

Procedure. Each study was reviewed by two evaluators to determine if the criteria were met for inclusion in the study. A third evaluator reviewed any studies for which the first two evaluators disagreed or one evaluator was undecided. The decision made by two of the three evaluators was the final decision. Of the original 182 identified articles, 53 met the criteria for inclusion in the study. The primary reason for exclusion was the article did not utilize a single-case design. Other reasons included participants did not have disabilities, the independent variable was not a VBM, and the study did not establish experimental control. The reference lists of these articles were then scanned in an effort to

find any studies that had not been identified in the initial search. This process resulted in 3

additional studies meeting the criteria for inclusion in the meta-analysis. All portions of

the search process resulted in a total of 56 single-case studies to be included.

Quality of Research Evaluation

The quality of research was evaluated utilizing a rubric based on the recommended

quality indicators of methodologically sound single-case research (Horner et al., 2005;

Kratochwill & Levin, 2010). The single-case quality indicators according to Horner

include: (a) describe participants and settings; (b) dependent variable; (c) independent

variable; (d) baseline; (e) experimental control/validity; and (f) social validity. Horner et

al. include components for each indicator, a total of 21 components, which further define

the conditions for meeting the criteria of the component. Kratochwill et al. (2010) include

similar indicators, however, the criteria for meeting each indicator is more explicit

particularly regarding what must exist to meet the standard (e.g., requires a minimum of 3

data points in a phase and specifies the number of phase changes necessary dependent on

design). Additionally, Kratochwill et al. (2010) place greater emphasis on evidence

regarding internal validity including guidelines that must be met when conducting a visual

analysis to address the question that the implementation of the intervention resulted in a

consistent and stable pattern of behavior change**.**

Quality Rubric. For the current review a 4-point rating scale approximating that

utilized by Chard et al. (2009) was developed for each of the 21 components with precise

and quantifiable descriptors for each anchor (see Appendix A), based on the criteria

established by Horner et al. (2005). The rubric is available upon request from the first

author. Each study was then scored by the first author for each of the 21 components utilizing the scale. The ratings for each component of a given standard were averaged in order to obtain an overall score for the quality standard (i.e., participant selection, dependent variable, independent variable.

An average score of 3 or higher was considered to meet the minimal standards for the indicator. However, if any component received a score of 1, then the indicator was considered not to meet minimal standards as it was lacking a significant aspect. An overall quality rating was then calculated by averaging the ratings across all components. Horner et al. (2005) indicate all indicators must meet the minimum standards in order to be considered methodologically sound. Given this, any study with one or more indicator that earned a score of less than three or a component with a score of 1 was not considered to meet the established criteria as a quality research study.

Extraction of Descriptive Information

Each study was reviewed to extract information on key features of the study including participant age, disability, and setting, study design, and dependent variables. The type of VBM, as well as additional intervention components included with the VBM, were noted for the independent variable.

Potential Moderators Coding

Moderator variables can account for differences in intervention effects, as factors that determine the impact, both the direction and potency, the intervention has on the outcome (Holmbeck, 1997) and can typically account for discrepancies in results.. Given the heterogeneity of individuals with disabilities, identification of for whom and under

what conditions an intervention yields meaningful acquisition of the target skill is essential (Odom et al., 2005). The studies in this review were coded for the following potential moderators: (a) VBM variable (VMO or VSM), (b) model variables, (c) and implementation variables.

VBM variable. This variable was divided into two levels: VMO and VSM.

Model variables. This included three variables: *type of model, familiarity, and production*. The *type of model* variable included three levels: other, peer, and self. *Familiarity*, which only applied the VMO studies, consisted of two levels, known and unknown. *Production* consisted of two levels, positive self-review and feedforward, and only applied to the VSM studies.

Implementation variable. The implementation variable consisted of three levels, similar to those defined by Shukla-Mehta et al (2010). The levels included alone, with reinforcement only, and component of a package. The reinforcement only level included studies that delivered reinforcement for attending to the video or for performance of the target skill. The component of a package level consisted of any study that employed other interventions, such as least-to-most prompting, social stories, role play and discrimination training.

Effect Size and Replication Analysis

Although there has been disagreement regarding the most appropriate effect size to utilize for single-case analysis, the improvement rate difference (IRD), has been applied in a number of recently published meta-analysis (Ganz, Earles-Vollrath, Heath et al., 2011; Ganz, Earles-Vollrath, Mason et al., 2011; Vannest, Davis, Davis, Mason, &

Burke, 2010; Vannest, Harrison, Parker, Harvey, & Ramsey, 2010).   Furthermore, IRD is

a popular statistical analysis employed in group research within the medical field, where it

is referred to as "risk difference" (Parker et al., 2009; Sacket, Richardson, Rosenbert, &

Haynes, 1997).

      Robust IRD is a variation of IRD that is less susceptible to variation due to

outliers, resulting in a more dependable measure of change (Parker, Vannest, & Davis,

2011).   Just like IRD, robust IRD is an effect size based on non-overlap of data between

the baseline and intervention phase, (Parker et al., 2009) which also provides a means for

the calculation of confidence intervals. Robust IRD quantifies the amount of

improvement that occurs in the intervention phase beyond the improvement that occurred

in the baseline phase (Parker et al., 2009). Robust IRD is equal to Phi in a 2x2 matrix

(Parker et al., 2011)

      With possible values ranging from -1 to 1, positive effects range from greater than

0 to 1 and deterioration effects range from -1 to less than 0. Robust IRD values of less

than .5 can be described as minute, or less than chance. Likewise, scores from .50 to .70

can be expressed as moderate change, whereas scores above .70 can be judged as

considerable effects (Parker et al., 2009).

      Phase contrasts. Effect sizes for phase contrasts were calculated in order to obtain

a measure of the magnitude of change that occurred between a minimum of two phases.

The baseline and initial intervention phase (A vs. B) were contrasted for each level

included in a study yielding individual effect sizes. For reversal designs ($A_1B_1A_2 B_2$),

only adjacent phases were contrasted (i.e., A_1 vs. B_1 and A_2 vs. B_2) and then aggregated as this method is congruent with visual analysis.

For multiple baseline studies utilizing more than one intervention phase (ABC), A vs. B contrast and A vs. C contrast were conducted and then individual effect sizes were combined when appropriate, depending on the intervention of interest. The levels were then aggregated to obtain an overall effect size for those multiple baseline designs in which the intervention, dependent variable, and contextual factors were similar. Maintenance and generalization were not included in phase contrasts.

Individual phase contrast effect sizes were than combined as appropriate in order to obtain a single effect size for a given dependent (moderator) variable. In order to combine the effect sizes, they are first weighted by the inverse of their standard error, which is automatically completed as part of the calculations in the statistical software.

Effect size calculation. Robust IRD was calculated for each study as well as for each participant and level in the included studies. Overlap was specified to be any data point in the baseline phase tied with or higher than a data point in the intervention phase or any data point in the intervention phase tied with or lower than a data point in the baseline phase. After the overlapping data points were identified, the minimum number of removed data points necessary to eliminate overlap was determined. The overlapping data points were then divided equally between each contrasted phase and the resulting ratios, improved to not improved totals, were entered into a 2 x 2 matrix for each contrast (Parker, Vannest, & Davis, 2011).

The data were then entered into the "risk analysis" module, specifically for meta-analysis research, of the Number Cruncher Statistical Software (NCSS: Hintze, 2002). This meta-analysis module is constructed from the difference of the separate ratios. The independent effect size (IRD) and combined effect sizes were then calculated. For the combined effect sizes, NCSS weights each individual effect size based on the "inverse of the standard error" and calculates a statistically average effect size utilizing algorithms developed for meta-analyses. The NCSS output also produces a Forest Plot of the bootstrap confidence interval for each calculated effect size, providing a graphic depiction allowing for visual analysis of replication and outliers (Hintze, 2002).

Statistical Significance

With the purpose of distinguishing differences between levels of the moderators, tests of statistical significance were conducted. This was achieved through the use of 83.4% confidence intervals. Statistically significant (p = .05) differences can be said to have occurred when the confidence intervals of a given measure, in this case robust IRD effect size, do not overlap at the upper and lower limits (Payton, Miller, & Raun, 2000; Payton, Greenstone, & Schenker, 2003; Schenker & Gentleman, 2001). This test of nonoverlap is equivalent to student t-test of statistical significance (*p*=.05) and, through the use of a forest plot , allows for a visual analysis of statistical significant differences between multiple effect sizes (Payton et al, 2000). All confidence intervals were computed at the 83.4% level.

Interrater Reliability

        Quality rubric. In order to assess the reliability of the quality coding, 33% of the studies were coded by a second evaluator who was unaware of the initial coding results. Simple percent agreement [(agreement/agreements +disagreements) x 100] was calculated for each of the 21 components. Percent of agreement for each component can be found in Table 1 noting that the mean percent of agreement was 90% with a range of 83-100%. As the rubric utilized ordinal data linear weighted (LW) Kappa was calculated. The obtained LW Kappa for the combined rubric was .80. Although percent of agreement was adequate, obtained LW Kappa ranged from .56 to 1 across components, for which individual scores can be found in Table 1. As the top end of the scale was predominantly utilized across most components, Kappa overcorrected for the imbalance in scale usage (Brennan & Prediger, 1981). Thus, prevalence and bias adjusted Kappa for ordinal scales (PABAK-OS) was calculated for each component, as well as the aggregated results for each indicator.

Table 1.

Summary of interrater reliability for quality indicator components.

Indicator	Component	% agreement	LW Kappa	Pabak 0S	95% CI	p-value
Participant and Setting	Participant Characteristics	100	1	1	[.77, 1.23]	<.000
	Selection	83	0.87	0.89	[.67, 1.1]	<.000
	Setting	94	0.92	0.95	[.72, 1.17]	<.000
	Overall	94	0.96	0.95	[.81, 1.1]	<.000
Dependent Variable	Description	89	0.59	0.85		<.000
	Description of Valid Measure	89	0.8	0.9	[.67, 1.13]	<.000
	Frequency of Measurement	83	0.67	0.85	[.62, 1.08]	<.000
	IOA	94	-	0.95	[.72, 1.18]	<.000
	Overall	88	0.7	0.89	[.77, 1]	<.000
Independent Variable	Description	94	0.77	0.95	[.72, 1.1]	<.000
	Manipulation	83	0.57	0.85	[.62, 1.08]	<.000
	Fidelity	94	0.96	0.95	[.72, 1.19]	<.000
	Overall	91	0.91	0.92	[.78, 1.05]	<.000
Baseline	Dependent Variable	83	0.75	0.85	[.62, 1.08]	<.000
	Baseline conditions	89	0.69	0.9	[.67, 1.13]	<.000
	Overall	88	0.72	0.87	[.71, 1.04]	<.000
Experimental Control/Internal Validity	Design establishes	94	0.92	0.95	[.72, 1.17]	<.000
	Threats restricted	83	0.72	0.85	[.62, 1.1]	<.000
	Pattern of results	94	0.93	0.95	[.72, 1.18]	<.000
	Overall	91	0.87	0.92	[.78, 1.05]	<.000
External Validity	Replication of Results	94	0.77	0.95	[.72, 1.18]	<.000
Social Validity	DV socially important	100	1	1	[.77, 1.23]	<.000
	Change socially impactful	89	0.68	0.9	[.67, 1.13]	<.000
	Practical and Cost effective	83	0.76	0.85	[.62, 1.08]	<.000
	Typical implementation	83	0.7	0.85	[.62, 1.08]	<.000
	Overall	89	0.8	0.9	[.79, 1.01]	<.000

PABAK-OS is a new measure of interrater reliability for ordinal scales that differentially weights disagreements based on degree of disagreement (Parker, Vannest, & Davis, 2011). Thus, with a 5-point scale, a 1-2 disagreement is given more credit than a 1-5 disagreement. Additionally, PABAK-OS does not overcorrect for unequal use of the scale nor give differential credit based on which portions of the scale are utilized,

problems frequently encountered with the more common reliability measure such as,

Cohen's Kappa (Parker et al., 2011). The scores for each rater were placed into a 4 x 4

agreement matrix which was then inserted into the PABAK-OS calculator with half

credit given for disagreements within one (Vannest, Parker, & Gonan, 2011).   The

obtained PABAK-OS for the components ranged from .89 to 1. PABAK-OS for each

individual score for each component can be found in Table 1. Based on reliability cutoff

standard of 80% agreement and a reliability index of at least 60% (Horner et al., 2005),

the findings suggests adequate inter-rater reliability for the quality indicator rubric.

Moderators. The coding for moderator variables was completed by both the

investigator and another rater for 37% of the studies in order to establish inter-rater

reliability. Interrater agreement was determined by a simple percent agreement. Percent

of agreement across variables was 100%.

Robust IRD. From the 56 studies a total of 233 robust IRD matrices were

calculated. Of these, 68% were individually computed by the author and another rater to

ensure reliable calculations. Interrater agreement was determined by a simple percent

agreement (agreement/agreements +disagreements) x 100) calculation. The overall

agreement was 94%. The 10 disagreements were discussed between the two raters and

then the IRDs were computed again until 100% agreement was achieved. The

disagreements were due to human error and difficult-to-read graphs.

Results

Descriptive Summary of Studies

Of the 56 studies included in the analysis, 42 utilized VMO as the intervention and 18 included VSM. The total exceeds the total number of studies because four of the articles included both VMO and VSM (Cihak & Shrader, 2008; Marcus & Wilder, 2009; Sherer et al., 2001; Van Laarhoven, Zurita, Johnson, Grider, & Grider, 2009). A total of 177 participants were included in the studies. The VMO studies included participants ranging in age from 2-72 with diagnoses of either ASD or intellectual disability. The VSM studies included participants ranging in age from 4 to the late 20s with diagnoses of ASD, intellectual disability, emotional behavioral disorders, or learning disabilities. Both the VMO and VSM studies addressed a variety of targeted outcomes including socio-communicative, play, academic, adaptive behavior, and independent living skills. A summary of descriptive information obtained from each of the included VMO and VSM articles can be found in Appendices B and C respectively.

Intervention components. Of the 42 studies implementing VMO, 40% utilized a model unknown to the participants and 31% utilized a model with whom the participants were familiar. Twenty-nine percent of the studies did not specify whether or not the participants were familiar with the model. Adults were the models for 55% of the studies peers were the models for 48%. One study (Keen, Brannigan, & Cuskelly, 2007) utilized animated models to teach toilet training to children with autism, and although this was considered a VMO intervention, the model does not fit into either the adult or peer category.

VMO was implemented as part of a package in 23% of the studies. Other interventions included with VMO were error correction procedures (e.g., least-to-most prompting) (Bidwell & Rehfeldt, 2004; Cihak & Shrader, 2008; Gena, Couloura, & Kymissis, 2005; Haring, Kennedy, Adams, & Pitts-Conway, 1987; Marcus & Wilder, 2009; Mechling, Pridgen, & Cronin, 2005; Murzynski & Bourret, 2007; Rehfeldt, Dahman, Young, Cherry, & Davis, 2003; Taber-Doughty, Patton, & Brennan, 2008; Taylor, Levin, & Jasper, 1999; Van Laarhoven et al., 2009), in-vivo modeling (Gena et al., 2005; Haring et al., 1987), social stories (Sansosti & Powell-Smith, 2008), and picture cards (Keen et al., 2007). VMO was used alone in 45% of the studies; 24% included reinforcement and no other components with the VMO intervention.

Of the 18 studies utilizing VSM, 55% implemented feedforward VSM and 50% implemented positive self-review. The total exceeds 100% because Buggey (2005) included two experiments utilizing feedforward and one utilizing positive self-review. Forty-four percent of the studies implemented VSM alone whereas 22% included reinforcement with the intervention. VSM was implemented with other components such as error correction procedures, role play, discrimination training and self-monitoring for 33% of the included studies.

Quality of studies. Utilizing the quality indicator rubric, each study included in the analysis was rated on the 21 components that comprised the seven quality indicators of single-case research (Horner et al., 2005). Fifteen of the VMO and 5 of the VSM studies met the minimum standards across all 7 indicators and 21 components (indicated by an asterisk by the QI average in Appendices B and C). Failing to meet minimal

standards for the *Independent Variable* indicator was the primary reason for not meeting

minimum standards overall, because only 28% and 45% of the VMO and VSM studies,

respectively, included measurement of fidelity of implementation. The 15 quality VMO

studies were completed by 15 authors and yielded 53 IRD effect size calculations; the

average IRD = .82 (CI [.80, 84]) with an IRD range of -.17 to .96. Twenty IRD effect size

calculations ranging from -.27 to .88 were obtained from the 5 VSM studies that met the

quality study criteria; the average IRD = .61 (CI [.56, 65]).

Effect Size and Replication Analyses

      Overall. When a study included participant data replicated for the same

independent and dependent variable the IRDs were combined. This resulted in a total of

233 IRD effect sizes across the 56 studies and 177 participants. The results yielded an

average IRD across studies of IRD = .81 (CI [.80, .82]) with an IRD range from -.26 to

.96. The large effect size and confined confidence interval lends further evidence that

VBM is an effective intervention for individuals with disabilities. However, the wide

range of results suggests variations in implementation yield differential effects. It was

hypothesized that the type of VBM utilized as well as protocol variation and inclusion of

other components facilitated disparate outcomes. Moderating variables were analyzed to

determine if differential effects occur based on implementation protocol. The initial

moderating variable analyzed was exploration of differential effects based on the three

levels of the *implementation variable* including alone, with reinforcement, or part of a

package. The IRD calculated for each of the variables indicated high effects across

variables; however, VBM with reinforcement IRD of .86 (CI [.84, .88]) was significantly

higher than VBM alone (IRD = .81, CI [.79, .82]) and VBM as part of a package (IRD= .75, CI [.73, .77]). Additionally, the calculated effect size for VBM as part of a package was significantly lower than both VBM alone and with reinforcement.

VBM variable. Of the 233 IRD effect sizes, 171 of the IRD effect sizes were for the VMO studies and 62 were for the VSM studies. No statistically significant difference was found when comparing VMO to VSM. VMO yielded a combined effect size of .82 with an IRD range of -.25 to .96. Likewise, VSM yielded a combined effect size of .79 with an IRD range of -.26 to .95. Again, the broad range of effect sizes suggests effects are likely moderated by other variables.

As is illustrated in Figure 1, the data from two VMO studies (Char lop-Christy & Daneshvar, 2003; Rosenberg, Schwartz, & Davis, 2010) and two VSM studies (Hepting & Goldstein, 1996; Hitchcock et al., 2004) yielded overall IRD effect sizes that were significantly smaller than the overall IRD of .81 and for the results obtained for the rest of the studies. Charlop-Christy and Daneshvar (2003), utilizing a multiple-baseline design across participants and tasks, implemented video modeling to increase perspective taking for three males. Although some improvement did occur, a pattern of results was not evident across all participants and tasks. Using commercially and custom created VMO to teach hand washing to three preschoolers with autism, Rosenberg et al. (2010) demonstrated skill acquisition with only one of the participants. Hepting and Goldstein (1995) implemented a VSM intervention to increase verbalizations with three preschool aged children. Improvement did occur across participants; however, all required several exposures to the VSM before gains were evident.
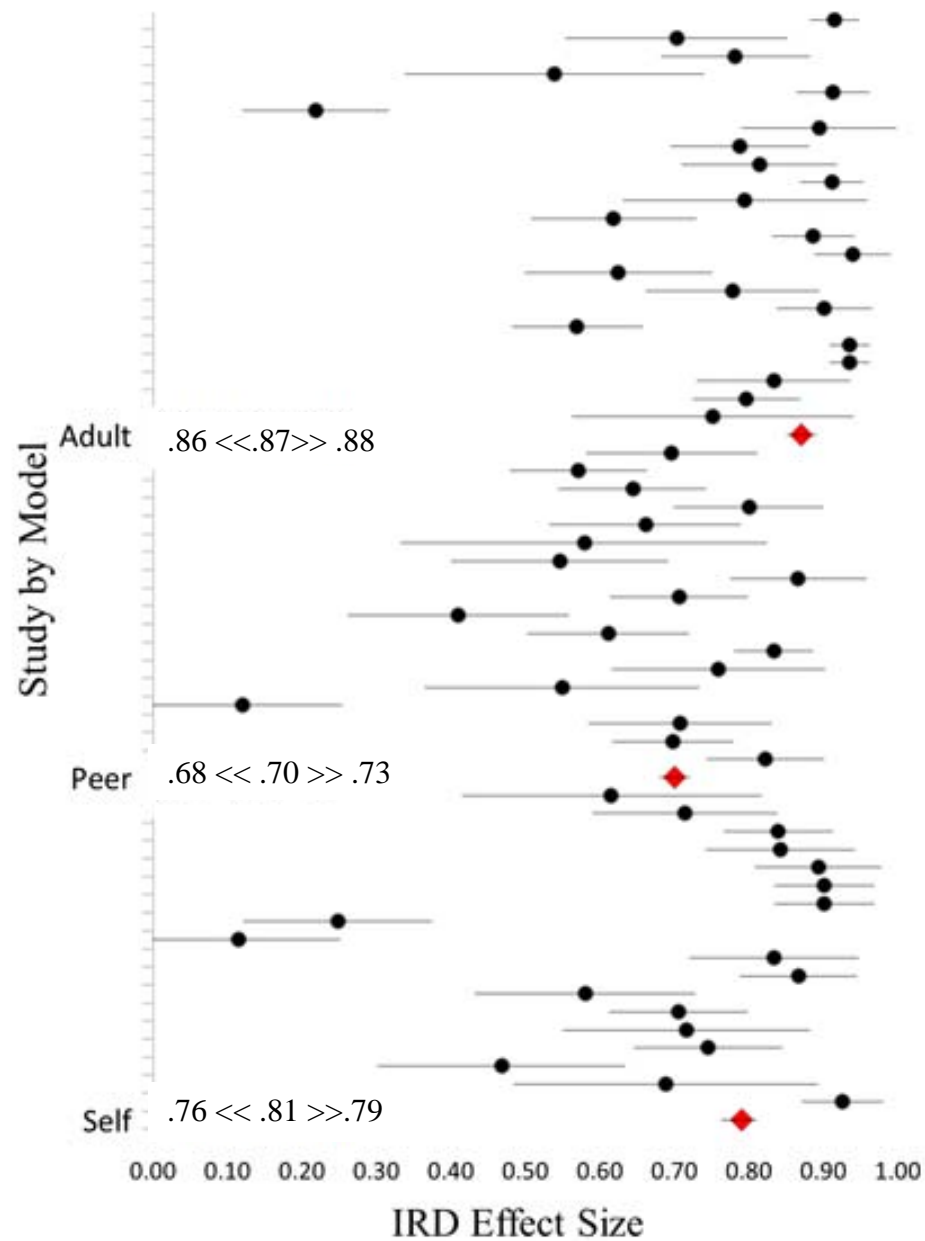
Figure 1. Forest plot of studies 83.4% CIs by model.

Note. Each circle represents the combined IRD effect size for a particular model within a single study. Diamonds represent the omnibus effect size for each level of the model variable**.**

It was hypothesized that one variable that likely moderates the outcomes was *type of model*. Eighteen VSM studies and 41 of the included VMO studies contributed to this analysis (Keen et al., 2007, was not included due to the animated model). The calculated effect sizes all yielded a high magnitude of change across the *type of model* variable, however, as is evident by non-overlapping confidence intervals (see Figure 1), statistically significant differences exist. Peer as Model yielded the smallest effect size (.70) and Adult as Model yielded the largest effect size (.87). Thus, when considered together there is not a differential effect between VMO and VSM. However, when disaggregated by type of model, VMO with adult as model is significantly more effective than VSM.

VMO. The 42 VMO studies were further analyzed to explore differential effects based on applicable *implementation and model* variables. The data for the *familiarity* variable, as well as *implementation* variable from the analyzed studies was further disaggregated based on type of VBM to explore applicable differential effects.

As was previously discussed, results of the analysis based on the *model* variable utilized indicated statistically significant differences in magnitude of change when an adult was used as model as opposed to a peer. The VMO studies were further analyzed based on the two levels of the *familiarity* variable, known and unknown, to determine if differences existed. The IRD calculated for both indicated large effect sizes with overlapping confidence intervals (see Figure 2) indicating no statistically significant differences between the two levels.
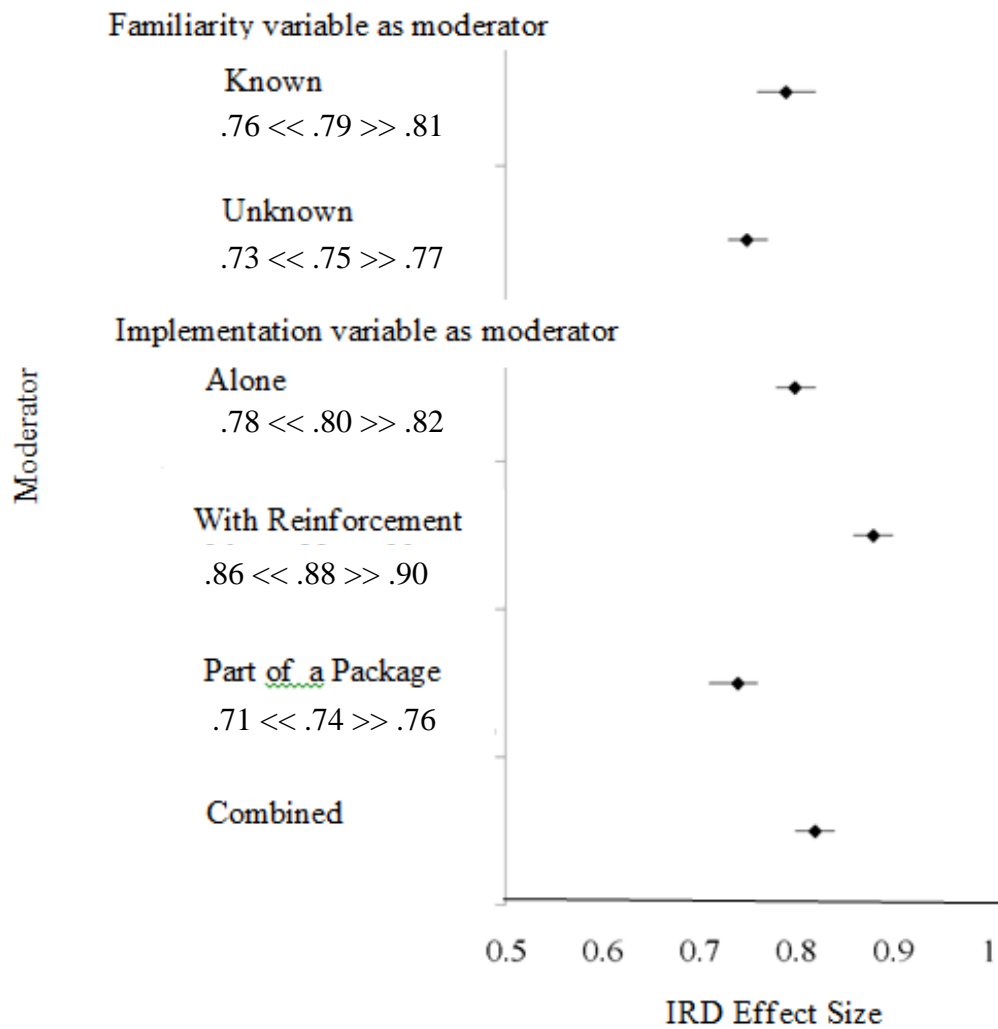
Figure 2. VMO IRD and 83.4% CIs for levels of familiarity and implementation variables.

In an effort to address the research question regarding implementation an analysis of the *implementation* variable with the levels alone, with reinforcement, and as part of a package was conducted. Again, effect size calculation based on *implementation* yielded large effect sizes (see Figure 2) across all levels. However, the obtained effect size for *VMO with* reinforcement (.88) is significantly higher than the other levels, implementation alone (.80) and as part of a package (.74). Additionally, the obtained effect size for *VMO implementation as part of a package* was significantly smaller than obtained results for the other two levels.

VSM. The18 studies employing VSM were analyzed to determine if differential effects occurred based on the *implementation* variable as well as the *production* variable. The *production* variable(applicable only to VSM) included the two levels previously described, positive self-review and feedforward Results (see Figure 3) indicated large effect sizes for both positive self-review (.82) and feedforward (.77) with no statistically significant differences as indicated by overlapping confidence intervals.

Again to address the research question regarding implementation of VSM an analysis of the *implementation* variable with the same levels as those utilized for the overall analysis of implementation and for VMO, was conducted. Eight of the VSM studies implemented VSM alone, 6 implemented VSM as part of a package, and only 4

Production variable as moderator

Positive Self- Review
.78 << .82 >> ,85

Feedforward

Implementation variable as moderator

Alone
74 << 78 << 82

With Reinforcement
.61 << .66 << .72

Part of Package
.80 << .83 << .86

Combined
.76 << .79 << .82

Moderator
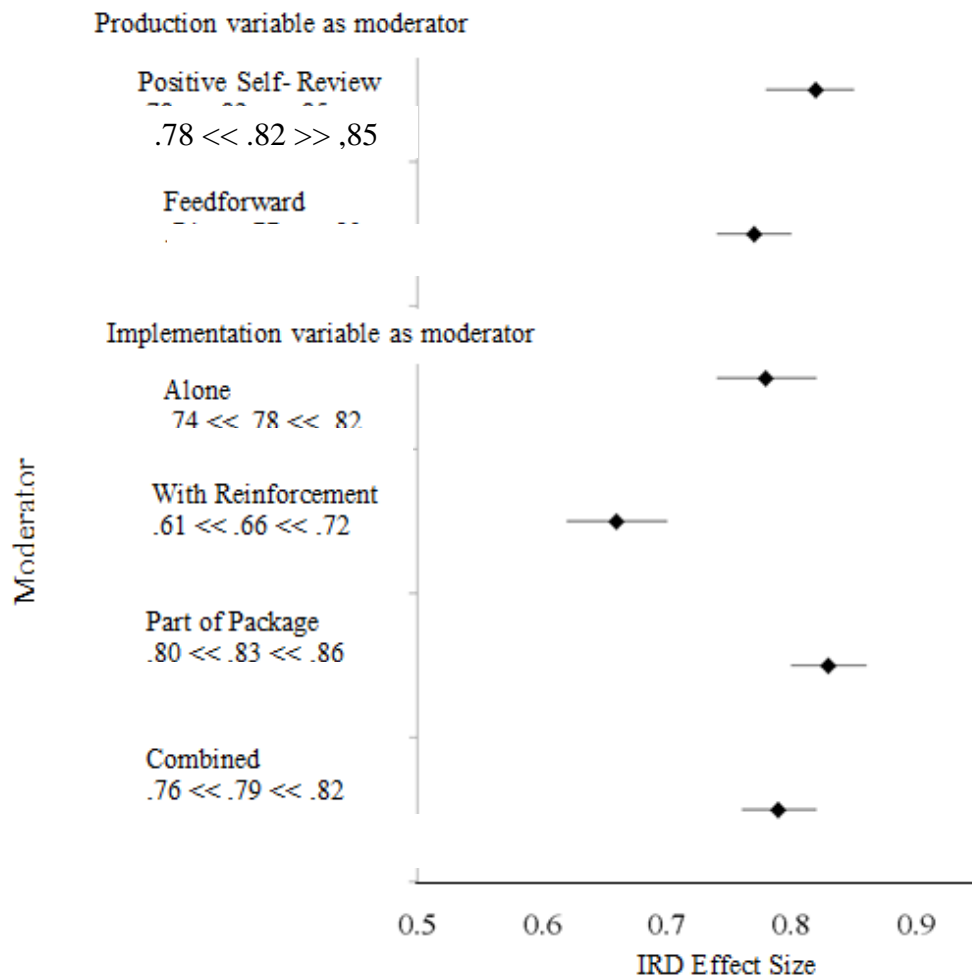
0.5   0.6   0.7   0.8   0.9

IRD Effect Size

Figure 3. Forest plot of VSM IRDs and 83.4% CIs by levels of production and implementation variables.

implemented VSM with reinforcement only. Results are visually depicted in Figure 3.

Large effect sizes were found for those studies utilizing VSM alone (.78) and VSM as part

of a package (.83), whereas those studies utilizing VSM with reinforcement yielded a

moderate effect size (.66). The smaller effect size obtained for VSM with reinforcement

demonstrated a statistically significant (p=.05) difference from VSM alone and as part of

a package as is evident by the non-overlapping confidence intervals (Figure 3).

Discussion

The purpose of this meta-analysis was to evaluate the evidence base for the use of

VBM as an intervention for individuals with disabilities. Overall results indicate large

effects with moderate to large effects for nearly all of the included studies which is

consistent with Bellini and Akullian's (2007) previous meta-analysis. Further analysis of

moderators provides clarity regarding which implementation variables produce the

greatest magnitude of change on participant outcomes. Results of the meta-analyses

answer the research questions regarding differential effects that occur based on the type of

VBM and the influence of the model on magnitude of effect, as well as differential effects

that occur based on whether VBM is used alone, with reinforcement, or as part of an

intervention package. Additionally, the evaluation of the quality of research addresses the

question regarding the methodological soundness of the VBM research.

The study first sought to identify if any differences in magnitude of effect occur

when VMO was implemented as opposed to VSM. When considered in aggregate,

statistical comparison of VMO to VSM indicate large effect sizes for both with no

differential effects based on the type of VBM utilized, consistent with results obtained by

Bellini and Akullian (2007). However, when disaggregated based on type of model,

differences in results are clear.

The second question addressed by this study is whether differential effects occur

based on the type of model utilized. Contrary to what might be expected given the

guidelines that modeling and thus VBM is most effective when the model resembles the

observer (Bandura, 1969), results of this analysis indicate VMO with adult as model is

more effective than VMO with peer as model and VSM. Familiarity of the VMO did not

result in differential effects nor were differential effects indicated between VSM

feedforward and positive self-review. From a practical standpoint this finding is

promising because VMO with an adult is the most efficient method for producing a video

model. VSM is likely the most challenging to produce due to the lengthy editing process

as well as the prompting required for feedforward and the lengthy recording time required

for positive self-review. In fact, the complexity for producing an effective VSM may be

one of the factors contributing to the lower effect sizes. However, this does not mean

VSM should not be used because large effect sizes were obtained for VSM. Additionally,

VMO studies were primarily implemented with participants with ASD and ID whereas

empirical evidence for VSM includes participants with other disabilities such as learning

disability and emotional disturbance. Given this, there is a lack of empirical evidence to

support the use of VMO with individuals with disabilities such as emotional behavior

disorders and learning disabilities. Given the lack of differential effects when comparing

positive self-review to feedforward, the implementation of feedforward is likely the better

option as positive self-review is the more challenging video to produce, potentially

requiring a considerable amount of time to collect an adequate sample of the target

behavior (Buggey, 2005; Dowrick, 1999).

Differential effects that occur due to variations in the implementation protocol are

another question addressed by this review. When disaggregated based on implementation

variables such as whether VBM was implemented alone or as part of package, differential

effects are clarified. Considering all types of VBM together, the results indicate VBM

with reinforcement yields greater effects than VBM alone and as part of a package which

is consistent with recommendations provided by Shukla-Mehta (2010) that reinforcement

be included as part of the intervention. Additionally, VBM implemented as part of a

package demonstrated significantly smaller effects on the targeted outcomes than alone or

with reinforcement. Analysis of the intervention variable for VMO yielded the same

results, however, results for VSM indicated no difference between VSM alone or as part

of a package, whereas VSM with reinforcement yielded statistically significant smaller

effect sizes. Given that the majority of the VMO studies involved participants with ASD,

it is hypothesized that the reason VBM as part of a package was less effective is due to

learner characteristics. Bellini and Akullian (2007) have suggested the efficacy of VBM

with individuals with ASD may be attributable to the removal of the need to communicate

with others while engaging in the learning component of the task. Additionally, VBM

capitalizes on the propensity of participants with ASD to focus on visual stimuli to

facilitate learning (Ayres et al., 2009). Thus, adding additional components such as least-

to-most prompting and role play might actually reduce the effectiveness of the

intervention. Given the small number of VSM studies included in each category, the results must be viewed with caution. Additional research utilizing VSM alone and with reinforcement in particular needs to be conducted before definitive statements can be made in regards to the most efficacious protocol.

Review of the quality of VBM research indicated both VMO and VSM can be considered to be evidence-based practices. Fifteen VMO studies, completed by 15 different authors with 53 IRD effect sizes, met at least minimum quality standards. Additionally, 5 of the 18 VSM, completed by five authors with a total of 20 IRD effect sizes met the minimum standards. The overall primary reason for studies not meeting quality standards was a lack of reported fidelity measurement. Lack of fidelity measures inhibits the capacity to make inferences that the results of the study were due to the effectiveness or ineffectiveness of the intervention (Cook, Tankersley, & Landrum, 2009; Gresham, MacMillan, Beebe-Frankenberger, & Bocian, 2000) Bellini and Akullian (2007) pointed out that the reason for the lack of fidelity measurement is likely due to the nature of VBM in that the video is produced and then viewed by the observer. However, such features of VBM interventions such as implementation schedule and whether or not reinforcement or prompts were used is important procedural information. The lack of fidelity measures limits the confidence in the results of the study and also limits the practical application of the intervention.

Limitations

Several limitations exist for this meta-analysis. Primarily, although many of the studies did not meet quality standards, they were not eliminated from the meta-analysis.

The decision not to eliminate the studies was twofold. First, the standards (Horner, et al, 2005; Kratochwill & Levin, 2011) utilized by the authors to establish criteria have been developed only as a guide and no steadfast guidelines for evaluation of rigorous single-case methodology currently exist. Second, many of the studies did not meet quality due to limited descriptions of factors that facilitate replication and failure to address social validity questions. Although these are omissions that limit the ability to expand on the knowledge regarding VBM as well as the practical application of VBM, the omissions do not interfere with internal validity, a necessary component for statements regarding intervention causality.

The lack of fidelity measures, however, does limit the confidence one can have in intervention efficacy and is a common limitation of educational research (Cochrane & Laux, 2008; Conroy, Stichter, Daunic, & Haydon, 2008). However, elimination of the studies would have resulted in a significantly smaller sample size that would have yielded inconclusive results. Inclusion of fidelity of implementation measures is an area to be addressed in future research to increase the empirical support of VBM, as well as to assist with the establishment of operationalized treatment procedures. Another limitation is the inclusion of only published research which, given the inherent bias towards only successful interventions, discounts those instances in which VBM was not effective in changing the targeted outcome.

Implications for Future Research

Results of this meta-analysis suggest several questions to be addressed by future research. Although the results indicate VMO with adults as models is most effective,

questions regarding for whom and for what remain unanswered. Future research, perhaps in the vein of a meta-analysis should further explore this area, considering participant age and disability categories as well as targeted outcomes. Expansion in this regard might lend evidence that VSM and/or VMO with peers are more effective for certain participants than VMO with adult. This more precise desegregation would further assist in establishment of specificity regarding the conditions for implementation of VBM. Further, as has been previously mentioned, expanded details to facilitate replication as well as measures to address social validity are necessitated.

In summary ample evidence exists to suggest VBM interventions are efficacious intervention across variables. Perhaps even more promising is that the most efficient method, VMO with adult as model, tends to produce the greatest magnitude of change on the targeted outcome. Additionally, greater effects occur when supplemented with reinforcement only, resulting in a highly practical and manageable educational intervention for individuals with disabilities.

CHAPTER III

MODERATING FACTORS OF VIDEO-MODELING WITH OTHER AS MODEL:

A META-ANALYSIS OF SINGLE-CASE STUDIES[*]

The implementation of evidence-based practices (EBPs) is critical to improving the behavioral and learning outcomes for individuals with disabilities (Cook, Tankersley, & Landrum, 2009). Guidelines for determining an EBP for single-case research include at least 5 published, peer-reviewed and methodologically sound studies carried out by a minimum of 3 different investigators from three different geographical locations with a minimum of 20 participants across studies (Horner et al., 2005; Kratochwill et al., 2010). However, quantitative identification of effective practices for individuals with disabilities is often more challenging given the heterogeneity across and within the various disability categories (Montgomery, 2006). Of particular importance is identifying under what circumstances, including target population traits, intervention package components, and targeted outcomes, a particular intervention does and does not yield meaningful changes (McDonald, Keesler, Kauffman, & Schneider, 2006).

Video based modeling (VBM) is a frequently investigated intervention in special education research literature and has been identified as an evidenced based intervention (Bellini & Akullian, 2007; Nikopoulos & Nikopoulos-Smyrni, 2008; Rayner, Denholm, & Sigafoos, 2009). VBM, the process of recording the performance of targeted behaviors with the anticipation that the observer will cognitively internalize and later reproduce the observed behaviors (Bellini & Akullian, 2007; C. H. Hitchcock, Dowrick, & Prater, 2003; McCoy & Hermansen, 2007), has several advantages. First, VBM techniques take advantage of the effectiveness of modeling and visual strategies for improving skills (Bellini & Akullian, 2007; Biederman & Freedman, 2007). Second, VBM integrates technological modalities into instruction providing a precise and accurate exemplar of the skills being taught (Ayres & Langone, 2005). Third, presentation of VBM is uniform across trials and can be repeated within and across participants (Ayres & Langone, 2005).

VBM may be presented in three distinct variations, including (a) video-modeling with other as model (VMO), (b) video self-modeling, and (c) point-of-view modeling (Shukla-Mehta et al., 2010). VMO requires recording an adult or peer acting out a script demonstrating the targeted skill (Allen, Wallace, Renes, Bowne, & Burke, 2010), whereas video self-modeling involves recording the targeted individuals performance of a skill (Hitchcock, Dowrick, and Prater, 2003). The third option, point-of of-view modeling, involves recording the model from the perspective of the model, thus the actual model is not seen in the video (Shukla-Mehta et al., 2010). Instead of recording the entire scene, the camera may be placed at the model's shoulder level, recording only the model's hands completing a task (e.g., assembling a sandwich). VMO and video self-modeling are cited

most frequently in the literature base and both have been identified as being effective for individuals with disabilities (Bellini & Akullian, 2007). Video self-modeling, however, is more complicated and requires the targeted individual to perform the skill via prompting or capturing enough footage of the individual accurately performing the skill in the natural setting without guidance from others (Dowrick, 1999; McCoy & Hermansen, 2007) . Video self-modeling requires more time for recording and editing, as prompting from others and inaccurate performances have to be deleted (Dowrick et al., 2006; Hitchcock, Prater, & Dowrick, 2004).

Given the complexity of video self-modeling and point-of-view modeling preparation, VMO has previously been noted as the more practical option (Bellini & Akullian, 2007; Shukla-Mehta et al., 2010). However, despite the potential of VMO for efficiently improving targeted outcomes for individuals with disabilities and the practical appeal (Bellini & Akullian, 2007; Shukla Mehta et al,, 2010), a thorough examination has not yet been conducted of the differential effects based on potential moderators such as (a) participant characteristics, (b) intervention package components by participant characteristics, and (c) targeted outcomes. Moderating variables are factors that either limit or enhance the magnitude of change that occurs on the dependent variable in the presence of the independent variable (Holmbeck, 1997). An evaluation of potential moderators based on these factors would potentially provide clarity and facilitate practical implementation by distinguishing with whom, and for what skills, VMO is most likely to produce desired results.

Three qualitative reviews (Delano, 2007; McCoy & Hermansen, 2007; Shukla-Mehta et al., 2010) and one quantitative review (Bellini & Akullian, 2007) have considered VMO as a category separately from other versions of VBM.   Delano (2007) included 12 studies utilizing VMO in a review of VBM studies for individuals with ASD published between the years of 1985 to 2005. The review did not describe participant characteristics such as age or gender nor were results disaggregated based on the type of VBM implemented. Nevertheless, the review did indicate that 25% of the VMO studies reported mixed results across participants. Delano hypothesized that VMO alone, that is without additional intervention package components, may not have been sufficient for changing the increased skill. The review did not indicate if VMO appeared to be more or less effective for certain skills.

McCoy and Hermansen (2007) reviewed the use of VBM interventions with participants with ASD, disaggregating the studies based on type of model utilized. Of the 34 included studies published between 1987 and 2006, 18 investigated VMO. The studies included a total of 46 participants, 22 of whom had ASD, ranging in age from 3 to 20. McCoy and Hermansen (2007) reported generally positive results, although it was noted that additional components, such as live modeling, feedback, prompting, and picture schedules may have influenced the results. The review did not address variations in results due to participant characteristics such as age, gender, or the presence of a comorbid disability. Again, differential effects based on targeted outcome were not reviewed.

 In the most recent literature review, Shukla-Mehta et al. (2010) reviewed published studies that implemented VBM to increase social and communication skills for

participants with ASD. Shukla-Mehta et al. considered the type of VBM utilized and also

separated the studies based on whether VBM was delivered alone or as a component of a

package. Based on17 studies that utilized VMO, Shukla-Mehta et al. concluded that

VBM implemented as part of a package was more effective in improving social and

communication skills than VBM implemented alone for participants with ASD.

Magnitude of change was not calculated and information regarding whether this was

consistent across participant characteristics was not included as a component of this

analysis. Additionally, published studies that targeted other skills such as play and/or

independent living were not included.

In a meta-analysis comparing the use of video self-modeling with VMO for

participants with ASD, Bellini and Akullian (2007) reviewed 23 studies published

between the years 1980 to 2005. Utilizing percent of non-overlapping data for, Bellini

and Akullian determined that both VMO and video self-modeling were effective

interventions for individuals with ASD. Further, a Kruskall-Wallis test found no

statistically significant differences in magnitude of change between the implementation of

VMO and video self-modeling for individuals with ASD indicating that both were equally

effective. However, as demonstrated in Chapter 2, when further disaggregated by type of

model as a moderator with levels of self, other, and peer, statistically significant ($p = .05$)

differences were obtained. VMO with adult as model yielded stronger effects than VMO

with peer as model and video self-modeling. Differential effects based on participant

characteristics and targeted skill were not analyzed, although a qualitative review

indicated variable results across skills. Additionally, differential effects based on the inclusion of additional intervention package components was not addressed.

The above noted reviews (Bellini & Akullian, 2007; Delano, 2007; McCoy & Hermansen, 2007; Shukla-Mehta et al., 2010) clearly indicate VMO as an advantageous intervention, however, contextual specificity, particularly as it relates to participant characteristics, intervention package components, and targeted outcomes is not addressed. Although Chapter 2 of this study disaggregated the studies of VMO based on implementation variables, finding VMO with reinforcement to be the most effective, disaggregation of the VMO studies based on disabilities and implementation variables has not occurred. Knowledge of additional components that may increase the effectiveness of VMO with known participant characteristics, such as primary disability, would assist with practical implementation. Despite the indication of the effectiveness for VMO for participants with ASD, limited information exists regarding the suitability for other populations such as those with developmental disabilities. Additionally, factors related to participant characteristics, such as age or gender that may influence the magnitude of change are not addressed. Further, information regarding for which outcome variables VMO may be most effective is not currently available. Without information regarding these contextual factors, conclusive decisions regarding the evidence-base of video modeling is not possible (Horner et al., 2005). Further, decisions about for whom VMO produces improvements and for what specified skills cannot be made, limiting the practical utility of the practice (McDonald et al., 2006).

The current study focuses on a quantitative meta-analysis of those studies within the single-case research literature utilizing VMO as the independent variable. Meta-analytic procedures can assist with filling current gaps in the literature by aggregating individual single-case studies to derive specific and practical information regarding the conditions under which the practice will provide the most beneficial results in real-world settings (Scruggs & Mastropieri, 1998). The goal of the current meta-analysis was to determine if participant characteristics, intervention components by participant characteristics, and targeted outcome moderated the effectiveness of VMO. In particular, the study focuses on the following questions: (a) Do participant characteristics (age, gender, and diagnostic category) moderate the effectiveness of VMO? (b) Do the implementation components moderate effects when participant diagnosis is considered? (c) Does the targeted outcome moderate the magnitude of change that occurs with the implementation of VMO?

Method

Study Identification

Search method. Potential studies for inclusion in the meta-analysis were identified by way of electronic searches of the ERIC, PsycINFO, and Education Full Text databases. The search was restricted to only peer-reviewed studies and included the following search terms: *modeling* or *observational learning;* (*disability, autism, ADHD, attention deficit hyperactivity disorder, behavior disorder*, *developmental delay,* or *mental*

retardation; and *video* or *videotape*. Use of these search terms yielded 182 potential manuscripts.

     Inclusion criteria. After the search, each potential study was reviewed to determine whether they met the following eight predetermined inclusion criteria: (a) implemented video based intervention using other as model as the independent variable; (b) published in English; (c) appeared in a peer–reviewed journal; (d) focused on communication, social, academic, behavior, or self-help skills as the dependent variable; (e) used a minimum of one participant with a disability; (f) used a single-case research design ; (g) demonstrated experimental control through three or more phase changes; and (h) reported scores with time sequence data available (i.e., readable line graph).

     Inter-rater agreement. The first author and a doctoral student in special education reviewed each study to ensure all identified studies that met the established criteria were systematically included. All disagreements were evaluated by a third rater, also a doctoral student in special education. The inclusion decision made by at least two of the evaluators was the final decision. This process resulted in the inclusion of 41 studies which met all 8 criteria. The reference sections for each of the 41 studies were then searched to identify any additional studies that might have been omitted from the initial search. One additional study was identified and included through this review, yielding a total of 42 single-case manuscripts.

Extraction of Descriptive Information

The included studies were carefully reviewed to extract targeted outcomes, participants' age and disability. Potential moderators were also coded for each study in this review including: (a) participant characteristics, (b) implementation variables plus primary disability, (c) and targeted outcomes.

Participant characteristics. The participant characteristics coded from each study included *age*, *gender*, *primary disability*, and *comorbid disability*. The *age* variable included four levels: preschool (2-5), elementary (6-10), secondary (11-17), and postsecondary (17 and older). Male and female were the dichotomous levels for *gender*. All of the participants were diagnosed with either autism spectrum disorder (ASD) or a developmental disability (DD) thus these were the two levels coded for the *primary disability* variable.

Implementation variable plus primary disability. Coding for this variable was twofold. First*, implementation variable* was coded based on the following three levels: (a) VMO alone, (b) VMO with reinforcement, and (c) VMO as part of a package. Consistent with Chapter 2, VMO alone was the code given to those studies that implemented only VMO as the intervention. VMO with reinforcement was utilized for those studies that only added reinforcement to the VMO intervention protocol. VMO as part of a package included other intervention components (e.g., error correction procedure and/or prompting). Following this, the *implementation variable* codes were then combined with the participant disability code yielding the levels for the *implementation variable plus*

primary disability category. For instance, participants with ASD would include the following levels: ASD alone, ASD with reinforcement, and ASD with package.

Outcome variables. The *outcome* variable was coded based on 5 levels: (a) socio-communication (e.g., language, initiation, reciprocity); (b) play (e.g., functional use of toys, imitation of play script, parallel play); (c) adaptive behavior (e.g., attention, on-task); (d) academic skills (e.g., math, writing, reading); and (e) independent living skills (e.g., making food, brushing teeth).

Effect size and replication analysis

The robust improvement rate difference (IRD; Parker, & Vannest, 2011) was calculated for each of the studies via an evaluation of the line graphs. IRD, known as "risk difference" in other fields of study (Parker, Vannest, & Brown, 2009; Sacket, Richardson, Rosenbert, & Haynes, 1997), quantifies the amount of change that occurs between the contrasted phases of single-case design studies (Parker et al., 2009). IRD has many advantages over other more commonly utilized nonoverlap techniques, such as PND, as its use is not limited to data that meets particular assumptions, allows for calculation of confidence intervals, and is highly correlated with Phi (Parker et al., 2009).

IRD is calculated by identifying the fewest number of data points that would need removal in order to eliminate overlapping data points between each contrasted phase. For example, if the goal is to increase a behavior from baseline during the intervention phase, any data point in the baseline phase equal or greater to an intervention data point is identified as overlapping. A data point in the intervention phase that is equal or lower than one or more baseline data points would be considered overlapping and "not improved."

Complete instructions for calculating IRD are available from Parker et al. (2009); robust

IRD calculation is described by Parker, Vannest, and Davis (2011). IRD effect sizes are

interpreted as small if below .5, moderate if the range is .5 to .70 and large if above .70

(Parker et al., 2009).

In this study, the procedures outlined in Parker et al. (2011) for calculating robust

IRD improved to not improved ratios were followed. The obtained ratios for each study

were calculated and then analyzed utilizing the "risk difference" module of the Number

Cruncher Statistical Software (NCSS: Hintze, 2002). The analysis weights each effect size

(ES) based on the inverse of the standard error and then combines the ES based on

identified moderators, yielding overall effect sizes, confidence intervals, and related

graphics (Parker, Vannest, & Brown, 2009; Parker et al., 2011). Statistically significant

differences (p=.05) were determined by the non-overlap of the upper and lower limits of

the robust IRDs at the 83.4% confidence intervals (Payton, Greenstone, & Schenker,

2003; Schenker & Gentleman, 2001). Utilization of the 83.4% CI allows for a visual

analysis of statistical significance ($p = .05$) and non-overlap is the statistical equivalent of

the student T-test at a 95% CI ($p = .05$).

Interrater Reliability

Moderators. To confirm accurate coding, 31% of the studies were coded by two

raters for each of the identified moderators. Interrater agreement was determined by a

simple percent agreement (agreements/[agreements + disagreements] x 100) calculation.

Interrater agreement was 100% across all moderators.

Robust IRD matrices. From the 42 studies a total of 171 robust IRD matrices were calculated. Of these, 90% were individually computed by the author and another rater to ensure reliable calculations. Interrater agreement was determined by a simple percent agreement (agreements/[agreements + disagreements] x 100) calculation. The overall agreement was 94%. All disagreements were then discussed between the two raters and the IRDs were computed again until 100% agreement was achieved.

Results

Descriptive Summary

Appendix B provides a descriptive summary for each of the 42 studies.

Participant gender and age. Of the 126 participants included in the studies, gender and age were specified for only 121. Of the 121 participants for which gender and age were specified, 84% were males. The participants were parceled into four *age* categories as previously described. Twenty-one participants (17.4%) were in the preschool category whereas 53 (43.8%) were in the elementary category. The secondary category was comprised of a total of 21 participants(17.4%) and 26 participants (21.4%) were classified as postsecondary.

Participant diagnoses. ASD was the primary diagnosis for 106 (84%) of the participants with the remaining 20 (16%), of which 19 were in the secondary or postsecondary age category, diagnosed with developmental disability. Of the participants with ASD, 27 (25%) had a comorbid diagnosis including developmental disability, speech impairment, epilepsy, learning disability, Tourette's, and visual impairment. Four

participants (20%) identified with a developmental disability as the primary diagnosis had a comorbid diagnosis that included either emotional behavior disorder or mood disorder.

  Targeted behavioral outcomes. The majority of studies addressed independent living skills (43%) and socio-communicative skills (33%). Specific independent skills that were targeted included shopping, meal preparation, job performance, clothing care, cooking, job skills, setting the table, doing laundry, housekeeping, and self-care. Socio-communicative skills included initiating, responding, perspective taking, answering, and nonverbal responding. Play was also addressed in 19% of the studies including reciprocal play, imitation of play, and imitation of actions with toys. It is important to note that all included studies with socio-communicative and play skills as the targeted outcomes included only participants with ASD as the primary disability. Academic and behavior skills were each targeted for 2% of the studies.

Effect Size and Replication Analyses

  A total of 171 IRD effect sizes were calculated across the 42 studies and 126 participants. The results yielded an overall IRD effect size of .82 with an 83.4% confidence interval of [.81, .83]. The small confidence interval indicates the precise nature of the obtained effect size, resulting in an 83.4% chance that the true IRD effect size of VMO is between .81 and .83. This obtained effect size and narrow confidence interval suggests a high magnitude of change in the targeted outcomes.

  Participant characteristics. The included studies were analyzed to determine if participant characteristics such as *age*, *gender*, and *primary disability* as variables yielded

differential effects.   Analysis by age with the 121 participants for which age was

specified for 41 of the studies yielded a total of 166 IRDs.

Combined IRDs by age, as well as the corresponding 83.4% confidence intervals

are displayed in Figure 4. Results indicate large effect sizes across all age groups, with

effect sizes ranging from .71 (postsecondary) to .86 (elementary). The non-overlap of the

CI for elementary (IRD = .86) when compared to the other age groups indicates a

statistically significant difference (p = .05) from the other age categories. That is,

elementary aged participants demonstrated the greatest magnitude of change. No

statistically significant differences (p = .05) were detected between the magnitude of

change between any of the other age categories or when gender, male (.82, CI[.83, .81])

and female (.84, CI [.80, .87], was analyzed as a moderator.

To address the research question regarding differential effects due to diagnosis, an

analysis of the *primary disability* variable was conducted. The difference between the

large IRD effect size for ASD (.83) was statistically significant (p = .05) when compared
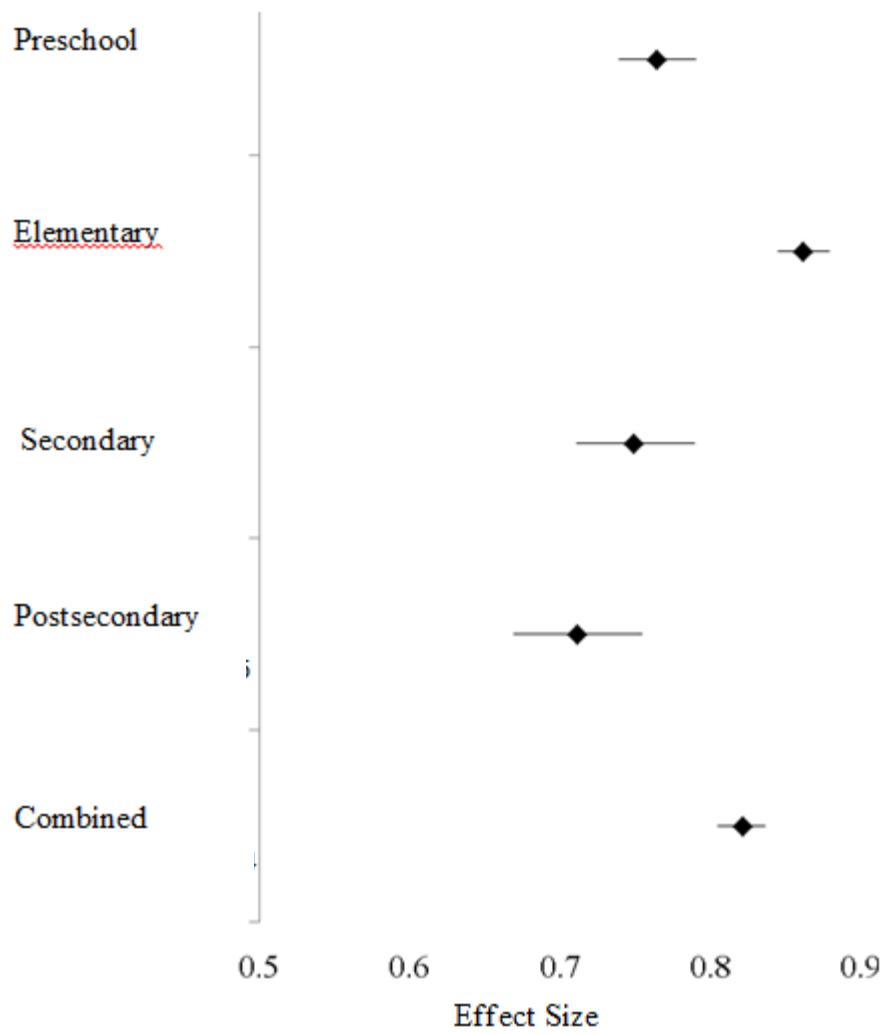
Figure 4. Forest plot of IRD and 83.4% CIs by age.

with the moderate IRD effect size obtained for developmental disability (.68).
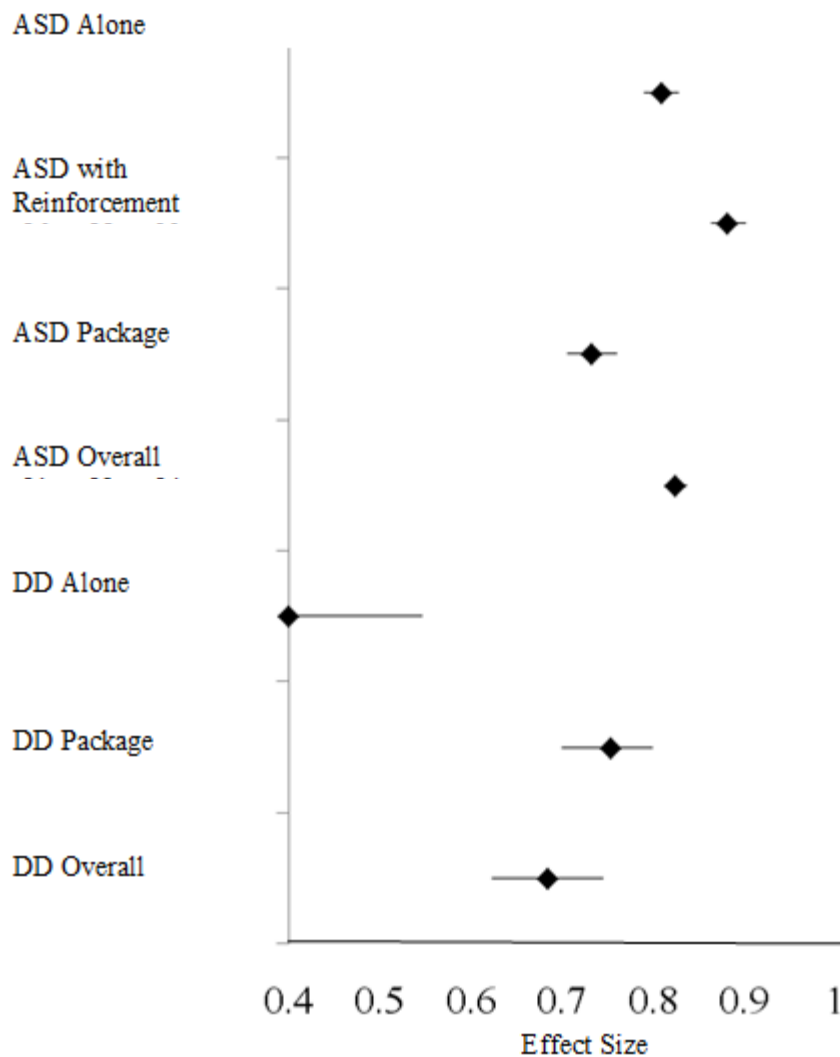
Figure 5. Forest plot of IRDs and 83.4% CIs by levels of primary disability plus implementation variable.

The studies were further analyzed to determine if the *implementation* variable for VMO impacted the effectiveness based on the *disability* variable. Statistically significant differences ($p = .05$) were obtained (See Figure 5) both within and across disabilities. Large effect sizes were obtained across the studies that included participants with ASD as the primary diagnosis regardless of implementation protocol. However, statistically

significant differences (p = .05) were noted for ASD with reinforcement when compared

to ASD with VMO alone and ASD with VMO as part of a package, which yielded the

greatest magnitude of change (IRD = .88). With an overall IRD effect size of .73 for the

use of VMO as part of a package intervention for individuals with ASD, results were

statistically smaller (p = .05) when compared to VMO alone and with reinforcement for

individuals with ASD.

When VMO is utilized alone for individuals with developmental disability , a

minimal effect (IRD = .40) was noted with a statistically significant (p = .05) smaller

effect than that obtained when VMO is utilized as part of a package for individuals with

developmental disability. Only one study (Conyer et al., 2004), which yielded 3 IRDs,

implemented VMO with reinforcement for individuals with developmental disability;

thus, obtaining a valid overall effect size for this variable was not possible. Comparing the

results of the implementation variable across disabilities, statistically significant

differences (p = .05) are noted when comparing VMO alone and with reinforcement for

individuals with ASD across all *implementation* variables of VMO with individuals with

developmental disability .

Targeted outcomes. The included studies were further analyzed to ascertain the

presence of differential effects based on targeted outcomes. Because only one study

targeted each academic and behavior skills, these were not included in the analysis. Figure

6 displays the forest plot with targeted outcomes as moderator. As is visually

demonstrated by the non-overlap of the CIs for play (IRD = .90) with the CIs for the
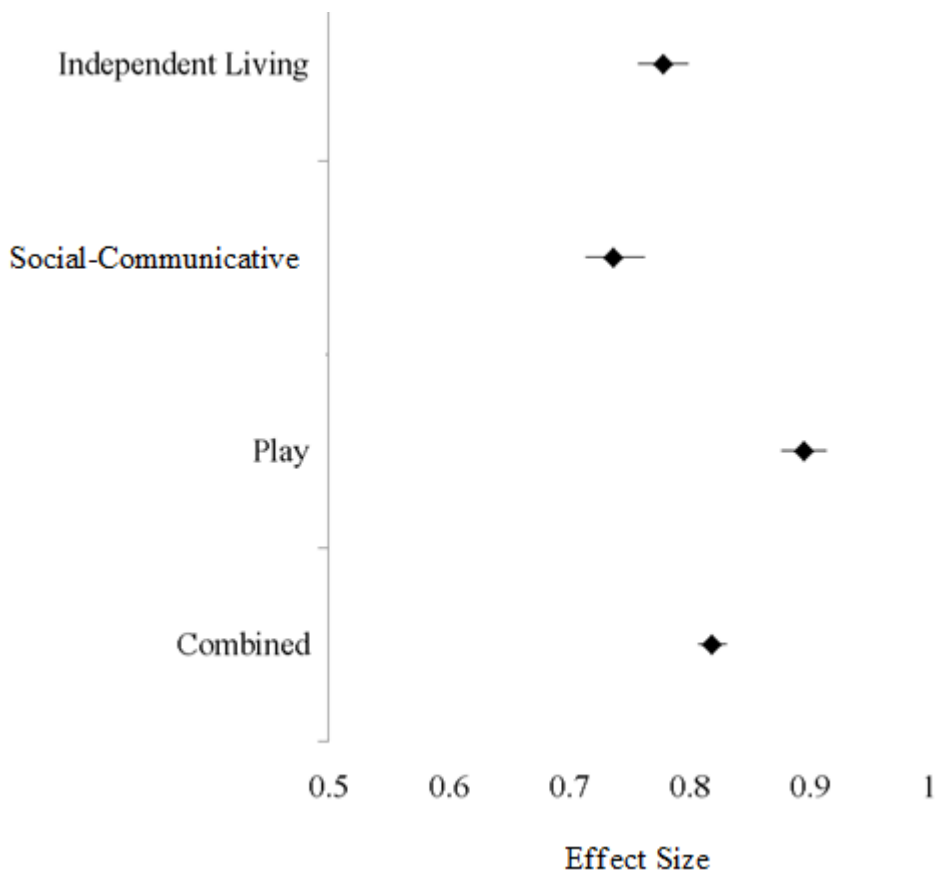
Figure 6. Forest plot of IRDs and 83.4% CIs by targeted outcome.

other two skills, statistically significant differences (*p* = .05) are evident. However, large

effect sizes for both independent living (IRD =.78) and social-communicative (IRD = .74)

skills were also obtained.

Discussion

This meta-analysis investigated factors, specifically, (a) participant characteristics,

(b) intervention components plus primary diagnosis, and (c) and targeted outcomes, that

moderate the effectiveness of VMO interventions with individuals with disabilities. Although previous literature has reviewed VMO in conjunction with other VBMI (Ayres & Langone, 2005; Baker, Lang, & O'Reilly, 2009; Bellini & Akullian, 2007; Delano, 2007; McCoy & Hermansen, 2007; Shukla-Mehta et al., 2010), none have specifically analyzed the effects of VMO. Additionally, information regarding for whom does VMO produce the greatest change, with which implementation variables, and for which targeted outcomes are these gains most likely to be realized has been a gap in the literature. Results of this aggregation of VMO research address these gaps and illuminate areas to explore in future research.

The first research question focused on determining whether the levels of participant characteristic variables, (i.e., age, gender, primary disability, and comorbid disability) moderate the effectiveness of VMO. Results indicated moderate to large effect sizes (Parker et al., 2009) across all levels of each participant characteristic variable. Findings indicate age and diagnosis do moderate the potency of VMO, although gender does not. Regarding age, results indicated that VMO is most effective for elementary aged individuals (ages 6-10), however, large effects sizes were obtained across all age groups. One reason for the differential effects may be attributed to the fact that only one participant with developmental disability was elementary age and the remainder were secondary and/or postsecondary. The lower effect sizes obtained for participants with developmental disabilities likely attributed to the lower omnibus IRDs obtained for the secondary and postsecondary levels. Differential effects based on age have not been

previously explored and these findings indicate VMO to be an effective intervention across age categories.

The second research question focused on the moderating effect of participants' diagnoses (i.e., ASD and developmental disabilities). In the research literature, the effectiveness of VMO with individuals with developmental disabilities has not been previously explored although it has frequently been explored for participants with ASD (Bellini & Akullian, 2007; Shukla-Mehta et al., 2010; McCoy & Hermansen, 2007). Results indicate that diagnosis does moderate the effectiveness of VMO as the large effect size obtained for participants with ASD was statistically significant when compared to the moderate effect size obtained for individuals with developmental disabilities. The large effect sizes obtained for participants with ASD is consistent with Bellini and Akullian's (2007) meta-analysis which found VMO to be an effective intervention for individuals with ASD. Further, although findings indicate VMO to not be as effective with participants with developmental disabilities, this must be viewed with caution given the small sample size.

The third research question focused on determining whether a combination of disability and implementation protocol further moderated the magnitude of change on the targeted outcome. Results indicated that the primary disability combined with implementation components does moderate the magnitude of change that occurs with VMO. The largest effects were obtained when VMO with reinforcement was utilized for participants with ASD, with no statistically significant differences found between VMO delivered alone or as part of a package. This is consistent with results obtained in Chapter

2 and inconsistent with the suggestion of Shukla-Mehta et al. (2010) that the effects of

VMO with individuals with ASD is enhanced with the addition of other interventions

such as prompts and error correction procedures. One plausible explanation for this is that

additional package components reintroduces the need to interact and/or focus on others, a

task particularly challenging for individuals with ASD (Bellini & Akullian, 2007; Rao,

Beidel, & Murray, 2008), distracting them from the video model. A small effect size was

obtained for VMO utilized alone for participants with developmental disabilities whereas

a large effect size was obtained for VMO as part of a package for participants with

developmental disabilities. This analysis calls into question the moderate effect size that

was obtained when analyzing the overall use of VMO for individuals with DD, suggesting

that the results are more indicative of the effectiveness of other included interventions

rather than the VMO. Given the need for additional components for participants with

developmental disabilities combined with the results of this meta-analysis, it appears that

VMO is not an effective intervention for individuals with developmental disabilities

without the inclusion of other components such as prompting and error correction

procedures. This may be due to the complexity of the skills being taught, given that the

analysis was limited to older participants, perhaps requiring more support for acquisition.

VMO may altogether be an inefficient treatment for individuals with DD, however given

the small number of participants with developmental disabilities for whom VMO has been

investigated as well as the truncated age range, definitive statements cannot be made.

Regarding individuals with ASD, VMO will likely yield significant improvements in

targeted skills; however, the addition of reinforcement may enhance the effectiveness of the intervention.

In addition to participant characteristics, this study analyzed potential moderating effects of targeted outcomes on the impact of VMO. The included studies allowed for analysis of social-communicative, play, and independent living as variables for targeted outcomes. Results indicated that VMO is highly effective for all targeted outcomes, however, the impact on play skills was significantly greater than for the other included skills. These results are likely due to two issues. First, only the studies targeting independent living skills included participants with developmental disabilities (See Appendix B). Given that previous analysis demonstrated disability moderates the effectiveness of VMO, it is likely that the overall effect size for independent living skills was moderated by the lower effects sizes of participants with developmental disabilities. The studies targeting play and socio-communicative skills included only participants with ASD. Given this, VMO is likely more effective for improving play skills given the complexity of socio-communicative skill acquisition for individuals with ASD (Reichow & Volkmar, 2010). Additionally, socio-communicative skill deficits are the main symptoms associated with ASD (American Psychiatric Association [DSM-IV-TR], 2000) and, likely the most challenging to improve. Nevertheless, results indicate VMO to be a highly effective intervention across both play and socio-communicative skills for individuals with ASD.

Limitations

   This meta-analysis of single-case research utilizing VMO has some limitations. The primary limitation is the small number of participants with developmental disabilities, limiting the ability to make conclusive statements regarding the utility of VMO with this population. Also, conclusions regarding implementation protocol based on primary diagnosis must be delivered with caution as the limited number of studies included in each category, when disaggregating studies based on disability and implementation protocol, precludes definitive suppositions. Additionally, given the limited effectiveness of VMO alone with participants with developmental disabilities, inclusion of these participants with those with ASD may have confounded some of the results. For instance, VMO may be as effective for independent living skills as it is for play skills for participants with ASD, however, the inclusion of the lower IRDs for participants with developmental disabilities lowered the overall effect size for independent living skills. Another limitation of this study is the lack of comparison of combinations of multiple variables (e.g., age, targeted outcome, and implementation components) to determine which variables enhance or confound results when combined. Such analyses, as would typically be done with larger group studies, would allow for more precise decision-making when practitioners are attempting to determine whether or not VMO is the most appropriate intervention for a given situation. Also, the study was limited to published, peer-reviewed studies, which leads to exclusion of unpublished data, which may include additional studies indicating that VMO did not yield positive results (Borenstein, Hedges, Higgins, & Rothstein (2009).

Implications for Future Research

Future research should focus on the use of VMO with individuals with disabilities other than ASD and developmental disabilities. Given the efficiency with which VMO can be delivered, it is important to analyze the effectiveness across a variety of participants with varying disabilities, including high incidence disabilities such as learning disability and emotional behavior disorders. Given that the results of this study indicate VMO to be an effective intervention for participants with ASD and developmental disabilities, there is a potential that VMO could serve as an appropriate intervention for individuals with high incidence disabilities as well, enhancing inclusive instruction. Video self-modeling has previously been investigated to increase the on-task behaviors for 3 elementary aged students with ADHD (Clare, Jensen, Kehle & Bray, 2000) and the reading fluency for two first graders with learning disability (Hitchcock et al., 2003). Given that VMO with an adult has been found to be significantly more effective than video self-modeling (Chapter 2) and requires less effort to implement (Bellini & Akullian, 2007), exploration across participants with a range of disabilities is warranted.

In the same vein, findings of this meta-analysis indicate there is insufficient research addressing challenging behavior and academic skills. Given evidence of its effectiveness for these areas, VMO would be a highly portable and efficient means for delivering interventions to address these targeted outcomes. Future research targeting participants with high incidence disabilities across a variety of targeted outcomes is

warranted in an effort to determine for which targeted outcomes VMO is most successful and for what populations.

Further, more research assessing the impact of VMO alone with individuals with developmental disabilities is also warranted as the current number of peer-reviewed studies is insufficient for decisions regarding the evidence base of VMO with this population. Furthermore, research addressing the implementation of VMO with individuals with developmental disabilities across a variety of targeted outcomes is also a gap in the current literature base. This current meta-analysis indicates that VMO alone, and thus VMO, is only minimally effective for independent living skills for individuals with developmental disabilities. It is plausible, however, that VMO may be more effective for other targeted outcomes, such as academic and social-communicative skill, for individuals with developmental disabilities. Additionally, component analysis research, comparing VMO alone with VMO as part of a package when utilized with individuals with developmental disabilities, will provide information that is practically useful in terms of the necessary components to include to maximize the effectiveness of the intervention.

In summary, results of this meta-analysis indicate VMO to be a highly effective intervention for individuals with ASD and are consistent with other studies that have designated VMO as an evidence-based practice (Bellini & Akullian, 2007). However, this meta-analysis broadens the practical utility of the evidence base as it delineates the effectiveness of implementation protocol and extends beyond individuals with ASD. When implemented alone or with reinforcement across a variety of skills, excluding academic and behavior, VMO has strong effects with individuals with ASD. Regarding

the use of VMO with participants with developmental disabilities, this study indicates

VMO is not an effective intervention for independent living skills unless included as part

of an intervention package. Although results also indicate strong effects when

implemented as part of an intervention package for participants with developmental

disabilities, this does not speak to the utility of VMO. The results indicate an intervention

package that includes VMO is effective, yet information regarding which components are

necessary to achieve desired effects is not indicated. The lack of effects obtained for

VMO alone compared to the high effects of VMO as part of a package, suggests that

VMO may not be an effective intervention for participants with developmental

disabilities. Additional research is needed to identify VMO as an evidence-based practice

for individuals with other disabilities.

CHAPTER IV

CONCLUSION: IMPLICATIONS FOR PRACTICE AND FUTURE RESEARCH

VBM maximizes the benefits of observational learning, more specifically imitation, providing a more efficient and flexible method than the more primitive in vivo procedure (Biederman & Freedman, 2007; McCoy & Hermansen, 2007). VBM has been implemented to address a variety of targeted outcomes across participants with varying abilities and diagnoses (Bellini & Akullian, 2007; Nikopoulos & Nikopoulos-Smyrni, 2008; Rayner, Denholm, and Sigafoos, 2009). However, the current evidence base lacks implementation and contextual specificity particularly regarding which VBM procedures are beneficial for which participants and targeted outcomes (Rayner et al., 2009; Shukla-Mehta, Miller, and Callahan, 2010). This dissertation addresses these gaps in the VBM research and provides information that will expand the practical value of this intervention.

In summary, both studies indicate differential effects are present based on procedural implementation. The types of VBM implemented, VSM and VMO, appear to be nearly equal in potency, both with high effects as was also indicated by Bellini and Akullian (2007) and Cihak and Schrader (2008). However, when disaggregated based on the type of model utilized, VMO with adult as model has higher potency than both VMO with peer and VSM. Whether or not the model is known or unknown for VMO does not appear to moderate the effectiveness. Furthermore, no statistical differences resulted when VSM is produced utilizing either positive self-review or feedforward. This extends the literature base as the analysis indicates that all types of VBM, regardless of type of model are effective; however, adult as model is the most effective based on the current literature.

The current VMO research is limited to participants with ASD and developmental disability; thus, application of these findings to participants with other disabilities (i.e., learning disability, emotional disturbance) is limited.

The effectiveness of variations in implementation protocol, with levels of VBM alone, with reinforcement, or as part of a package, was also analyzed. Results indicate VBM with reinforcement to be the most efficacious with significantly smaller effects when VBM was implemented as part of a package, which is consistent with results obtained for VMO when analyzed separately from the VSM studies. However, results indicate VSM is equally efficacious when implemented alone or as part of a package, whereas VSM with reinforcement alone yielded a significantly smaller effect size. Thus, the previous finding of Shukla-Mehta et al. (2010) that the inclusion of reinforcement with VBM will likely enhance effectiveness holds true for VMO but not VSM.

As VMO was found to be the most efficacious in study one (Chapter 2), study two (Chapter 3) further analyzed moderators as they relate to participant characteristics and targeted outcomes. Results indicate VMO to be highly efficacious across targeted outcomes (i.e. independent living, socio-communicative, and play skills), particularly play skills. Although the included studies were limited to participants with ASD and developmental disabilities, it was found to be highly effective across age groups, although more effective for participants with ASD than developmental disabilities. Moderator potency did not vary based on the presence or absence of a comorbid disorder. Findings suggest VMO results in a greater magnitude of change when reinforcement is included as part of the intervention protocol for participants with ASD. However, diverging from

previous studies (Shukla-Mehta et al., 2010), additional components do not increase potency beyond VMO delivered alone for participants with ASD. On the contrary VMO delivered alone appears to be less effective than VMO as part of a package for individuals with developmental disabilities suggesting VMO may not be an efficacious intervention for participants with developmental disabilities.

Implications for Practice

These findings have several implications for practitioners including those in educational and clinical settings. First, clearly both VMO and VSM are highly effective for producing change in targeted behaviors across a range of skills for participants with disabilities. Given the simplicity of producing a VMO intervention when compared to the challenges of producing a VSM intervention (Bellini & Akullian, 2007; Rayner et al., 2009), which requires editing, VMO with an adult as model is the most efficient method and will likely have the desired impact on targeted outcomes. Findings indicate that whether or not the target person is familiar with the model will unlikely have an impact on the effectiveness.

Furthermore, if VSM is the desired intervention, study one of this meta-analysis indicates no difference in potency when feedforward versus positive self-review is implemented. Again, feedforward is the more practical of the two as it merely requires prompting the individual so that he/she can perform each step of the skill accurately and then editing out the prompts. Positive self-review on the other hand, requires recording

the target person, potentially for several hours, in a natural setting until there is enough footage of accurate skill performance to produce the video model.

In terms of moderating effects due to variations in participant characteristics, this study only investigated the impacts for VMO. VMO is clearly an evidenced base practice for participants with ASD across all age ranges for improvement of play, independent living, and socio-communicative skills. The addition of positive reinforcement in tandem with VMO for participants with ASD is recommended. For participants with developmental disabilities, the second study indicates VMO has been implemented primarily with secondary and postsecondary participants and is most effective when implemented as part of a package including other procedures such as performance feedback, error correction, and prompting. Although the results indicate minimal change in behavior when VMO is implemented alone for older participants with developmental disabilities, it should not necessarily be ruled out as a viable option. Consideration should be given to the social impact of the desired skill, and if small changes in behavior would have socially significant ramifications. Given the simplicity of VMO as an intervention, the social validity of small changes with minimal resource expenditure may outweigh greater changes that might occur with more costly interventions. One possible means for implementing VMO to maximize both the efficiency and benefit is to begin with VMO alone. If the skill improves but does not reach desired levels or criterion, supplementing the VMO with additional components such as reinforcement and error correction procedures is recommended. There is insufficient evidence for the use of VMO to enhance academic and adaptive behavior skills.

Limitations

Aside from the limitations noted at the end of each individual study, some additional limitations exist. First, both studies exclude publications that were not peer reviewed such as dissertations. Such studies could potentially provide further information regarding populations for whom VBM is effective as well as indicating contextual variables and participant characteristics that do not benefit from VBM intervention.

Another limitation of these studies is that analyses of participant characteristics and targeted outcomes that moderate the effectiveness of VSM were not included. This current work provides information that VSM is a highly effective intervention for evoking change in skills, however contextual specificity such as for whom and under what circumstances was not addressed. Although this was a purposeful exclusion due to the known benefits of VMO, information regarding the potency of various moderators as they relate to VSM would further guide the process of choosing the most appropriate VBM intervention.

Additionally, neither study analyzed dosage including length of video and typical number of sessions required to achieve desired outcomes. Such information is important as the video exemplars in the current evidence base vary in length as well as the amount of time that was necessary before criterion was achieved.

Future Research

Future research that expands upon these studies and addresses the limitations will further enhance the practical implementation of VBM. First, an evaluation of potential

moderators for VSM would assist in providing information regarding when VSM might be a more appropriate intervention despite the increased requirement of resources such as time and personnel. VSM may be more effective and easier to implement for participants and targeted skills for which VMO is not effective, such as participants with high incidence disabilities, as well as targeted skills such as academic and adaptive behavior.

Additionally, research that evaluates the required dosage, length of video, and number of viewings required to achieve maximum effect is an area to be addressed. Research that provides specific information regarding appropriate dosage based on type of VBM, participant characteristics, and targeted outcomes would further assist practitioners with intervention decision-making. Potentially, VBM techniques that require more initial resources could require shorter videos and result in more rapid acquisition for certain participants.

As was evident in the second study, VMO for older participants with developmental disabilities is most effective when implemented as part of a package. More specific analysis regarding which specific intervention components, such as error correction or prompting, facilitate more potent results is necessary. Additionally, research that analyzes if these interventions are equally effective alone or in conjunction with VBM is necessary. For example if research indicates error correction procedures are more effective in improving targeted outcomes for participants with developmental disabilities than VMO with error correction procedures, implementation of VMO would not be warranted. Moreover, research that compares the effects of VBM in addressing specific targeted outcomes (i.e., independent living, socio-communicative) compared to other

interventions is necessary to assist practitioners in choosing the most appropriate intervention. For instance, research that analyzes differential effects of social stories, peer-mediated training, and VBM for improving the socio-communicative skills of individuals with ASD would further the field.

More broadly, these two studies indicate the importance of disaggregating the evidence base for specific interventions based on a variety of potential moderators. An overall assessment of VBM indicates that this intervention, whether implemented as VMO or VSM, is effective. Further analysis would not have provided the information that adult as model works best with reinforcement for participants with ASD across all ages to improve play, independent living, and socio-communicative skills. Furthermore, information that VMO may require the addition of other intervention techniques to obtain desired results for participants with developmental disabilities would not be available. Information regarding the potency of an intervention in light of given contextual factors is necessary for practitioners to make informed decisions regarding the most appropriate interventions to use when assisting participants with disabilities given the heterogeneity across and within disability categories (Montgomery, 2006). Choosing appropriate interventions for individuals with disabilities requires practitioners to select the most appropriate intervention that will maximize outcomes for their particular students (Cook, Tankersley, & Landrum, 2009). Careful analysis of the evidence base can assist practitioners in choosing effective interventions with a high degree of expectancy that the intervention will work for their particular contexts.

REFERENCES

"*" Indicates studies used in the meta-analysis

*Alcantara, P. R. (1994). Effects of videotape instructional package on purchasing skills of children with autism. *Exceptional Children, 61*, 40-55.

*Allen, K. D., Wallace, D. P., Greene, D. J., Bowen, S. L., & Burke, R. V. (2010). Community-based vocational instruction using videotaped modeling for young adults with autism spectrum disorders performing in air-inflated mascots. *Focus on Autism and Other Developmental Disabilities, 25*, 186-192. doi:10.1177/1088357610377318

*Allen, K. D., Wallace, D. P., Renes, D., Bowen, S. L., & Burke, R. V. (2010). Use of video modeling to teach vocational skills to adolescents and young adults with autism spectrum disorders. *Education and Treatment of Children, 33*, 339-349.

American Psychiatric Association [APA]. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text rev.). Washington, DC: Author.

*Apple, A. L., Billingsley, F., Schwartz, I. S., & Carr, E. G. (2005). Effects of video modeling alone and with self-management on compliment-giving behaviors of children with high-functioning ASD. *Journal of Positive Behavior Interventions, 7*, 33-46. doi:10.1177/10983007050070010401

Ayres, K. M., & Langone, J. (2005). Intervention and instruction with video for students with autism: A review of the literature. *Education and Training in Developmental Disabilities, 40*, 183-196.

*Ayres, K. M., & Langone, J. (2007). A comparison of video modeling perspectives for

students with autism. *Journal of Special Education Technology*, *22*, 15-30.

Ayres, K. M., Maguire, A., & McClimon, D. (2009). Acquisition and generalization of

chained tasks taught with computer based video instruction to children with autism.

Education and Training in Developmental Disabilities, 44, 493-508.

Baker, S. D., Lang, R., & O'Reilly, M. (2009). Review of video modeling with students

with emotional and behavioral disorders. *Education and Treatment of Children, 32*,

403-420. doi:10.1353/etc.0.0065

Bandura, A. (1969). Social learning of moral judgments. *Journal of Personality and

Social Psychology*, *11*, 275-279.

Bellini, S., & Akullian, J. (2007). A meta-analysis of video modeling and video self-

modeling interventions for children and adolescents with autism spectrum disorders.

Exceptional Children, 73, 264-287.

*Bellini, S., & McConnell, L. L. (2010). Strength-based educational programming for

students with autism spectrum disorders: A case for video self-modeling.

Preventing School Failure, *54*, 220-227. doi:10.1080/10459881003742275

*Bidwell, M. A., & Rehfeldt, R. A. (2004). Using video modeling to teach a domestic

skill with an embedded social skill to adults with severe mental retardation.

Behavioral Interventions, 19, 263-274. doi:10.1002/bin.165

Biederman, G. B., & Freedman, B. (2007). Modeling skills, signs and lettering for

children with down syndrome, autism and other severe developmental delays by

video instruction in classroom setting. *Journal of Early and Intensive Behavior Intervention,* 4, 736-743.

Borenstein, M., Hedges, L.V., Higgins, J.P.T., & Rothstein, H.R. (2009). *Introduction to meta-analysis.* West Sussex, United Kingdom: Wiley.

*Boudreau, E., & D'Entremont, B. (2010). Improving the pretend play skills of preschoolers with autism spectrum disorders: The effects of video modeling. *Journal of Developmental and Physical Disabilities*, *22*, 415-431. doi:10.1007/s10882-010-9201-5

Brennan, R. L., & Prediger, D. J. (1981). Coefficient kappa: Some uses, misuses, and alternatives. *Educational and Psychological Measurement, 41*, 687-699. doi:10.1177/001316448104100307

*Buggey, T. (2005). Video self-modeling applications with students with autism spectrum disorder in a small private school setting. *Focus on Autism and Other Developmental Disabilities*, *20*(1), 52-63. doi:10.1177/10883576050200010501

*Buggey, T., Toombs, K., Gardener, P., & Cervetti, M. (1999). Training responding behaviors in students with autism: Using videotaped self-modeling. *Journal of Positive Behavior Interventions*, *1*, 205-214.

*Cannella-Malone, H., Sigafoos, J., O'Reilly, M., De La Cruz, B., Edrisinha, C., & Lancioni, G. E. (2006). Comparing video prompting to video modeling for teaching daily living skills to six adults with developmental disabilities. *Education and Training in Developmental Disabilities*, *41*, 344-356.

Chard, D. J., Ketterlin-Geller, L., Baker, S. K., Doabler, C., & Apicatabutra, C. (2009). Repeated reading interventions for students with learning disabilities: Status of the evidence. *Exceptional Children, 75*, 263-281.

*Charlop, M. H., Dennis, B., Carpenter, M. H., & Greenberg, A. L. (2010). Teaching socially expressive behaviors to children with autism through video modeling. *Education and Treatment of Children*, *33*(3), 371-393.

*Charlop, M. H., Gilmore, L., & Chang, G. T. (2008). Using video modeling to increase variation in the conversation of children with autism. *Journal of Special Education Technology*, *23*, 47-66.

*Charlop, M. H., & Milstein, J. P. (1989). Teaching autistic children conversational speech using video modeling. *Journal of Applied Behavior Analysis*, *22*, 275-285. doi:10.1901/jaba.1989.22-275

*Charlop-Christy, M. H., & Daneshvar, S. (2003). Using video modeling to teach perspective taking to children with autism. *Journal of Positive Behavior Interventions, 5*, 12-21. doi:10.1177/10983007030050010101

*Charlop-Christy, M. H., Le, L., & Freeman, K. A. (2000). A comparison of video modeling with in vivo modeling for teaching children with autism. *Journal of Autism and Developmental Disorders*, *30*, 537-552.

*Cihak, D. F., & Shrader, L. (2008). Does the model matter? comparing video self-modeling and video adult modeling for task acquisition and maintenance by adolescents with autism spectrum disorders. *Journal of Special Education Technology, 23*, 9-20.

*Clare, S. K., Jenson, W. R., Kehle, T. J., & Bray, M. A. (2000). Self-modeling as a treatment for increasing on-task behavior. *Psychology in the Schools, 37*, 517-522.

Cochrane, W. S., & Laux, J. M. (2008). A survey investigating school psychologists' measurement of treatment integrity in school-based interventions and their beliefs about its importance. *Psychology in the Schools, 45*, 499-507.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement, 20*(1), 37-46. doi:10.1177/001316446002000104

Conroy, M. A., Stichter, J. P., Daunic, A., & Haydon, T. (2008). Classroom-based research in the field of emotional and behavioral disorders: Methodological issues and future research directions. *Journal of Special Education, 41*, 209-222. doi:10.1177/0022466907310369

*Conyers, C., Miltenberger, R. G., Peterson, B., Gubin, A., Jurgens, M., Selders, A., & Barenz, R. (2004). An evaluation of in vivo desensitization and video modeling to increase compliance with dental procedures in persons with mental retardation. *Journal of Applied Behavior Analysis*, *37*, 233-238.

Cook, B. G., Tankersley, M., & Landrum, T. J. (2009). Determining evidence-based practices in special education. *Exceptional Children, 75*, 365-383.

Corbett, B. A., & Abdullah, M. (2005). Video modeling: Why does it work for children with autism? *Journal of Early and Intensive Behavior Intervention, 2*, 2-8.

*Coyle, C., & Cole, P. (2004). A videotaped self-modeling and self-monitoring treatment program to decrease off-task behavior in children with autism. *Journal of

Intellectual and Developmental Disability, 29, 3-15.
doi:10.1080/08927020410001662642

*D'Ateno, P., Mangiapanello, K., & Taylor, B. A. (2003). Using video modeling to teach complex play sequences to a preschooler with autism. *Journal of Positive Behavior Interventions, 5*, 5-11. doi:10.1177/10983007030050010801

Delano, M. E. (2007). Video modeling interventions for individuals with autism. *Remedial and Special Education, 28*, 33-42.

Delano, M. (2007). Improving written language performance of adolescents with asperger syndrome. *40*, 345-351.

Dowrick, P. W. (1999). A review of self-modeling and related interventions. *Applied and Preventative Psychology, 8*, 23-39.

Dowrick, P. W., Kim-Rupnow, W. S., & Power, T. J. (2006). Video feedforward for reading. *Journal of Special Education, 39*, 194-207.

*Dowrick, P. W., & Ward, K. M. (1997). Video feedforward in the support of a man with intellectual disability and inappropriate sexual.. *Journal of Intellectual & Developmental Disability, 22*, 147-160.

Ganz, J. B., Earles-Vollrath, T. L., Heath, A. K., Parker, R. I., Rispoli, M. J., & Duran, J. (2012). A meta-analysis of single-case research studies on aided augmentative and alternative communication systems with individuals with autism spectrum disorders. *Journal of Autism and Developmental Disorders, 42*, 60-74. doi: 10.1007/s10803-011-1212-2

Ganz, J. B., Earles-Vollrath, T. L., Mason, R. A., Rispoli, M. J., Heath, A. K., & Parker, R. I. (2011). An aggregate study of single-case research involving aided AAC: Participant characteristics of individuals with autism spectrum disorder. *Research in Autism Spectrum Disorders, 5*, 1500-1509. doi:10.1016/j.rasd.2011.02.011

*Gena, A., Couloura, S., & Kymissis, E. (2005). Modifying the affective behavior of preschoolers with autism using "in-vivo" or video modeling and reinforcement contingencies. *Journal of Autism and Developmental Disorders, 35*, 545-556. doi:/10.1007/s10803-005-0014-9

Gresham, F. M., MacMillan, D. L., Beebe-Frankenberger, M., & Bocian, K. M. (2000). Treatment integrity in learning disabilities intervention research: Do we really know how treatments are implemented? *Learning Disabilities Research & Practice (Lawrence Erlbaum), 15*, 198-205.

*Haring, T. G., Kennedy, C. H., Adams, M. J., & Pitts-Conway, V. (1987). Teaching generalization of purchasing skills across community settings to autistic youth using videotape modeling. *Journal of Applied Behavior Analysis, 20*, 89-96. doi:10.1901/jaba.1987.20-89

*Hepting, N. H., & Goldstein, H. (1996). Requesting by preschoolers with developmental disabilities: Videotaped self-modeling and learning of new linguistic structures. *Topics in Early Childhood Special Education, 16*, 407-427.

Hintze, J. (2002). *NCSS and pass: Number cruncher statistical systems*. Kaysville, UT: NCSS LLC

Hitchcock, C. H., Dowrick, P. W., & Prater, M. A. (2003). Video self-modeling

    intervention in school-based settings: A review. *Remedial and Special Education,*

    24, 36-46. doi:10.1177/074193250302400104

*Hitchcock, C. H., Prater, M. A., & Dowrick, P. W. (2004). Reading comprehension and

    fluency: Examining the effects of tutoring and video self-modeling on first-grade

    students with reading difficulties. *Learning Disability Quarterly, 27*, 89-103.

Holmbeck, G. (1997). Toward terminological, conceptual, and statistical clarity in the

    study of mediators and moderators: Examples from the child-clinical and pediatric

    psychology literature. *Journal of Counseling and Clinical Psychology, 65*, 599-610.

Horner, R. H., Carr, E. C., Halle, J., McGee, G., Odom, S., & Wolery, M. (2005). The

    use of single-subject research to identify evidence-based practice in special

    education. *Exceptional Children, 71*, 165-179.

Individuals with Disabilities Education Improvement Act Regulations, 34 C.F.R. §

    300.35 *et seq.* (2004).

Jitendra, A., Burgess, C., & Gajria, M. (2011). Cognitive strategy instruction for

    improving expository text comprehension of students with learning disabilities: The

    quality of evidence. *Exceptional Children, 77*, 135-159.

Kavale, K. A. (2010). Meta-analysis: A primer. *Exceptionality, 9*, 177-183.

    doi:10.1207/S15327035EX0904_2

*Keen, D., Brannigan, K. L., & Cuskelly, M. (2007). Toilet training for children with

    autism: The effects of video modeling category. *Journal of Developmental and*

    Physical Disabilities, 19, 291-303. doi:10.1007/s10882-007-9044-x

*Kehle, T. J., Clark, E., Jenson, W. R., & Wampold, B. E. (1986). Effectiveness of self-observation with behavior disordered elementary school children. *School Psychology Review, 15*, 289-295.

*Kleeberger, V., & Mirenda, P. (2010). Teaching generalized imitation skills to a preschooler with autism using video modeling. *Journal of Positive Behavior Interventions, 12*, 116-127. doi:10.1177/1098300708329279

Kratochwill, T. R., Hitchcock, J., Horner, R. H., Levin, J.R., Odom, S. L., Rindskopf, D. M., et al. (2010). Single-case designs technical documentation. Retrieved from http://ies.ed.gov/ncee/wwc/pdf/wwc_scd.pdf

Lane, K., Kalberg, J., & Shepcaro, J. (2009). An examination of the evidence base for function-based interventions for students with emotional and/or behavioral disorders attending middle and high schools. *Exceptional Children, 75*, 321-340.

*Lasater, M. W., & Brady, M. P. (1995). Effects of video self-modeling and feedback on task fluency: A home based intervention. *Education and Treatment of Children*, 18, 389 - 407.

*LeBlanc, L. A. (2010). Using video-based interventions with individuals with autism spectrum disorders: Introduction to the special issue. *Education & Treatment of Children, 33*, 333-337. doi:10.1353/etc.0.0099

*Lonnecker, C. (1994). Video self-modeling and cooperative classroom behavior in children with learning and behavior problems: Training and generalization effects. *Behavioral Disorders*, *20*, 24-34.

*MacDonald, R., Clark, M., Garrigan, E., & Vangala, M. (2005). Using video modeling to teach pretend play to children with autism. *Behavioral Interventions, 20*, 225-238. doi:10.1002/bin.197

*MacDonald, R., Sacramone, S., Mansfield, R., Wiltz, K., & Ahearn, W. H. (2009). Using video modeling to teach reciprocal pretend play to children with autism. *Journal of Applied Behavior Analysis, 42*, 43-55.

*Maione, L., & Mirenda, P. (2006). Effects of video modeling and video feedback on peer-directed social language skills of a child with autism. *Journal of Positive Behavior Interventions, 8*, 106-118. doi:10.1177/10983007060080020201

*Marcus, A., & Wilder, D. A. (2009). A comparison of peer video modeling and self video modeling to teach textual responses in children with autism. *Journal of Applied Behavior Analysis, 42*, 335-341. doi:10.1901/jaba.2009.42-335

*Martin, J. E., & And Others. (1992). Effects of picture referencing on PVC chair, love seat, and settee assemblies by students with mental retardation. *Research in Developmental Disabilities, 13*, 267-286.

McCoy, K., & Hermansen, E. (2007). Video modeling for individuals with autism: A review of model types and effects. *Education and Treatment of Children, 30*(4), 183-213.

McDonald, S., Keesler, V. A., Kauffman, N. J., & Schneider, B. (2006). Scaling-up exemplary interventions. *Educational Researcher, 35*(3), 15-24.

McLeskey, J. & Waldron, N.L. (2011). Educational programs for elementary students with learning disabilities: Can they be both effective and inclusive? *Learning Disabilities Research & Practice, 26*, 48-57.

*Mechling, L. C. (2005). The effect of instructor-created video programs to teach students with disabilities: A literature review. *Journal of Special Education Technology, 20*(2), 25-26.

Mechling, L. C., Pridgen, L. S., & Cronin, B. A. (2005). Computer-based video instruction to teach students with intellectual disabilities to verbally respond to questions and make purchases in fast food restaurants. *Education and Training in Developmental Disabilities, 40*, 47-59.

Montgomery, J. K. (2006). Current status of evidence-based practices for low-incidence disabilities: Introduction to the special series. *Communication Disorders Quarterly, 28*, 37-38.

*Murzynski, N. T., & Bourret, J. C. (2007). Combining video modeling and least-to-most prompting for establishing response chains. *Behavioral Interventions, 22*, 147-152. doi:10.1002/bin.224

*Nikopoulos, C. K., Canavan, C., & Nikopoulou-Smyrni, P. (2009). Generalized effects of video modeling on establishing instructional stimulus control in children with autism: Results of a preliminary study. *Journal of Positive Behavior Interventions, 11*, 198-207. doi:10.1177/1098300708325263

*Nikopoulos, C. K., & Keenan, M. (2003). Promoting social initiation children with autism using video modeling. *Behavioral Interventions, 18*, 87-108. doi:10.1002/bin.129

*Nikopoulos, C., & Keenan, M. (2004). Effects of video modeling on social initiations by children with autism. *Journal of Applied Behavior Analysis*, *37*, 93-96.

*Nikopoulos, C., & Keenan, M. (2007). Using video modeling to teach complex social sequences to children with autism. *Journal of Autism and Developmental Disorders, 37*, 678-693. doi:10.1007/s10803-006-0195-x

Nikopoulos, C. K., & Nikopoulos-Smyrni, P. (2008). Teaching complex social skills to children with autism; advances of video modeling. *Journal of Early and Intensive Behavior Intervention, 5*, 30-43.

No Child Left Behind Act of 2001, 20 U.S. C. 70 §6301 *et seq*. (2002)

Odom, S. (2009). The tie that binds: Evidence-based practice, implementation science, and outcomes for children. *Topics in Early Childhood Special Education, 29*, 53-61. doi:10.1177/0271121408329171

Odom, S., Brantlinger, E., Gersten, R., Horner, R., Thompson, B., & Harris, K. S. (2005). Research in special education: Scientific methods and evidence-based practices. *Exceptional Children, 71*, 137-148.

Parker, R. I., Hagan-Burke, S., & Vannest, K. (2007). Percent of all non-overlapping data (PAND): An alternative to PND. *The Journal of Special Education, 40*, 194-204.

Parker, R. I., Vannest, K. J., Davis, J.L. (2011). Effect size in single-case research: A review of nine nonoverlap techniques. *Behavior Modification, 35*, 303-322. doi:10.1177/0145445511399147

Parker, R. I., Vannest, K., & Davis, J. (2011). Inter-rater reliability for multi-categorical scales: Problems and solutions. Manuscript in preparation.

Parker, R. I., Vannest, K. J., & Brown, L. (2009). The improvement rate difference for single-case research. *Exceptional Children*, *75*, 135-150.

Paterson, C. R., & Arco, L. (2007). Using video modeling for generalizing toy play in children with autism. *Behavior Modification*, *31*, 660-681. doi: 10.1177/0145445507301651

Payton, M. E., Greenstone, M. H., & Schenker, N. (2003). Overlapping confidence intervals or standard error intervals: What do they mean in terms of statistical significance? *Journal of Insect Science, 3*(34), 1-6.

Payton, M. E., Miller, A., & Raun, W. R. (2000). Testing statistical hypotheses using standard error bars and confidence intervals. *Communications in Soil Science and Plant Analysis, 31*, 547-552.

*Pigott, H. E., & Gonzales, F. P. (1987). Efficacy of videotape self-modeling in treating an electively mute child. *Journal of Clinical Child Psychology, 16*, 106-110.

*Possell, L. E., Kehle, T. J., Mcloughlin, C. S., & Bray, M. A. (1999). Self-modeling as an intervention to reduce disruptive classroom behavior. *Cognitive and Behavioral Practice, 6*, 99-105.

Rao, P.A., Beidel, D.C., & Murray, M.J.(2008). Social skills interventions for children with Asperger's Syndrome or high-functioning autism: A review and recommendations. *Journal of Autism and Developmental Disorders, 38*, 353-361. doi: 10.1007/s10803-007-0402-4

*Rayner, C. S. (2010). Video-modeling to improve task completion in a child with autism. *Developmental Neurorehabilitation, 13*, 225-230. doi:10.3109/17518421003801489

Rayner, C., Denholm, C., & Sigafoos, J. (2009). Video-based intervention for individuals with autism: Key questions that remain unanswered. *Research in Autism Spectrum Disorders, 3*, 291-303. doi:10.1016/j.rasd.2008.09.001

*Rehfeldt, R. A., Dahman, D., Young, A., Cherry, H., & Davis, P. (2003). Teaching a simple meal preparation skill to adults with moderate and severe mental retardation using video modeling. *Behavioral Interventions, 18*, 209-218. doi:10.1002/bin.139

Reichow, B., & Volkmar, F. R. (2009). Social skill interventions for individuals with autism: Evaluation for evidence-based practices within a best evidence synthesis framework. *Journal of Autism and Developmental Disorders*, *40*, 149-166.

Rosenberg, N. E., Schwartz, I. S., & Davis, C. A. (2010). Evaluating the utility of commercial videotapes for teaching hand washing to children with autism. *Education & Treatment of Children. 33*, 443-455. doi:10.1353/etc.0.0098

Sacket, D. L., Richardson, W. S., Rosenbert, W., & Haynes, R. B. (Eds.). (1997). *Evidence-based medicine: How to practice and teach EBM*. London: Churchill Livingston.

*Sansosti, F. J., & Powell-Smith, K. A. (2008). Using computer-presented social stories and video models to increase the social communication skills of children with high-functioning autism spectrum disorders. *Journal of Positive Behavior Interventions, 10*, 162-178.

*Scattone, D. (2008). Enhancing the conversation skills of a boy with Asperger's disorder through social stories and video modeling. *Journal of Autism and Developmental Disorders*, *38*, 395-400. doi: 10.1007/s10803-007-0392-2

Schenker, N., & Gentleman, J. F. (2001). On judging the significance of differences by examining overlap between confidence intervals. *The American Statistician,* 55, 182-186.

Scruggs, T. E., & Mastropieri, M. A. (1998). Summarizing single-subject research: Issues and applications. *Behavior Modification, 22*, 221-242.

*Sherer, M., Pierce, K. L., Paredes, S., Kisacky, K. L., Ingersoll, B., & Schreibman, L. (2001). Enhancing conversation skills in children with autism via video technology: Which is better, "self" or "other" as a model? *Behavior Modification, 25*, 140-158. doi:10.1177/0145445501251008

Shukla-Mehta, S., Miller, T., & Callahan, K. J. (2010). Evaluating the effectiveness of video instruction on social and communication skills training for children with autism spectrum disorders: A review of the literature. *Focus on Autism and Other Developmental Disabilities, 25*, 23-36. doi: 10.1177/1088357609352901

*Taber-Doughty, T., Patton, S. E., & Brennan, S. (2008). Simultaneous and delayed
video modeling: An examination of system effectiveness and student preferences.
Journal of Special Education Technology, 23, 1-18.

*Taylor, B. A., Levin, L., & Jasper, S. (1999). Increasing play-related statements in
children with autism toward their siblings: Effects of video modeling. *Journal of
Developmental and Physical Disabilities, 11*, 253-264. doi:
10.1023/A:1021800716392

*Van Laarhoven, T., Zurita, L. M., Johnson, J. W., Grider, K. M., & Grider, K. L.
(2009). Comparison of self, other, and subjective video models for teaching daily
living skills to individuals with developmental disabilities. *Education and Training
in Developmental Disabilities, 44*, 509-522.

Vannest, K., Parker, R. I. & Gonan, O. (June 14, 2011). *Single case research: Web-
based calculators for SCR analysis.* Available from singlecaseresearch.org.

Vannest, K. J., Davis, J. L., Davis, C. R., Mason, B. A., & Burke, M. D. (2010).
Effective intervention for behavior with a daily behavior report card: A meta-
analysis. *School Psychology Review, 39*, 654-672.

Vannest, K. J., Harrison, J., Parker, R. I., Harvey, K. T., & Ramsey, L. (2010).
Improvement rate differences of academic interventions for students with emotional
behavioral disorders. *Remedial & Special Education, 32,* 521-534. doi:
10.1177/0741932510362509

Walker, H. (2004). Commentary: Use of evidence-based interventions in schools: Where

we've been, where we are, and where we need to go. *School Psychology Review,*

33, 398-407.

*Wert, B. Y., & Neisworth, J. T. (2003). Effects of video self-modeling on spontaneous

requesting in children with autism. *Journal of Positive Behavior Interventions, 5*,

30-34. doi:10.1177/10983007030050010501

APPENDIX A

QUALITY INDICATOR RUBRIC

Participants and Setting	1	2	3	4
Participant Characteristics (ex. age, gender, disability)	Information includes only 1 of the 3(age, gender, disability)	Information includes 2 of 3 (age, gender, disability)	Information includes all 3 (age, gender, disability) without any additional information	Sufficient detail provided including age, gender, disability (including how measured). Additional information such as measure of functioning, current levels of skill performance etc. provided
Selection Criteria	No information regarding how the participants were chosen	Description of selection criteria minimal (ex. individuals with ASD) without specific information regarding how chosen – (ex. Know they want to use participants with autism because of purpose however,	Selection criteria is provides some specific information regarding how the participants were chosen (ex. individuals with autism who used 1 word vocalizations)	Significant information regarding how participants were chosen including assessments used (ex. individuals with autism who used 1 word vocalizations and scored x on GARS)

		no info regarding how they were chosen…2 would be vague and only implied		
Setting Description	Setting not identified and limited details mentioned	Setting identified (ex in school) but no explicit details regarding the setting (ex, number of students, size of room, other people in the room)	Setting identified with limited details (ex. In a resource classroom of the school with one teacher)	Explicit details regarding the setting (ex. in a public school resource classroom with 5 students, viewed on a 5 X5 screen) allowing replication
Dependent Variable(DV)	1	2	3	4
Description	Names DV only without clear depiction(ex. aggression)	Narrow information of DV but lacks measurable definition (ex no hitting)	Dependent variable described in a manner that could be measured (ex. hands and feet to self) but without specifics such as criterion information	Clear, operational description of DV (ex. incidents of verbal and physical aggression as evidenced by hands and feet to self and no loud disruptions for 30 m interval)
Description of valid Measurement	Process not reported	Limiting information regarding how the DV was evaluated limiting ability for the reader to employ the same measure	Limited description of procedure provided (tells how measured but not schedule or tools utilized) and/or not all variables are measureable	Clear description of DV measurement included tools utilized, schedule, and measurable variables and the process is repeatable

Frequency of measurement	Measurement inconsistent across phases or DV measured for only 1 data point in one or more of the phases (excluding follow-up)	Less than 3 data points in one or more of the phases	At least 3 data points per phase	More than 3 data points per phase (unless explanation for 3 given)
IOA	Not provided	Data provided, however, method not reported or calculated inaccurately.(i.e.,on less than 20% of the data)	Either both or one of the measures, IOA and Kappa, are less than minimum standards : less than 20% of data for each phase **OR** IOA<80% and/or Kappa < 60%,	IOA assessed for at least 20% of the data points in each phase and IOA = 80% and/or Kappa = 60% (if only one reported it meets the minimum standard)

Note. Rubric is based on the quality indicators for single case research as established by Horner et al. (2005). The 4-point scale

established by Chard et al. (2009) served as a guide for the development of the above rubric.

SUMMARY OF VMO ARTICLES INCLUDED IN THE ANALYSIS

Authors	Participant Characteristics	Targeted Outcomes	Known (K)or Unknown (U)	Model	Alone, package or other	Quality Indicator Average
Alcantara, 1994	2 M, age 9 y and 1 F age 8 with ASD and DD	independent living	U	adult	+	3.81*
Allen, Wallace, Greene, Bowen, & Burke, 2010	3 M, ages 19-22 with ASD	independent living	U	peer	a	3.64
Allen, Wallace, Renes, Bown, & Burke, 2010	4 M, ages 16-25 with ASD	independent living	U	peer	a	3.76
Apple et al., 2005	3 M, 1 F ages 4-5 with ASD	Social-communication	K	peer	a	3.40*
Ayres & Langone, 2007	3 M, 1 F ages 6-8 with ASD	independent living	U	adult	a	3.12
Bidwell & Rehfeldt, 2004	3 F, ages 33-72 with DD	independent living	K	peer	++	3.57
Boudreau & DaEntremont, 2010	2 M, age 2 with ASD	Play	U	adult	a and +	3.83*

Cannella-Malone et al., 2006	3 M, 1 F ages 27-41 with ASD; 1 M, age 36 with DD and EBD	independent living	U	peer	a	3.42
Charlop et al., 2010	3 M, ages 7-11, with ASD	social – communication	K	adult	a	3.75
Charlop, et al., 2008	2 M, ages 8-9 with ASD	Social-communication	U	adult	+	3.85*
Charlop-Christy & Milstein, 1989	3 M, ages 6-7 with ASD	Social-Communication	K	adult	+	3.14
Charlop-Christy & Daneshvar, 2003	3 M, ages 6-9 with ASD	Social-Communication	K	adult	a	3.82
Charlop-Christy, Le, & Freeman, 2000	1 F, 4 M ages 7-11 with ASD	Play	K	adult	a	3.73
Cihak & Shrader, 2008	4 M, ages 16-21 with ASD	independent living	U	adult	++	3.73*
Conyers et al., 2004	2 F, 1 M ages 43-54 with DD	Behavior	K	peer	+	3.67*
D'Ateno et al., 2003	1 F, age 3 with ASD	Play	_	adult	a	3.75

Gena et al., 2005	1 F, 2 M ages 3-5 with ASD	Social-Communication	U	peer	++	3.74*
Haring et al., 1987	1 F, 2 M age 20 with ASD	independent living	K	peer	++	3.76
Keen et al., 2007	3 M, ages 4-6 with ASD	independent living	U	animated	++	3.36
Kleeberger& Mirenda, 2010	1 M, age 4 with ASD	Play	_	adult	++	3.86*
LeBlanc et al., 2003	3 M, ages 7 -13 with ASD	Social-communication	_	adult	++	3.61
MacDonald et al., 2005	2 M, ages 4 and 7 with ASD	Play	_	adult	a	3.61
MacDonald et al., 2009	2 M, ages 7 and 5 with ASD	Play	_	adult	a	3.64
Maione & Mirenda, 2006	2 M, ages 5-6 with ASD	Social - communication	U	adult	a	3.79*
Marcus & Wilder, 2009	1 F, 1 M age 9 and 1 M age 4 with ASD	Academic	K	peer	++	3.55
Martin et al., 1992	5 participants ages 15-20 with DD	independent living	_	peer	a	2.95
Mechling et al., 2005	1 M age 17 with ASD; 1 M, 1 F ages 17-20 with DD	independent living	K	adult	++	3.75*

Murzynski & Bourret, 2007	2 M, ages 8-9 with ASD	independent living	K		adult	++	3.26
Nikopoulos et al., 2009	2 M, 1 F ages 7-9 with ASD	independent living	U		peer	a	3.49
Nikopoulos & Keenan, 2004	3 M, ages 7-9 with ASD	Social-communication	_		peer	a	3.66
Nikopoulos & Kenan, 2003	6 M, 1 F ages 9-15 with ASD	social - communication	K		adult	+	3.65
Nikopoulos & Keenan, 2007	3 M, 1 F ages 6-8 with ASD	Social-communication	_		peer	+	3.74
Paterson & Arco, 2007	2 M, ages 6-7 with ASD	Play			adult	+	3.81
Rayner, 2010	1 M, age 12 with ASD	independent living	U		adult	+	3.60
Rehfeldt et al., 2003	2 M, 1 F ages 22-37 with DD	independent living	_		peer	++	3.62
Rosenberg et al., 2010	3 M, ages 3-5 with ASD	independent living	U		peer	+	3.68*
Sansoti & Powell-Smith, 2008	3 M, ages 6-8 with ASD	Social-communication	_		peer	a	3.93*
Scattone, 2008	1 M, age 9 with ASD	Social-communication	_		adult	a	3.81*
Sherer et al., 2001	5 M, ages 4-11 with ASD	Social-communication	U		peer	a	3.70

Taber-Doughty et al., 2008	3 M, ages 13-15 with ASD	independent living	U	adult	++	3.79*
Taylor et al., 1999	2 M, ages 6 and 8 with ASD	Play	K	peer	++	3.54
Van Laarhoven et al., 2009	2 F, ages 12 and 15 and 1 M, age 17 with DD	independent living	-	adult	++	3.64*

Note. An "*" by the QI average indicates the study met at least the minimum standards across indicators; Diagnostic codes:

ASD = Autism spectrum disorder; DD = Developmental disability; Alone, package, or other codes: a = alone; + = with

reinforcement; and ++ = package

SUMMARY OF VSM STUDIES INCLUDED IN THE ANALYSIS

Citation	Participant Characteristics	Targeted Outcomes	Type of VSM	Alone, package or other	Quality Indicator Average
Bellini et al., 2007	1M, 1F ages 4-5 with ASD	Social-communication	Feed	a	3.38
Buggey et al., 1999	2 M, 1 F, ages 8-11 with ASD	Social - communication	PSR	a	3.40
Buggey, 2005	5 M, ages 5-11 with ASD	Social-communication and behavior	Feed and PSR	a	3.52
Cihak & Shrader, 2008	4 M, ages 16-21	independent living	Feed	++	3.73*
Clare et al., 2000	3 M, ages 9-11 with LD/ED	Behavior	PSR	a	3.43
Coyle & Cole, 2004	3 M, ages 9-11 with ASD	Behavior	PSR	++	3.15
Dowrick & Ward, 1997	1 M, late 20s with ID	Behavior	Feed	+	3.32
Hepting & Goldstein, 1996	1 F, 2 M, age 4 with ID	Social - communication	Feed	+	3.51*

Hitchcock et al., 2004	2 M, 1 F; ages 6 and 8 with LD and 1 male age 6 with ID	Academic	Feed	++	3.87*
Kehle et al., 1986	4 M, ages 10-13 with EBD	behavior	PSR	+	3.42
Lasater & Brady, 1995	2 M, ages 14 and 15, with ASD	independent living	Feed	++	3.24
Lonnecker et al., 1994	2 M, ages 9 and 7 with LD	behavior	PSR	++	3.57
Marcus & Wilder, 2009	1 M , 1 F age 9 and 1 M, age 4 with ASD	academic	Feed	a	3.55*
Pigott & Gonzales, 1987	1 M, age 9 with EBD	behavior	PSR	a	3.10
Possell et al., 1999	4 M, ages 5-8 with EBD	behavior	PSR	+	3.56*
Sherer et al., 2001	4 M, ages 4-11 with ASD	Social-communication	Feed	a	3.49
Van Laarhoven, 2009	2 F, 1 M ages 12-17 with ID	independent living	Feed	++	3.52

| Wert et al., 2003 | 4 M, ages 4-5 with ASD | Social-communication | Feed | a | 3.44 |

VITA

Rose Ann Katherine Mason received her Bachelor of Science degree in psychology from Ambassador University in Big Sandy, TX in 1996. She entered the clinical psychology program with specialization in school psychology at The University of Texas at Tyler in August of 1996 and received her Master of Science degree in May 2008. After 7 years working as a licensed specialist in school psychology, she began her doctoral program in educational psychology with an emphasis in special education and graduated with her Ph.D. in May 2012. Her research interests include behavioral interventions for individuals with disabilities and implementation fidelity.

Dr. Mason may be reached at: 704 Harrington Tower, 4225 TAMU, College Station, TX, 77843. Her email address is: rosemason519@gmail.com.