

2010

The Use of Organizational Behavior Management in Training Applied Behavior Analysis Methods in Residential Settings: A Review

Katherine M. Stegman

Southern Illinois University Carbondale, kmstegman@gmail.com

Follow this and additional works at: http://opensiuc.lib.siu.edu/gs_rp

Recommended Citation

Stegman, Katherine M., "The Use of Organizational Behavior Management in Training Applied Behavior Analysis Methods in Residential Settings: A Review" (2010). *Research Papers*. Paper 35.

http://opensiuc.lib.siu.edu/gs_rp/35

This Article is brought to you for free and open access by the Graduate School at OpenSIUC. It has been accepted for inclusion in Research Papers by an authorized administrator of OpenSIUC. For more information, please contact opensiuc@lib.siu.edu.

THE USE OF ORGANIZATIONAL BEHAVIOR MANAGEMENT IN TRAINING
BEHAVIOR ANALYTIC METHODS IN RESIDENTIAL SETTINGS: A REVIEW

by

Katherine M. Stegman

B.G.S., University of Kansas, 2006

A Research Paper

Submitted in Partial Fulfillment of the Requirements for the
Master of Science Degree.

Department of Rehabilitation
in the Graduate School
Southern Illinois University Carbondale
December 2010

RESEARCH PAPER APPROVAL

THE USE OF ORGANIZATIONAL BEHAVIOR MANAGEMENT IN TRAINING
BEHAVIOR ANALYTIC METHODS IN RESIDENTIAL SETTINGS: A REVIEW

By

Katherine M. Stegman

A Research Paper Submitted in Partial

Fulfillment of the Requirements

for the Degree of

Masters of Science

in the field of Behavior Analysis and Therapy

Approved by:

Dr. Nicole A. Heal, Chair

Graduate School
Southern Illinois University Carbondale
November 12, 2010

AN ABSTRACT OF THE RESEARCH PAPER OF

Katherine M. Stegman, for the Masters of Science degree in Behavior Analysis and Therapy, presented on November 12, 2010, at Southern Illinois University Carbondale.

TITLE: THE USE OF ORGANIZATIONAL BEHAVIOR MANAGEMENT IN TRAINING BEHAVIOR ANALYTIC METHODS IN RESIDENTIAL SETTINGS: A REVIEW

MAJOR PROFESSOR: Dr. Nicole A. Heal

Applied behavior analysis has a long history of success in showing therapeutic gains within human service settings. The need to train all types of staff in behavior analytic methods is crucial to the continuation of this tradition. Research has shown organizational behavior management (OBM) to be effective in teaching others to implement behavior analytic programs. However, OBM appears to be largely absent from human service settings. The present research paper reviews the literature in this area, examines trends, and makes suggestions for future research.

DEDICATION

This paper is dedicated to my future husband and my family, especially to my brother for sparking my interest in this field.

ACKNOWLEDGMENTS

The author would like to acknowledge Jami Evans for her helpful comments on a previous draft of this research paper.

LIST OF FIGURES

<u>FIGURE</u>	<u>PAGE</u>
Figure 1	22
Figure 2	23
Figure 3	24
Figure 4	32
Figure 5	36
Figure 6	55
Figure 7	56
Figure 8	57
Figure 9	58
Figure 10	68
Figure 11	69

TABLE OF CONTENTS

<u>CHAPTER</u>	<u>PAGE</u>
ABSTRACT	i
DEDICATION	ii
ACKNOWLEDGMENTS	iii
LIST OF FIGURES	iv
CHAPTERS	
CHAPTER 1 – Introduction	1
CHAPTER 2 – Procedure	11
CHAPTER 3 – Staff Characteristics.....	14
CHAPTER 4 – Implementer Characteristics	25
CHAPTER 5 – Results.....	33
CHAPTER 6 – Social Validity	37
CHAPTER 7 – Maintenance	59
CHAPTER 8 – Conclusion	70
REFERENCES	78
VITA	86

CHAPTER 1

INTRO

Organizational behavior management has been an area of research within behavior analysis for decades. Broadly defined, organizational behavior management (OBM) is the examination of the behavior of those at work. More specifically, it is the application of Skinner's science of behavior to the behavior of those at work (Bucklin, Alvero, Dickinson, Austin, & Jackson, 2000; Hyten, 2002). As such, it is considered a sub-discipline of applied behavior analysis (ABA), not a separate field. OBM can include all work settings such as private and public organizations, and many categories of work-behavior including but not limited to: safety skills, efficiency, quality control, and the prevention of employee absenteeism (e.g., Bucklin et al., 2000).

In particular OBM has been utilized in many types of human service settings, including schools, residential settings for the developmentally delayed, special education settings, in-patient clinics, and autism centers. Organizational behavior management and the field of developmental disabilities in particular have been described as "old friends" (Sturmey, 2010). This may be due to the fact that the behavior of those served by human service settings has traditionally been a focus of ABA. Behavior analysts may enter these settings to write behavior plans for those served, but quickly realize staff training is needed to ensure they are carried out.

For this reason, behavior analysts have identified direct-care staff training as important since the 1960s (Sturmey, 1998). Behavior analysts have produced a significant body of research addressing staff training in human service settings. Additionally, researchers have published discussion pieces and meta-analyses to further analyze this research. The largest finding from this body of research has been a great deal of hard evidence showing OBM's success in adult residential settings (e.g., Reid 1998; Riley & Frederiksen, 1984). One meta-analysis examined the utility of OBM in community group homes specifically (Harchik & Campbell, 1998). This analysis found OBM to be useful in these settings as well, an important finding in light of the trend toward deinstitutionalization occurring over the past several decades (Harchik & Campbell, 1998).

Additionally, meta-analyses and discussion pieces have analyzed OBM literature to find the most effective procedures and provide general recommendations to practitioners. Some general findings of the research included, workshops alone are ineffective and frequent specific feedback is critical to reaching desired levels of behavior (Harchik & Campbell, 1998). In addition to meta-analyses researchers have published empirical studies to provide further analysis of OBM programs. For example, studies have compared role-playing versus lectures in training the use of positive reinforcement (Adams, Tallon, & Rimell, 1980), different types of feedback in improving client teaching skills (Ford, 1984), and evaluated client learning as the basis for providing reinforcement to staff (Azrin, Jamner, & Besalel, 1989).

Unfortunately, most human service settings do not use OBM routinely. Therefore, the effectiveness of OBM in these settings does not appear to be enough for the adoption and dissemination of OBM. This concern has been raised for some time and has been discussed by many within the literature (e.g., Christian, 1983; Fisher, 1983; Parsons, 1998; Schell, 1998; Sturmey, 1998). The *Journal of Organizational Behavior Management* published a special issue regarding the usage of OBM in the field of developmental disabilities in 1998. In this issue Reid (1998) provided an introduction to the special issue. In that introduction, Reid notes that every article in the issue mentions the lack of routine use of OBM in these settings as a problem. This was not a new assertion, in that the lack of OBM programs in human service settings has been discussed within the literature over the past few decades (e.g., Christian, 1984; Fisher, 1983; Parsons, 1998; Schell, 1998). These authors have presented a wide-range of possible reasons for this problem. A lack of social validity (Parsons, 1998), a restricted scope of applications (Reid, 1998), a lack of maintenance and generalization measures (Williams & Lloyd, 1992) and a lack of research to develop a comprehensive model of staff behavior (Sturmey, 1998) have all been hypothesized as contributing factors to the absence of OBM. Further discussion and research suggestions regarding the lack of OBM have also been published. The conclusion made by Fisher seems to have been made by many others: a behavior analyst needs more than an advanced degree and positive data to be successful in the settings in which they wish to work.

Researchers often investigate staff behavior unique to human service settings. Examples of these behaviors include staff providing basic daily care, and the usage of behavior analytic techniques by staff with the clients they serve (e.g., Arco, 1991; Azrin & Pye, 1989; Baker, Fox, & Albin, 1995; Fleming, Oliver, & Bolton, 1996). The usage of ABA techniques by staff appears to be an especially common focus of OBM interventions within the literature. In many residential settings, direct-care staff members are present at all times, especially agencies serving individuals with more severe disabilities. For this reason, many professional staff within an agency relies on direct-care staff to be their “eyes and ears.” In particular, behavior analysts may especially rely on direct-care staff in that behavior analysts frequently require direct-care staff to carry out their own programs, which may be one reason why staff behavior is often the focus of research. Additionally, these types of situations may be of extra interest to researchers as they represent ABA on two levels, basic behavior plans for clients *and* organizational behavior management. It has become apparent to researchers and practitioners alike; staff training is as important as developing behavior change programs to successfully change client behavior (Williams & Lloyd, 1992). Simply put, if a behavior plan is not implemented correctly, it will not work, no matter how well-designed the program. Other professionals in human service settings (e.g., occupational therapists, counselors, speech therapists) may require staff assistance in carrying out their programs. However these professionals may not rely on others to the degree behavior analysts do. This may be due to the fact that behavior analysts usually address behavior in

the natural environment where staff members work, rather than holding scheduled therapy sessions as these therapists frequently do.

Further explanation for the importance of OBM in these situations is the link between client gains and staff performance (Greene, Willis, Levy, & Bailey, 1978). The majority of studies later analyzed in this meta-analysis also emphasize this link by providing client outcome data in addition to staff performance data (e.g., Arco, 1991; Fleming & Sulzer-Azaroff, 1992; Green, Parsons, & Reid, 1993; Harchik, Sherman, Sheldon, & Strouse, 1992; Parsons, Rollyson, & Reid, 2004; Shore, Iwata, Vollmer, Lerman, & Zarcone, 1995). Staff behavior has an enormous impact on client behavior. Thus, it is important that staff be trained in a variety of behavior analytic techniques, not just client skill development. Often problem behavior exhibited by a client (e.g., aggression, self-injurious behavior, self-stimulatory behavior) are viewed as a problem exclusive to the client, a view that minimizes the effects of the client's environment (Williams & Lloyd, 1992). Even clients with no previous record of unwanted behavior may engage in problem behavior when staff members fail to provide leisure materials or interact with clients in a positive manner.

There could be many reasons OBM is not used more routinely to ensure staff are utilizing behavior analytic principles. Adding a staff training component to any program adds more unknown variables into the research (Williams & Lloyd, 1992). While changing behavior is difficult, affecting behavior change through others is even more so. Unfortunately, very few direct-care or professional staff members in human service settings have any formal

background in behavior analysis, adding further difficulty (Fisher 1983).

Moreover, few staff members have any formal education in working with special populations. Many staff members do receive a basic introduction to behavior analysis when they begin working in these residential settings (Fisher 1983).

These trainings are often classroom-type trainings and include many different aspects of staff member's job such as paper work and basic first-aid. However, these short classroom-type trainings are unlikely to provide new staff ABA skills they can apply to their job settings, or even to impress upon them the idea that behavior analytic techniques are useful. Additionally, the skills learned in these trainings are unlikely to be used long-term by staff without any follow-up by the agency employing them (Harchik & Campbell, 1998). These difficulties may lead practitioners to determine that training staff to implement behavior analytic techniques is simply too difficult.

One potential problem practitioners may face when researching this topic is locating relevant literature. As this type of research has components in two areas of ABA, reviewing the literature can be difficult. Many researchers place these studies in the OBM literature due to its staff training components. For this reason, many empirical studies and discussion pieces can be found in the OBM flagship journal the *Journal of Organizational Behavior Management*. Others however, place this topic in with the more traditional ABA literature due to the usage of behavior analytic techniques with clients. Therefore, many articles and discussion pieces on this topic can also be found in the research flagship journal the *Journal of Applied Behavior Analysis*. The present meta-analysis attempted

to gather the research from multiple journals, including the two discussed here, and present cohesive findings.

It appears the most frequent type of residential setting OBM has been implemented in is settings serving those with developmental disabilities (Sturmey, 2010). Many of the discussion pieces on OBM in human service settings discuss facilities for those with developmental disabilities exclusively (e.g., Harchik & Campbell 1998; Parsons, 1998; Schell, 1998). However, in this investigation, research taking place in any adult residential facility was considered for inclusion. Adult residential facilities could include those serving individuals with developmental disabilities, chronic mental illnesses, traumatic brain injuries, or other intellectual disabilities. For the sake of simplicity, this analysis does not include nursing homes or other settings meant to serve the elderly. Although, many different types of settings could have been considered for inclusion, nearly all the literature reviewed involved agencies that served those with developmental disabilities. One study (Huberman & O'Brien, 1999) took place in an agency serving clients with chronic mental illness. This study was the only study utilizing OBM in a setting that did not serve individuals with developmental disabilities. Interestingly, the data obtained from this study did differ from the other studies including those with developmental disabilities in several aspects. These differences will be discussed later within this analysis.

In that OBM has thus far been unsuccessful in gaining wide-spread adoption in human service settings, when will behavior analysts be successful in disseminating these programs? Williams and Lloyd (1992) characterized the

1950s and 1960s as a time when researchers focused on changing client behavior. They further characterized the 1970s as a time when researchers began to use these same principles to change staff behavior. Based upon these research trends, Williams and Lloyd predicted research in the 1990s would show the development of large scale management and organizational-level interventions. It was hoped that with these larger-scale interventions the organizational environment would be more favorable for the implementation of OBM programs. Unfortunately, neither the development of organizational-level interventions nor the increased usage of OBM appears to have occurred.

This paper presents a meta-analysis on the use of OBM by researchers to ensure staff is correctly implementing behavior analytic techniques in adult residential settings. There are many reasons this meta-analysis was completed. First and foremost, this analysis was done to help answer the question: why is OBM not implemented more commonly in human service settings? More specifically, why is it not utilized to implement the programs behavior analysts themselves frequently use in their daily work? Second, this analysis hoped to uncover patterns in the existing literature, both to recognize our successes and uncover gaps in the research, which could be addressed in the future literature. Hopefully, recognizing both OBM's successes and failures will provide further guidance for behavior analysts to disseminate OBM in adult residential settings.

It has become clear that if behavior analysts are to be successful in these types of settings the support of others is imperative. Behavior analysts require those working with clients served to implement written behavior plans. When

direct-care staff members implement behavior plans incorrectly (or fail to implement altogether), valuable time and effort is wasted. Furthermore, past researchers have found management to be ineffective in nearly every instance staff members have been found to be ineffective in their duties (Reid & Whitman, 1983). Staff must be trained to correctly utilize ABA procedures, if behavior analysts are to be successful. OBM has been proven to be effective in training staff to use ABA when working with clients and encouraging the continued use of ABA procedures.

One example of a study included in the current meta-analysis investigated the effects of behavioral contracting between a supervisor and employee (Azrin & Pye, 1989). The behavior of interest was client training components completed correctly (e.g., presence of edible reinforcers, presence of data collection materials). Another representative study (Mozingo, Smith, Riordan, Reiss, & Bailey, 2006) examined the effects of an in-service, systematic supervisor feedback, and supervisor presence without feedback on the correct recording of client problem behavior. In this meta-analysis, studies contained interventions aimed at the initial training of staff in behavioral analytic methods, and interventions designed to facilitate staffs' usage of previously developed programs were examined. Analyzed studies focused on both client skill acquisition and addressing problem client behavior.

First, a description of how relevant studies were identified, and the criteria for their inclusion in this review is discussed. Second, the resulting data of this meta-analysis is presented. Data developed from this analysis is divided into five

categories. Categories included are as follows: (1) staff characteristics, (2) implementer characteristics, (3) results obtained by each study, (4) social validity, and (5) maintenance. After data for each of these categories is presented a brief summary will be presented. Important/interesting aspects of the presented data and implications of each data set will be discussed. Finally, suggestions for future research will be presented and why these suggestions are being made will be discussed.

CHAPTER 2

PROCEDURE

In order to identify relevant research, the internet search engine PsycInfo was used. This search engine was used to identify articles within the following three journals: *Journal of Applied Behavior Analysis*, *Journal of Organizational Behavior Management*, and *Behavioral Residential Treatment*. When searching the journal *Behavioral Residential Treatment*, PsycInfo included articles from the journal *Behavioral Interventions* by default. If the articles from this journal (*Behavioral Interventions*) met the inclusion criteria, these articles were included in the analysis as well. Searches were conducted using the following search, “staff training,” “behavior analysis,” and “staff management.” All articles meeting specified criteria were included in this meta-analysis. Further articles were then identified and included using the reference sections of those articles.

The criteria for an experiment’s inclusion were: (1) must be an empirical study (i.e., include the systematic manipulation of an independent variable), (2) must be an OBM intervention measuring paid staff behavior, (3) have taken place at a residential facility for adults (i.e., the authors described clients as “adults”, the facility was described as an adult facility, or the clients ages were specified as all 18 or older, (4) have been published between the years of 1980-2010, and (5) have the main dependent variable of interest be some measure of staff’s performance when implementing a common behavior analytic method.

Examples of behavior analytic methods included assessments such as functional analyses and preference assessments (e.g., Green, Reid, Perkins, & Gardner, 1991), proper implementation of programs to develop skill acquisition with clients served (e.g., Realon, Lewallen, & Wheeler, 1983), proper implementation of programs to decrease unwanted behavior exhibited by clients (e.g., Methot, Williams, Cummings, & Bradshaw, 1996), graphing or data collection skills related to client behavior (e.g., Mozingo et al., 1996), and programs to lead to environment enrichment such as increased staff-client interactions or the delivery of noncontingent reinforcement (e.g., Harchik et al., 1992). Some non-examples of behavior analytic methods would be assessing staff's attitudes towards behavior analysis (e.g., Reid & Parsons, 1996), assessing their performance on a written test without later applying learned skills, or assessing the amount of paperwork completed if the paperwork was unrelated to client behavioral plans. Within some of the included articles, more than one experiment was conducted. In these instances all experiments were considered individually, were included or excluded individually, and then analyzed individually throughout the analysis. Twenty-nine experiments were included in this meta-analysis. These 29 experiments were drawn from a total of 26 empirical journal articles.

Many of the included articles used different terms for similar groups of people. Some of the terms used to describe special populations served by an agency included: residents, adults, participants, clients, men and women served, and students. For the sake of simplicity, in this meta-analysis those who were served by the residential facility or workplace are labeled as "clients." Many

authors used different terms for staff working with the clients (i.e., those who were the main participants of the OBM programs) as well. Some terms used to describe this group of people included: direct-care staff, paraprofessionals, teachers, float staff, and job coaches. For the sake of this analysis, “staff” delivered programmed antecedents or consequences to clients. When the term “staff” is used here it refers to those implementing behavior analytic techniques with clients and having their behavior measured in doing so by “implementers.” Implementers were anyone who trains or manages staff in using behavior analytic techniques. Thus, “Implementers” delivered antecedents or consequences to “staff.” Thus, the terms “staff” and “implementers” could include both direct-care staff and supervisors/managers. Individuals were placed in the two categories based upon their duties within the research project (who they delivered antecedents or consequences to).

CHAPTER 3

STAFF CHARACTERISTICS

Subcategories included in the analysis of staff characteristics were; the type of staff serving as participants (direct-care or professional/supervisory), staff education levels, staffs' ABA background, and the number of staff participants. All of the included studies reported some general information on staff characteristics. First, studies were organized by the type of staff serving as participants. Categories included; no data, direct-care staff, and supervisors, managers, or professionals. The direct-care versus professional staff categories were not mutually exclusive; a singular study could have multiple individuals warranting inclusion in both categories. For example, an experimenter could use an OBM program to facilitate positive interactions between clients and both direct-care staff and a program director (e.g., Parsons & Reid, 1993). The one study (Reid, Green, & Parsons, 1999) which did not specify the role of the staff used the term "job coaches." Although it was implied job coaches were direct-care staff, this assumption was not made for the sake of this analysis. Thus, the Reid study is included in the "no data" category. The majority of the remaining studies (93%) included direct-care staff. Seven of the twenty-eight remaining studies (25%) described supervisors, managers, or professionals as the subjects of the OBM program. Most of the studies analyzed here which included professional staff included direct-care staff as well. Of the seven studies including supervisors, managers, or professionals, five of them also studied

direct-care staff behavior. Thus only 2 studies of the 29 total examined the behavior of professionals/supervisors exclusively (Huberman & O'Brien, 1999; Parsons & Reid, 1995).

Information provided in the articles was further analyzed to identify staff education levels. As discussed in the introduction, one reason it may be difficult to train others to use behavior analytic techniques may be a lack of education or past experience with these methods. The majority of studies (79%) gave at least some information regarding staff education. Many gave detailed information, not only specifying if staff had college degrees but the general field in which each participant earned their degree (e.g., Baker, Fox & Albin, 1995; Fleming, Oliver, & Bolton, 1996). Not included in the 79% mentioned above, were studies that stated staff "were comparable to other staff in similar settings" (e.g., Parsons, Cash, & Reid, 1989). These studies were not included due to the lack of specificity needed for the placement in categories listed here. Of those reporting education data every study reported all staff had obtained high school degrees or equivalent certificates. Four studies, (17%) reported staff had further education relating to a human service field. Examples included in this category were: certified teachers, psychologists, or staff for whom the term "held a degree in a related field" was used by the author (e.g., Fleming, Oliver, & Bolton, 1996).

The four studies that included staff with related education were further organized into staff with or without education or experience specific to ABA. Only one of these studies (Arco, 1991) reported staff with any education/experience specific to ABA. Arco described all staff as having completed at least high

school. The author further reported that all staff had to complete an on-the-job certificate course which included ABA instruction. However, the details of the training described in this study are unclear. Details regarding the training such as: how much of this training was ABA-orientated, experience of the trainer, details provided to the staff, teaching methods used (i.e., handouts, lectures, role-plays) and covered topics could have been useful. Additionally, it is unclear if the ABA topics in this training course were related to the prompting and praising of clients, the variable of interest in the subsequent OBM program and focus of the Arco study. Therefore, no study reported those with prior behavior analytic experience or those with behaviorally-orientated degrees as staff. To summarize staffs' prior ABA experience, only one study (3%), of all experiments involved staff with stated prior ABA background (Arco 1991). However, it should be noted that this data only includes studies directly stating staff as having ABA education or experience. Thus, the 3% figure could be an underestimate. It is possible that more staff participants had past behavior analytic experience, but this information was not explicitly presented in the article. Additionally, staff may have had prior knowledge/experience of ABA the authors of an article were unaware of.

Next, studies were organized into categories based upon the number of staff participants included in the OBM program. Of the total 29 studies, a large majority (93%) reported the number of staff participants included in their research. The number of staff participants included in each study was broken down into three categories. These three categories were; 1-5, 6-10, and 11 or

more staff members. Just over one-third (37%) of studies included 1-5 staff, slightly more (44%) included 6-10, and lastly, five studies (19%) included 11 or more staff. While studies with fewer staff participating appear to be more common in these types of investigations, one study did include approximately 110 direct-care staff participating in an OBM project (Parsons, Cash, & Reid, 1989). The second largest investigation included 41 staff members (Williams, Di Vittoria, & Hausherr, 2002). Although there was some ambiguity, most of the articles in this analysis gave relatively clear information regarding the characteristics of their staff participants. In many instances, authors were not clear on the exact number of participants. However, authors frequently did give information sufficient to categorize a study (i.e., explaining two or three different staff worked two different shifts, thus allowing the reader to derive the number of staff members).

To provide a brief summary of staff characteristics the large majority of studies included three features. These three features were: a small number of staff subjects, the use of direct-care staff, and staff without prior knowledge of ABA or a related field. The majority of studies (81%) included fewer than 10 staff. As well, the majority included direct-care workers, and did not state staff had any prior knowledge of ABA (93% and 97%, respectively). This finding replicates the discussion presented by Reid (1998) in that most OBM in these settings are done with direct-care staff. Only two studies (Huberman & O'Brien; Parsons & Reid, 1995) of the 28 examined here focused exclusively on supervisors, managers, or other professionals. Perhaps the most encouraging

general finding is that every study gave at least some information regarding the role of staff included in the OBM program within the agency.

There could be several reasons why few staff may serve as subjects. The trend of deinstitutionalization may have some effect. Adult residential settings are now serving fewer clients in smaller settings as services move from large institutions to agencies with multiple community-based homes (Harchik & Campbell, 1998). Additionally, it may be difficult for researchers to gain access to multiple homes within an agency. For instance, an agency may have seven residences with just a small number of clients and staff living and working in each residence. The researcher in this example may need cooperation from seven different supervisors to begin a research project. The logistics of traveling between multiple locations to train staff members, collect data, and deliver consequences may be extremely difficult. Researchers focused on program development that may simply recruit staff working at one location. Some possible ways experimenters might be easing these difficulties are recruiting assistance from other researchers, training others at an agency's various locations to assist, or in the case of classroom-type training, gathering staff in one central location. Even if there is a larger number of staff available in one location, researchers may include a limited number of participants to ease data collection. For example, a researcher may include only the morning shift staff at a facility as a method of simplifying a research program.

Although investigations with fewer participants may dominate the literature, it may be that behavior analysts need to complete larger investigations

with more participants. However, as mentioned in the previous paragraph the logistics of including larger numbers of staff in community-based services may be difficult. Many practitioners will experience these difficulties if they attempt to apply the research to their daily work. Do researchers need to experience first-hand the difficulties of practitioners? As these contingencies likely operate on the behavior of practitioners attempting to affect agency-wide change, researchers should attempt to experience these contingencies themselves when applicable. By experiencing these difficulties the researcher may decide to slightly alter a proposed program, decide another intervention is more appropriate in an applied setting, or even develop new programs and OBM techniques. However, the ideal number of subjects for a proposed program should depend on the goals of the researcher. If the goal of the study is merely to show a functional relation, a study with a smaller number of staff participants may be appropriate. This would be suitable for researchers with a “research and development” type focus. Studies with a more applied focus, such as those testing existing programs, or attempting to replicate programs in applied settings may necessitate larger numbers of staff participants.

Based upon the data regarding staff characteristics and discussion presented here, it is recommended that researchers include larger numbers of staff in some of their studies when appropriate (e.g., when attempting dissemination of an intervention, when an agency wishes to serve a larger number of clients). Studies with smaller groups are still valuable, especially when more experimental in nature (e.g., testing new interventions). However, if

OBM is to gain more acceptance in these settings, behavior analysts must ensure OBM programs are sufficiently applicable to the settings practitioners will face in real work situations. Every study in this meta-analysis used a single-subject design (one used a group design in conjunction). Single-subject designs are largely considered one of the hallmarks of behavior analysis. One of the benefits of these designs is a large number of subjects is not necessary in showing a functional relation between the target behavior and environmental variables, making it easier for many researchers to simply include fewer staff participants. However, larger numbers of participants can be used within single-subject designs, and should be used when the focus of the research deems appropriate (as mentioned previously).

It is also suggested that researchers include supervisors, managers, and professional staff more frequently. A previous meta-analysis (Schell, 1998) found very few OBM interventions with professional staff as subjects. This finding is replicated here. By reviewing a bibliography of 244 OBM studies Schell found that most studies utilizing professional staff included teachers as subjects (30 studies included teachers compared to 16 including all other professionals). Teachers, however, could be considered the direct-care staff within a school setting (Schell, 1998). Thus, even fewer authors focused on professional staff in residential settings or with adults. Including professional staff as participants of OBM programs more frequently could have multiple benefits. Usually supervisors or managers are required to demonstrate the skills those working for them are required to possess. As mentioned previously, the logistics of involving

larger numbers of staff in OBM procedures can be difficult. Professionals, supervisors, or managers may assist in alleviating some of these problems.

Training professionals/supervisors in ABA is the first step in training them to later supervise direct-care staff in the usage of ABA. Additionally, all outside consultant influence within an agency must end at some time. Training professionals/supervisors may facilitate long-term maintenance of OBM programming and ensure continued usage of ABA techniques by all staff.

Figure 1. Staff Participants: The type of staff serving as participants in the described OBM studies. If a study was not applicable or did not provide sufficient information, it is not represented. The two categories are not mutually exclusive.

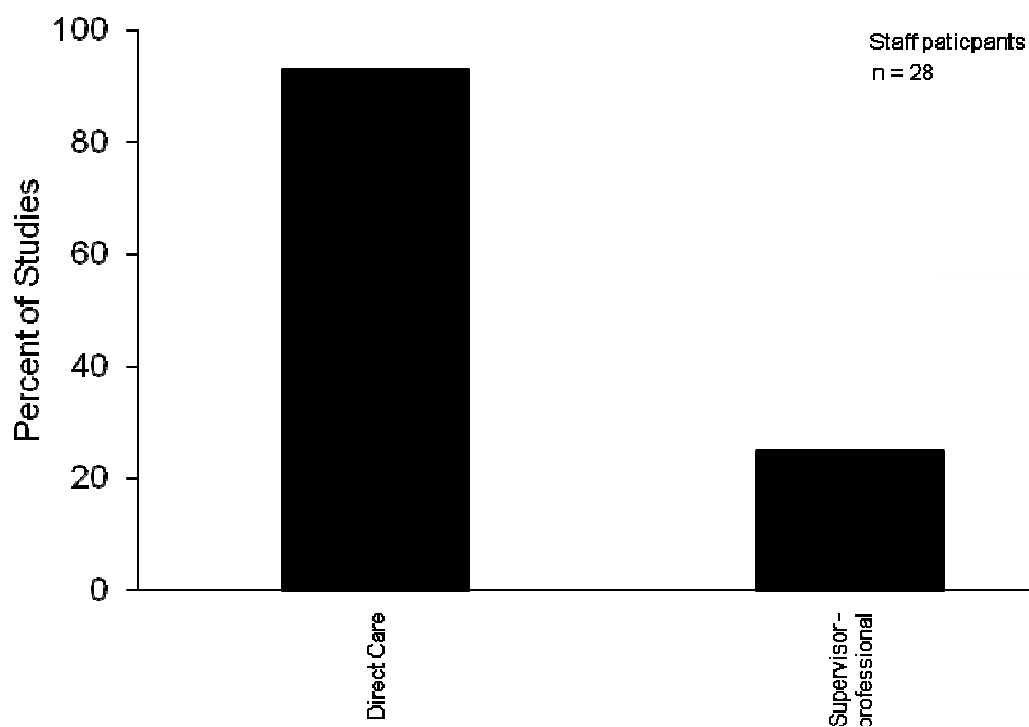


Figure 2. Staff Education Levels: The amount of education obtained by staff participants as reported by the authors of each study. Categories are not mutually exclusive.

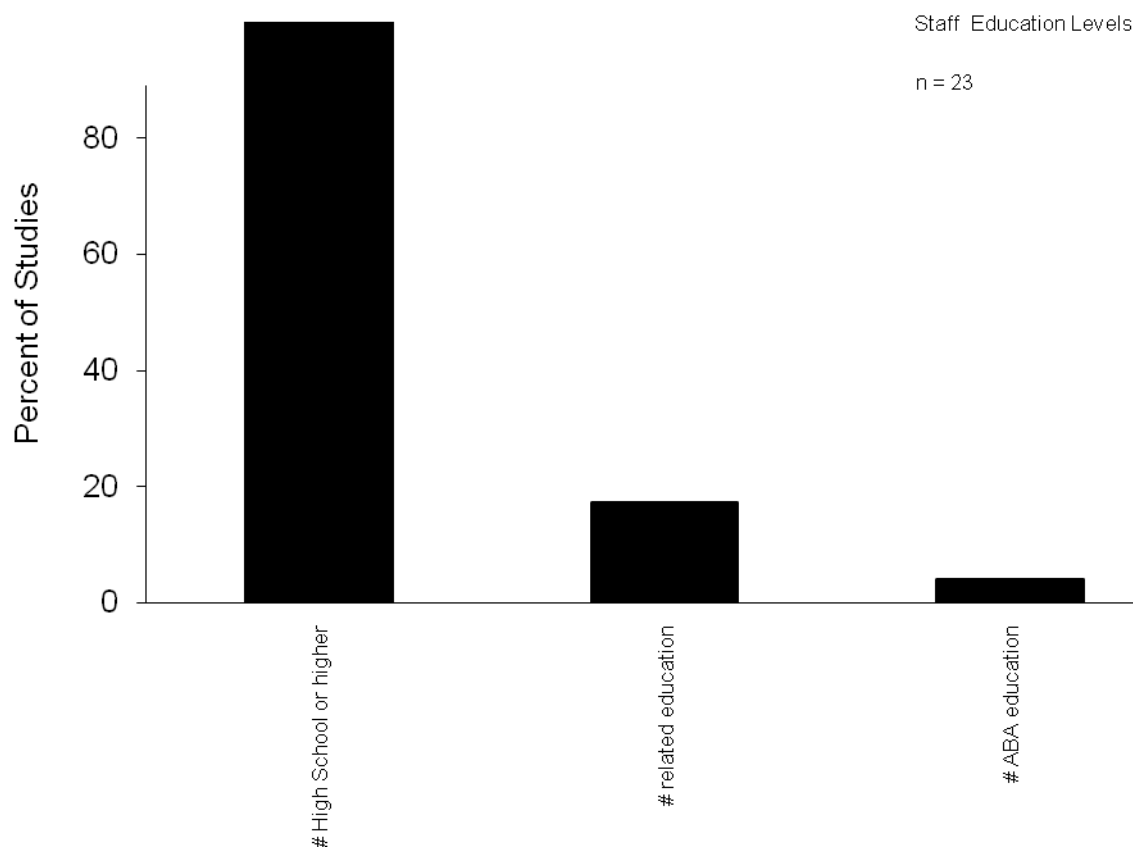
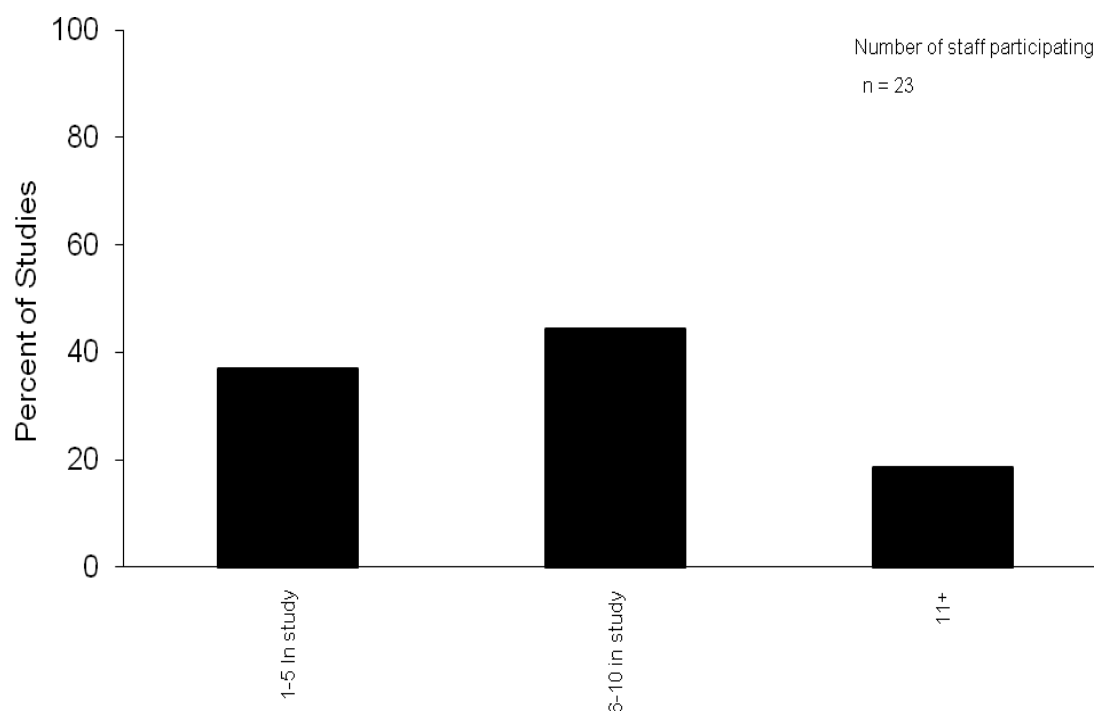


Figure 3. Number of Staff Participating in each OBM intervention as reported by the authors. Columns are mutually exclusive.



CHAPTER 4

IMPLEMENTER CHARACTERISTICS

As mentioned in the introduction, for the purpose of this meta-analysis, “implementer” first had to be defined. To review, an implementer had to do at least one of the following within the described OBM program to qualify for inclusion as an implementer: deliver some consequence to staff based upon staff behavior, or deliver some type of training to staff. Thus to be considered an implementer, that individual had to have direct contact with the staff implementing behaviorally-orientated programs in regards to those programs. If an individual was involved in planning or assisting with the OBM program in some other way (such as scheduling observation times), but did not meet one of the above two criteria, they were not classified as an implementer.

First, studies were organized by the presence or absence of information regarding implementers within the article. The majority, 23 out of 29 studies (79%), provided at least some details regarding program implementers. These 23 studies were then organized by type of implementer for at least one intervention component. The categories included here were: (1) experimenter or author, (2) supervisor, manager, or director, and (3) other team member. The categories were not mutually exclusive and studies were placed in all applicable categories. Of the 23 studies with implementer information, 14 studies (61%) mentioned an experimenter/author as implementing at least some component of

the program being investigated. Fifteen studies, (65%) indicated a supervisor, manager, or director as the implementer of at least one intervention component. For the category, “other team member,” only one study was found, representing 4% of the studies giving implementer information. The one study utilizing an “other team member” (Green, Parsons, & Reid, 1993) had a team psychologist implement the OBM program, along with an experimenter. Studies where implementer information was not given or sufficiently described were placed in the “no data/unclear” category. If the term “consultant” was used this information was considered unclear and was included in the “no data/unclear” category. It was determined the consultant may or may not have been an author/experimenter or held supervisory duties amongst staff. In summary, in the majority of studies presenting implementer information, an author or experimenter implemented at least some component of the intervention themselves. Additionally, a majority utilized staff with a supervisory/professional role. Only one study reported utilized an implementer outside these two categories. Perhaps, the most surprising and discouraging finding of this analysis was that six studies, or 21% of all studies, gave no information concerning implementers whatsoever. Implementer information could be useful to those deciding if an intervention is appropriate for a given setting. Additionally, it may assist a reader in replicating the intervention.

There appears to be a great deal of overlap across categories describing implementers. Two categories (experimenter/author, and supervisor/manager/director) were well represented. One potential reason

overlap may occur frequently is that the implementer fulfills multiple roles. For instance, Parsons et al. (2004) included an implementer who was both the director of the agency and listed as an experimenter. Thus, the same individual could warrant an implementer's inclusion in more than one category. Another reason for this overlap is the described OBM procedure involved multiple components with different individuals implementing various components (e.g., Mozingo et al., 2006; Williams et al., 2002). For example, an experimenter may deliver an initial staff training and a different individual, a supervisor, may provide follow-up feedback. Lastly, a diverse group of people could implement the same components of a program at different times or as teams (e.g., Harchik et al., 1992), such as both an experimenter and two managers delivering feedback. In this example, the person delivering feedback, and thus the implementer, could vary based upon who witnesses the staff behavior of interest in the moment.

It is encouraging to note that although several studies gave no information regarding implementers, the majority discussed them in some way. However, several studies went into further detail of staff characteristics describing education, age, gender, length of employment, and past experience with behavior analysis (e.g., Parsons & Reid, 1995; Realon et al., 1983; Sigafoos, Roberts, Couzens, & Caycho, 1992), giving more complete information than for other aspects of the study. The more complete information regarding staff characteristics could be due in part to the fact that most research articles included a section specifically referencing participants. In general, most behaviorally-orientated journals do not have separate sections for implementer

characteristics as they do for participants. Confusion could occur when a study lists the author as the implementer but fails to mention the author's role within the agency (e.g., Ducharme & Feldman, 1992; Fleming & Sulzer-Azaroff, 1992). The author may or may not have been in a supervisory position, and may or may not have had a strong relationship with the staff they were observing.

Additionally, many studies analyzed here provided more complete information when discussing data collectors. Often authors discussed data collectors in the context of reducing reactivity (e.g., Richman, Riordan, Reiss, Pyles, & Bailey, 1998). Thus, authors may not have been discussing data collectors for the sake of describing their procedures in detail, but to defend their investigations from questions of reactivity. This also explains why many studies provided information regarding data collectors and their relationships with those they were observing. For example, Richman et al. (1998) explained that data collectors included students and the first author. Richman et al. further clarifies that data collectors did not interact with staff in any way and were rotated amongst residences in order to assist the researchers in detecting reactivity effects.

Multiple studies discussed implementer information for some components of an intervention but not all components. Incomplete information made some sections of the present analysis more difficult. For example, Huberman and O'Brien (1999) specified that a director and experimenters delivered feedback to staff based upon performance. However, no information regarding implementers for the initial training described within the article was given. In the present meta-

analysis studies were classified by the information provided. When information was unclear or missing the study was placed in a category based upon the information given. In the example mentioned previously, (Huberman & O'Brien, 1999) this study was placed in both the experimenter/author and supervisor categories.

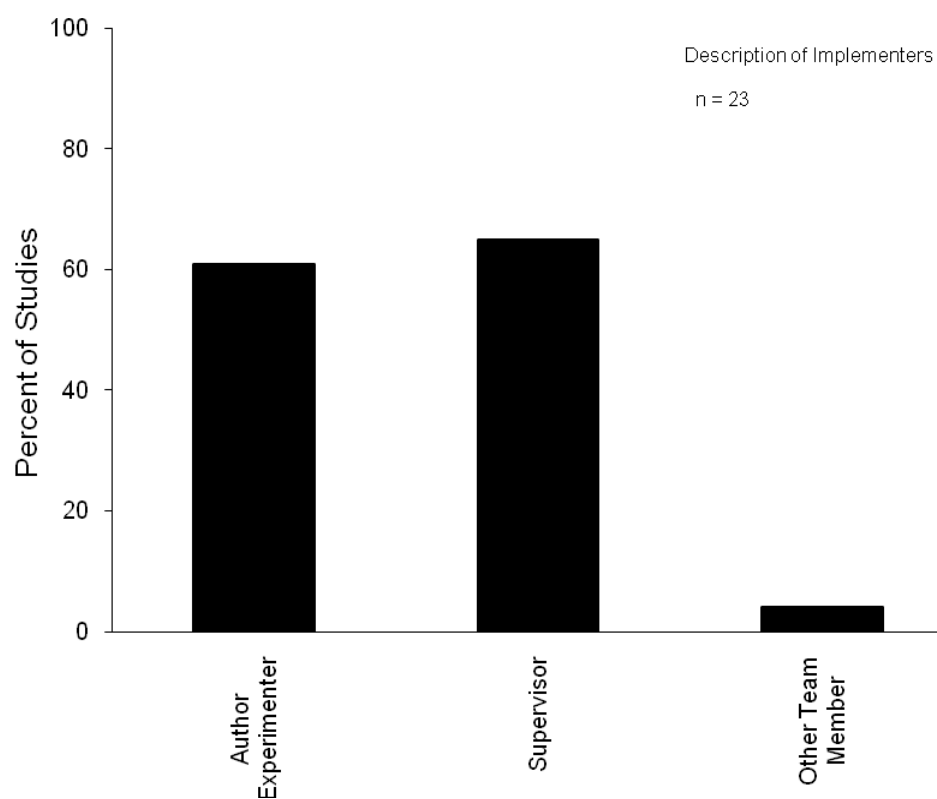
Many studies discussed recruiting individuals unfamiliar to staff (e.g., students, interns) as data collectors for the purpose of reducing reactivity (e.g., Richman et al., 1998). It is likely that data obtained from a supervisor/observer is more susceptible to reactivity. This may be especially true if the staff has had their behavior reinforced or punished by the supervisor previously. The supervisor may serve as a discriminative stimulus for the delivery of a consequence. One study included in this meta-analysis (Mozingo et al., 2006) did investigate to some degree the extent to which a supervisor served as a discriminative stimulus. In one phase of this study, a supervisor delivered feedback to staff based upon the target behaviors. In the subsequent phase, supervisors were present but did not deliver feedback. The researchers found behavior did maintain at levels similar to those occurring in the supervisor presence plus feedback treatment condition. As a result, Mozingo et al. (2006) found evidence supervisors were serving as discriminative stimuli. However, it should be noted that these results may be partly due to the fact that supervisors had been systematically paired with feedback previously.

Based upon the data presented and discussed, the chief recommendation from this meta-analysis is for authors to present more details on implementers. Clearer information regarding who implemented the various components of an intervention could be helpful for replication purposes. Additionally, more information could be useful about an implementer's role in the agency, if they held a position of authority, their education, and their past experience implementing the given procedures. An example of an article specifying an implementer's role within the host agency was found among the studies in this analysis (Green et al., 1991). In this study the authors specify that the implementer of the program was an experimenter. However, the authors provided further detail explaining she was a representative of the facility's education department and had programming responsibilities for the clients on the unit. Thus, the implementer may have already been associated with client's behavior plans and could serve as a discriminative stimulus for carrying out these plans. This information might be important to readers of an article, especially if they plan to implement such procedures without a similar implementer.

A general statement about an author's relationship to the staff could also be useful. For example, one study included in this meta-analysis (Arco, 1991) described the implementer/author as a consultant psychologist and as having established a good relationship with the staff over the previous year. Information such as this could be beneficial to practitioners interested in implementing these procedures, as it reveals the level of rapport building that may be needed to successfully implement the described OBM program with staff.

The second general suggestion here is to include supervisors or in-house staff (i.e., long-term employees of the agency) more often. As mentioned previously, all outside influence within an agency must end at some time (Sigurdsson & Austin, 2006). Thus, creating a program which relies entirely on outside consultants will most likely end when the consultant's time with an agency does. Utilizing in-house staff could have several possible benefits. Some benefits may include; in-house staff may be able to assist in the fading out of outside influence; facilitate better results by giving staff more "ownership" in the intervention, and allow these staff to implement procedures in other settings throughout the agency.

Figure 4. Description of Implementers: Categories of staff implementing at least one intervention component. Categories are not mutually exclusive.



CHAPTER 5

RESULTS

All of the studies included in this meta-analysis presented data on staff behavior, as the presentation of OBM data was a criteria for inclusion. First, each analyzed study was placed into one of three categories based upon the results obtained by the researchers: (1) those obtaining positive results (i.e., studies where staff were implementing programs at an acceptable level), (2) those with mixed results and finally, (3) those with no results. This analysis was based upon data of staff behavior presented by the author of each article. Although, many studies reported various types of client data (e.g., Arco, 1991; Fleming & Sulzer-Azaroff, 1992; Huberman & O'Brien, 1999) related to the behavioral programs staff were applying (e.g., length of engagement in leisure materials), the focus of this review was staff behavior. Additionally, one study (Harchik et al., 2001) reported dependent variables regarding staff behavior in addition to their usage of ABA (e.g., knowledge of company procedures). In this case only the data presented as measuring staffs' use of ABA was analyzed. If the authors of a study described the data obtained as positive but highly variable, this study was categorized as obtaining mixed results. The majority of studies analyzed here reported positive results (69%), just under one-third (31%) reported mixed results. No study in this analysis reported a lack of results or behavior lower than baseline levels. The data obtained here suggests that OBM is overwhelmingly successful in training staff and ensuring the implementation of

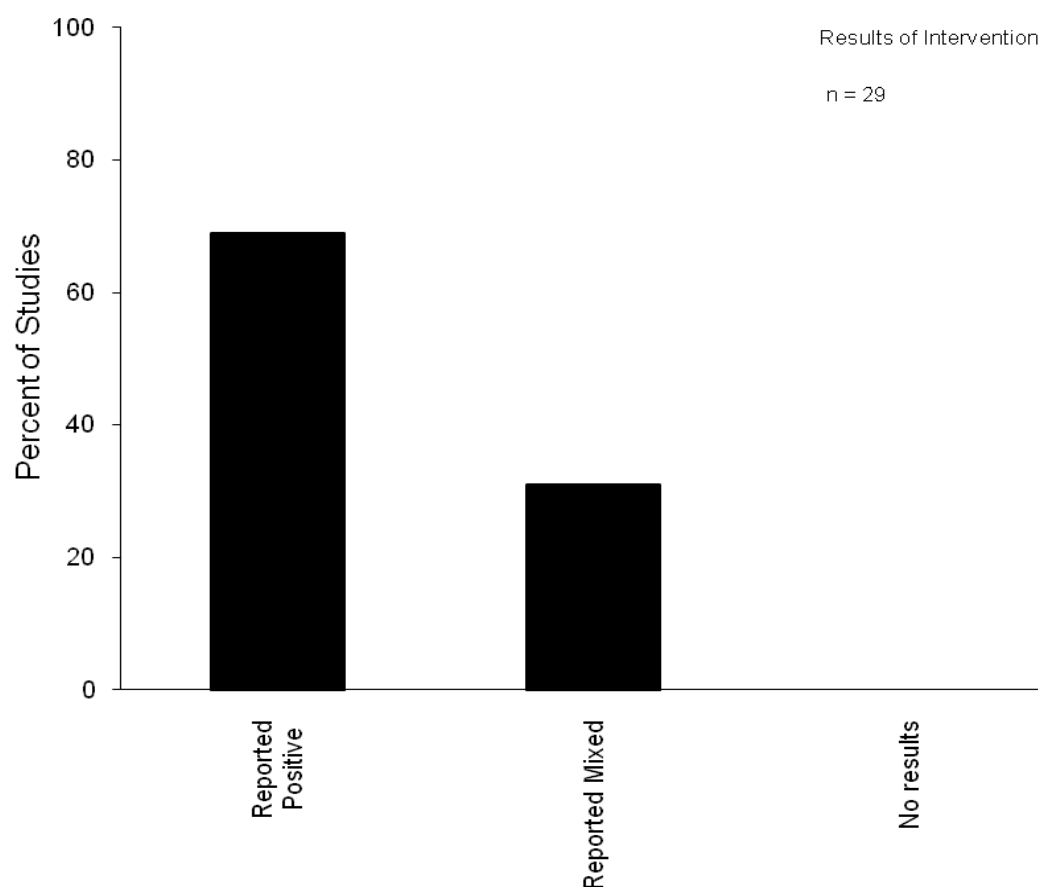
client behavior programs. This finding strengthens past conclusions presented in the literature. Such conclusions have been presented for over 20 years. For example, a discussion piece by Reid (1998) stated the research on OBM shows success when used in developmental disabilities settings. Further discussion by Riley and Frederiksen (1984) discussed OBM's success in human service settings in general. While the success of OBM researchers is an encouraging finding, it does raise some questions. One question to be raised is if a publication bias is affecting these results. Publication biases have been discussed by behavior analysts and within psychology research as a whole (Schwartz & Baer, 1991). Perhaps, the strongest conclusion we can gather from this data is that positive results alone are not enough to ensure the success of OBM programs. While OBM may be shown to be effective, this effectiveness alone does not appear to facilitate the usage of OBM procedures in these settings or as a tool in carrying out client behavior plans (Reid 1998).

It is suggested that to reduce any publication biases, more studies with mixed or no results be published. This suggestion is made for three main reasons: (1) researchers may learn if OBM programs are truly as effective as data suggests, (2) allow researchers to address gaps in the existing literature, and (3) allow researchers to search for patterns in unsuccessful programs. The final reason presented here may be of special importance. When researchers learn about problems faced by "failed" OBM programs, it could serve as a warning sign if similar problems are occurring within their own research. Researcher could then address the problem before it results in program

termination. It is foolish for behavior analysts to repeat the mistakes other researchers have made previously.

A successful program could be terminated for many reasons typically discussed in the behavior analytic literature. Issues directly related to the research itself such as, no programming for maintenance or generalization, or an absence of social validity, may cause problems. Additionally, a publication bias may extend to programs that obtained desirable results concerning the variables of interest, but were ultimately terminated for other indirect reasons (Fisher, 1984). It could be discontinued for reasons we do not typically consider within the realm of behavior analysis: no room within the host agency's budget to continue the program, competition with professionals from other disciplines, scheduling concerns, a lack of support by those higher in the company structure, or a lack of cooperation from even a small group within the setting. Fisher (1984) emphasized this point when he explained that many programs may fail due to public relations, marketing, management, and special interest troubles. Possible publication biases have been suggested by others with one discussion piece asking if a "program obituary" should be included in published journals (Schwartz & Baer, 1991). Without exposing problems experienced (both directly and indirectly related to the program itself) behavior analysts may be likely to repeat them. In addition these discussions of failed studies should included potential solutions to the problems causing a program's failure.

Figure 5. Results of Intervention as described the authors. Categories are mutually exclusive.



CHAPTER 6

SOCIAL VALIDITY

Social validity was first described by Kazdin (1977) and Wolf (1978) over three decades ago. In general, they described social validity as a way to ensure behavior analysis is making socially meaningful differences. Wolf further described social validity as ensuring those affected by ABA programs are satisfied with the goals, procedures, and results obtained. Soon after the establishment of behavior analysis as a field, researchers learned the importance of social validity. In his introduction of social validity, Wolf mentions his early successes with Achievement Place. In spite of their accomplishments at Achievement Place the researchers were 'fired,' thus learning early the importance of social validity if a program is to survive. Social validity data was initially avoided due to its subjective nature. However, Wolf points out that social validity data crept into the literature from the beginning, as the founders of the *Journal of Applied Behavior Analysis* struggled to define the journal. Whether assessed or ignored, social validity affects all applied research, be it positively or negatively. Social validity has been further defined by authors as "consumer satisfaction," the second opinion of a layperson, or the likelihood consumers will seek out services (Hawkins, 1991). Kazdin describes social validity as a tool to ensure the behavior changes brought about are clinically significant, sufficiently applicable, and its outcomes sufficiently therapeutic. How does one decide what

level of behavior change is therapeutic? In addition, how do we determine our goals and procedures are acceptable to those we serve? To answer these questions, he concludes social validity data must be presented in the literature.

While it is important to define social validity as a concept, social invalidity must also be defined. Social invalidity is described as not only the behavior of consumers dissatisfied with a program, but the behavior of those willing to do something regarding their dissatisfaction (Schwartz & Baer, 1991). Perhaps most importantly, social invalidity could provide an early warning of program rejection. The resources used for a rejected program are ultimately wasted, thus representing high opportunity costs as more acceptable programs are not utilized.

In the absence of a social validity assessment, participants may let an experimenter know a program is unacceptable by other less desirable means, such as refusing to comply with procedures (Parsons, 1998). In addition, there may be more difficulty in assessing social validity with staff than with the dependent populations commonly found as research participants. Members of the dependent populations typically served may express dissatisfaction more overtly than staff (Schwartz & Baer, 1991). Dependent populations may refuse to participate when the researcher is present, shove undesired materials away, aggress towards others, or leave the area where procedures are taking place. Staff however, may wait until the researcher is not present to show dissatisfaction. They may implement only half of the procedures, quit their job entirely with no explanation, or discuss their dissatisfaction with the OBM

program with other employees (Schwartz & Baer, 1991). These more subtle methods of indicating social invalidity by staff may be missed by researchers. Additionally, an OBM program may be acceptable to the staff serving as participants, but unacceptable to agency administrators. Thus, a behavior analyst could implement an effective program, find it to be preferred by participants, and the program still discontinued by agency administrators. The potentially disastrous effects of social invalidity cannot be ignored. Low social validity regarding OBM programs in settings for those with developmental disabilities has been directly identified as a potential problem by other reviewers (Parsons, 1998). Not only could programs with low social validity fail to be disseminated, but programs already deemed successes may be terminated at a later date.

First, studies here were categorized based upon the presence or absence of social validity information. Of the studies analyzed here, approximately half (52%) presented some type of social validity information. Secondly, studies were divided by the type of assessment used. Categories were not mutually exclusive, and an article that mentioned multiple types of assessment was placed in all applicable categories. The types of assessments found amongst the studies included: Likert scales, choice measures, normative data and anecdotal information. The majority of studies providing social validity information (60% or nine studies total), used Likert scales. Fewer studies used choice measures (20%), or normative data (20%). Slightly less than half of studies presenting social validity information (40%, or six studies total) presented anecdotal

evidence. Of the nine studies using Likert scales, a slight majority of those studies (56% or five studies total) used Likert scales without additional social validity measures. Half of the studies using anecdotal evidence (three out of six studies) provided anecdotal information without additional measurement. Next, studies utilizing anecdotal information alone were removed to formulate a percentage of studies providing quantitative social validity data. Removing the three studies with anecdotal evidence alone meant that of the 29 total studies in this meta-analysis, only 41% presented quantitative social validity data. The three studies using anecdotal evidence only are not included in further analysis of social validity within this paper. The information provided by anecdotal evidence was not sufficiently detailed for further analysis. Furthermore, the main focus of the present analysis included scientific assessments (i.e., quantitative data) which could be replicated by other researchers.

Next, when social validity assessments were completed by respondents in relation to the intervention was analyzed. Social validity data is usually assessed either pre- or post-intervention (Kennedy, 1992). Within the current analysis, studies were categorized by dividing studies into two categories, pre- and postintervention assessment. The two categories were not mutually exclusive, and studies were placed in both measurement categories when multiple assessments were described. Normative data was considered both pre- and post-intervention assessment and included in both categories. Of the studies providing quantitative social validity data in this meta-analysis, 33% (four studies

total) completed a pre-intervention assessment. Every study assessing social validity in this analysis included a post-intervention assessment.

Studies were then sorted into categories based upon which individuals responded to social validity assessment. These categories included the following groups: staff, supervisors, clients, and others. Categories were not mutually exclusive and studies were placed within multiple categories when applicable. The “other” category included studies assessing the opinions of other agency employees, clients’ guardians and family members, and community members. The largest category represented was staff (89%), followed by supervisors (22%). Only one study (11%) asked clients for their opinion regarding the OBM program utilized (i.e., Huberman & O'Brien, 1999). No study was included in the “other” category.

Finally, studies were divided by the findings authors presented from social validity measures. Categories were the same as those used in the “results” section: positive, mixed, and negative reports of social validity. The results of the social validity measures completed by studies in this analysis were mostly positive (55%). Slightly fewer studies (45%) reported mixed results. No study found a program to be disliked by those responding to the social validity assessment.

Several interesting aspects emerge from this data. The most noticeable and perhaps most disappointing aspect was the number of studies assessing social validity. Fewer than half of all studies (41%) provided quantitative social validity data. As frequently as social validity and its importance are discussed,

one would expect its measurement to be more prevalent, especially in these applied settings (Kennedy, 1992). Additionally, it appears that the majority of studies relied on weaker, more subjective methods of assessment. The two largest categories represented here were Likert scales (60%) and anecdotal information (40%), two methods of assessment noted to be prone to subjective data (Kennedy, 1992). While staff may rate all components or intervention packages favorably, they may still have preferences among components (Parsons, 1998). Only three studies analyzed here used choice measures (20%) which are considered to be a stronger method of assessment. Additionally, all of these choice measures asked staff to indicate preferences on a written survey used in conjunction with a Likert scale. No author discussed the usage of a social validity measure in which staffs' choices changed the OBM program they experienced. If staff's choices had actually impacted the program, one wonders if these choices would have differed.

Anecdotal statements appeared to present very different information from quantitative data. As mentioned previously, six studies presented anecdotal information; three of them without other measures in conjunction. No study using quantitative measures presented information suggesting social invalidity. Social invalidity was suggested by anecdotal evidence, however. One study in the meta-analysis (Arco, 1991) had to be closed for further assessment due to those within the agency's dissatisfaction. In another investigation the program was discontinued at the end of the research project (Huberman & O'Brien, 1999). While yet another study (Fleming & Sulzer-Azaroff, 1992) presented quantitative

data suggesting high social validity, anecdotal information suggested social invalidity. In this investigation (Fleming & Sulzer-Azaroff, 1992) two subjects withdrew participation before the study began, citing they were 'uncomfortable' with the procedures outlined to them. These two subjects never participated in the OBM program or later completed the Likert scale presented to staff participants at the conclusion of the study. Thus, the absence of these two staff most likely skewed the results of the Likert scale data. It should be noted that not all anecdotal information provided evidence of social invalidity. Perhaps the strongest evidence of social validity includes the extension of the described program to other facilities within a program. This evidence was described by a study included in the current meta-analysis (Richman, et al. 1988). Further positive evidence presented by an anecdotal statement in an analyzed study was included in Harchik et al. (1992). In this investigation of an OBM program, state surveyors did not find any deficiencies after the program was implemented.

While no study presented an assessment regarding opinions supplied by others, (e.g., clients' guardians, community members) if anecdotal information had been analyzed, this category would have been represented. Both administrative staff and committee members were represented by anecdotal data. For example, one study (Huberman & O'Brien, 1999) explained administrative staff's request that the OBM program be terminated. In this study, Huberman and O'Brien (1999), the administrative staff found the paperwork generated by the program difficult to incorporate into their present filing system. Interestingly, this research had formally assessed social validity with two other

groups (staff and clients) and found the program to be moderately preferred. In spite of the shown efficacy of the program and encouraging social validity data, the program was terminated at the end of the study per request of the administrative staff.

Anecdotal information also provided preintervention evidence of social validity, when relatively few quantitative assessments (33%) did so. One study analyzed here (Arco, 1991) did provide anecdotal information which could be considered a pre-intervention social validity assessment. In the Arco investigation, the researchers and staff discussed which skills staff felt it would be beneficial for the clients to possess. This discussion led researchers to target client-to-client interactions in an effort to improve client social skills. The authors describe this as an informal pre-intervention social validity assessment of the goals of a program.

Interestingly, only one study (Huberman & O'Brien, 1999) assessed social validity amongst client participants. This could be due in part to differences among the clients themselves. Clients in the Huberman and O'Brien study appeared to be of higher functioning than those participating in other programs analyzed here. While Huberman and O'Brien included patients with chronic mental illness as clients, all others included those with developmental delays. Most of the studies describing clients with developmental delays mention various levels of mental retardation, most commonly severe to profound. Huberman and O'Brien gave examples of the social validity data solicited from clients. Such questions included: how bothersome the procedures were, if they felt their

therapist's work improved, and if they thought these procedures could help others. These questions suggest clients were of higher functioning. One meta-analysis (Fleming & Reile, 1993) found the majority of clients included in the type of research analyzed here were of lower functioning. Fleming and Reile (1993) found the majority of authors to report severe to profound mental retardation. Most likely, using a traditional Likert scale would not be practical with the clients being served in these studies. The same would hold true for many other assessments (e.g., choosing intervention components on a paper-and-pencil assessment, verbal surveys).

Several implications could be drawn from the above data. First and foremost, too few studies assess social validity. Social validity has been characterized as one of the reasons many behavior analysts enter the field and as one of the most important tools to help those served. The landmark article first describing social validity referred to it as "how applied behavior analysis is finding its' heart" (Wolf 1978). If social validity is to hold this importance, behavior analysts must ensure it is measured frequently and properly. Less than half of all studies here (41%) quantitatively measured social validity and slightly over half (52%) mentioned it in some way. This finding is consistent with the findings of other reviews of the behavior analytic literature (e.g., Kennedy, 1992). The data presented here does suggest a slight improvement over the Kennedy (1992) meta-analysis, with rates of studies completing a social validity assessment somewhat higher here. Kennedy found 20% of articles reported social validity measures and 91% of those 20% utilized quantitative data. It

should be noted however, Kennedy's analysis included a much wider scope of behavior analytic literature (all empirical studies within the *Journal of Applied Behavior Analysis* from 1968-1990 and *Behavior Modification* from 1977-1990).

Another possible implication from the data presented here is the high social validity of most OBM programs in these settings. Unfortunately, there is also the possibility that some of these findings represent false positives. Many explanations may be responsible for false positives. First, subjects may be "faking good." The potential problem of participants "faking good" has been experienced and discussed by behavior analysts and test-developers in other psychological fields (e.g., Schwartz & Baer, 1991). Very little discussion of the prevalence of "faking good" appears in behavior analytic literature, or on how to avoid the problem. Second, a disconnect may exist between subjective measures and actual behavior (Reid & Whitman, 1983) allowing for more false positives within the data. For instance, staff may respond favorably on a Likert-type scale, but later fail to implement the procedures of an OBM program. In one of the flagship articles introducing social validity, Wolf (1978) warns that subjective data may not accurately predict quantitative measures of behavior. Wolf goes on to present three reasons for this disconnect. First, he highlights the impossibility of collecting interobserver agreement. Second, participants may be responding to some changes in behavior not being recorded. Last, Wolf states, individuals may not know their own situation accurately, and therefore be unable to report their true impressions.

One reason for a possible disconnect between actual behavior and subjective measures might be the Likert scales themselves. Questions on a Likert scale may not evoke the type of information they were designed to collect. The term “satisfaction” may mean different things to different consumers (Hawkins, 1991). An instrument designed to assess the acceptability of a program may more accurately measure staff attitudes towards the individuals implementing the program rather than the program itself. Additionally, there may be questions absent from questionnaires that should be present. Perhaps, an intervention is acceptable but the cost of its implementation is not. In these particular settings and variables of interest (e.g., staffs’ usage of ABA procedures) researchers should address satisfaction on two levels. It is recommended that social validity be assessed both with the ABA techniques used by staff and the OBM program ensuring that usage.

Additionally, the data generated by this meta-analysis suggest researchers over-rely on post-intervention measures. No study here took pre-intervention data only, as all studies included post-intervention measurement as well. Only one third as many studies assessed post-intervention social validity as those assessing pre-intervention. This finding replicates similar findings of related meta-analyses. Kennedy (1992) also found the majority of studies only assessed post-intervention, with 155 of 198 reviewed articles assessing post-intervention social validity. While post-intervention measurement can be crucial to behavior analytic research, it should not be done to the exclusion of pre-intervention measurement. Both pre- and post-intervention assessments should

be completed when applicable. Researchers should also assess social validity during the implementation of an OBM program if appropriate and useful.

However, behavior analysts should take care not to assess too frequently, lest they annoy those responding to the assessments.

Several problems may result from a lack of pre-intervention assessments. First, pre-intervention assessments provide an early warning sign of procedure rejection (Hawkins, 1991). Second, these measures allow one to more closely examine the acceptability of the goals of a program. If program participants are included in the selection of program goals it may give them more “ownership” in the program and thus increase the chances they will work towards a program’s success (Schwartz & Baer, 1991). Finally, the relationship many researchers hold with an OBM program’s host agency ends once the research program is complete. If an assessment is completed only as the researcher is leaving the agency this may present the impression that the researcher is unconcerned with the results of the assessment. Participants’ suggestions must be used if they are to learn their feedback is truly important. If not, participants may learn this assessment is a mere formality, their feedback is not actually needed, and behavior analysts are deceptive (Schwartz & Baer, 1991). For this reason, it is hypothesized that consumers filling out post-intervention questionnaires may be less likely to accurately record their responses. This problem may continue to result in false positives amongst social validity data.

Perhaps the most important suggestion based upon this research is to increase the use of social validity measurement in general. If it is to become a

hallmark of behavior analysis it must be treated as such. The fact that less than half of the studies reviewed here presented quantitative social validity data is concerning. This may be especially problematic, as this information could indicate clues as to the lack of OBM adoption in human service settings (Parsons, 1998). However, as noted by others (e.g., Schwartz & Baer, 1991) if all behavior analysts used current social validity procedures more frequently applied behavior analysis would still not progress. Behavior analysts must ensure the methods used are reliable and valid. One of the earliest descriptions of social validity warned that face validity alone was not enough to ensure an assessment device was useful (Wolf, 1977). Other types of validity and reliability, (e.g., test-retest reliability) must be present in any scale a researcher constructs. Many authors offer advice on the construction of social validity assessments. For example, Fawcett (1991) offers ten general procedures in constructing an assessment. Hawkins (1991) discusses methods of conducting both subjective (e.g., how to construct a Likert-type scale question) and objective assessment (e.g., the usage of normative data). Researchers should familiarize themselves with the advice of these and other reviewers and apply such information to their work.

A publication bias may exist within social validity data, as program developers appear to assume a positive social validity assessment is needed for publication (Schwartz & Baer, 1991). Researchers may implement a program, obtain desired behavior change, but find low social validity. These researchers may then decide against publishing altogether or omit social validity data from

the published article. Schwartz and Baer (1991) have suggested publishing more studies with poor social validity (i.e., social invalidity). This recommendation is made with the hopes that researchers will have a rich sample of social invalidity to compare with published examples of social validity. This recommendation is similar to the advice made in the results section of the current paper. More “failed” studies should be published, both amongst social validity data and among the results of the independent variable.

Few studies in this analysis asked staff to rank the components of an intervention by preference. For example, in one included study by Green et al. (1991) experimenters found overall staff acceptance of the management program. However, when the data was further analyzed clear preferences emerged. While two of the three components still averaged on the “like” side of the scale, one component did not. Had these researchers only completed an assessment for the general program, this important information would have been missed. No study allowed staff to choose an intervention, where researchers then implemented the chosen components. Another study in this analysis (Fleming & Sulzer-Azaroff, 1992) did ask staff which components they would be most interested in continuing. Fleming and Sulzer-Azaroff found all components to be acceptable by staff, but clear preferences among the components were observed. Unfortunately, they did not indicate those choices were later implemented, or if the lesser-preferred components were discontinued. It is unknown if the paper-and-pencil measures would have corresponded with staff behavior. As mentioned previously, staff may have more subtle ways of

indicating a program is not favored, and staff may continue these more subtle methods of expressing discontent while completing formal assessments positively (Schwartz & Baer, 1991). Therefore, it is recommended that researchers include more choice measures within the social validity research. Additionally, when possible, participants' choices should affect the OBM program.

While usually considered the weakest type of evidence within behavior analysis, anecdotal information may be preferred over no social validity discussion. Anecdotal evidence should be presented if researchers were unable to implement formal measures or if anecdotal evidence would strengthen existing formal measures. As previously mentioned, state surveyors did not find additional deficiencies (Harchik et al. 1999) in a study included in the current meta-analysis. The authors presented this information in addition to Likert-type scale measures. In this example, the state survey would have occurred regardless of the implementation of the OBM program. While some other factor may have been partially responsible for the successful state survey, this anecdotal evidence does strengthen the authors' hypothesis that social validity was present. While, this information may not have been necessary for the authors to include in a report of their research, this naturally-occurring data should be included when possible. Additionally, this type of data could be helpful in assisting practitioners when selecting an intervention to use in their work. It is hypothesized here that anecdotal evidence within an article might provide a

solution to a problem experienced by a practitioner. As well, it may present an outcome the practitioner desires in their own work.

While the main focus of the studies examined here is staff training, this training ultimately affects clients served. Only one study (Huberman & O'Brien, 1999) assessed client's attitudes towards an OBM program and their therapist's performance. As discussed, this may be due to differences with their client population. Increased social validity assessment with the clients served by an agency should be done. Likert-type scales are currently one of the most popular methods of measuring social validity (Parsons, 1998). However, these methods are difficult to use with clients with severe cognitive delays. Alternative methods of assessment need to be developed if behavior analysts are to fully assess social validity with those served, including those with developmental or physical disabilities.

Another suggestion, based upon the information presented here, is to include more pre-intervention measures. This would allow researchers to examine the social validity of their goals in advance. Thus, pre-intervention assessments could allow behavior analysts the opportunity to edit programs and discard procedures likely to be met with resistance, thus preventing future problems. Pre-intervention assessments could also provide an opportunity for researchers to educate others regarding OBM programs (Hawkins, 1991). Participants' negative opinions concerning specific procedures could change if they receive further education and information. Additionally, pre-intervention measures could be used in conjunction with post-intervention measures to see if

perceptions changed over time, and perhaps as a result of the studied intervention. One investigation in this analysis (Williams et al., 2002) collected staffs' opinions regarding client training both before and after the intervention phase. After the intervention was applied, staff felt they knew how clients were progressing, they were able to express their opinions regarding client training easier, and they believed client training was more effective. In this way, pre-intervention social validity measures could support the researchers' assertion that the OBM program was responsible for this change.

In addition to recommendations regarding how social validity is measured, how findings are discussed should also be examined. In most behaviorally-orientated journals the effects of the intervention on the dependent variable are discussed, along with hypotheses on why those results were obtained. The same should be done for social validity data. Rather than merely presenting data and drawing a conclusion, authors should present hypotheses as to why social validity assessments yielded these results. One study reviewed here (Azrin & Pye, 1989) did present some discussion why the program they used could have been acceptable. In this study, the authors described a behavior contracting procedure to ensure a staff member was correctly implementing client training programs. They explained the behavior contract may have been acceptable as the staff member was able to choose both her standards of performance and the reinforcers earned. However, Azrin and Pye present these reasons as why behavior contracting in general is useful, and why it may have been successful in

their investigation. The term social validity is not used, nor do they provide other evidence of social validity.

Furthermore, it is recommended that researchers provide discussion regarding programming for social validity in advance and troubleshooting should problems arise. Nevertheless, researchers need to ensure they are actually manipulating the acceptance of their procedures and not the verbal behavior of participants (Fawcett, 1991). Additionally, it is suggested that the results of the general investigation and social validity assessment be compared. While social validity is vital to behavior analysis as a whole, it must be assessed frequently, and as accurately as possible to maintain this high standard of applicability (Kennedy, 1992). And finally, the resulting conclusions should be used for the benefit of all those affected by an intervention.

Figure 6. Type of Social Validity Assessment: Categories are not mutually exclusive.

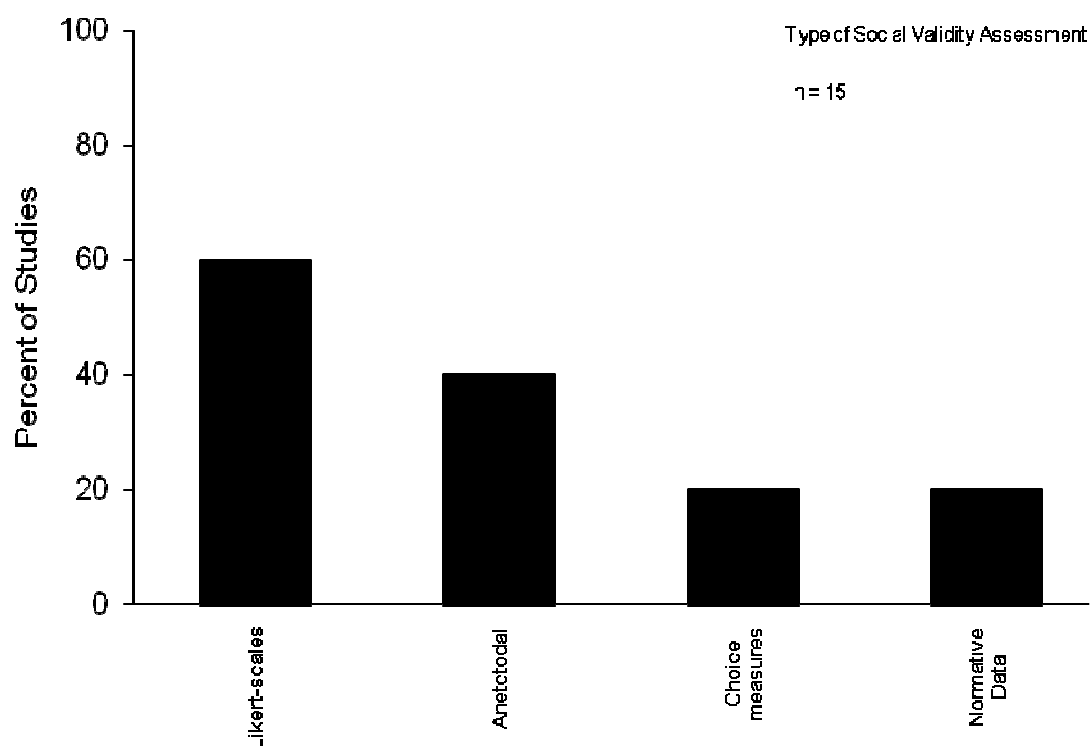


Figure 7. When Social Validity is assessed in relation to the intervention amongst studies in this meta-analysis.

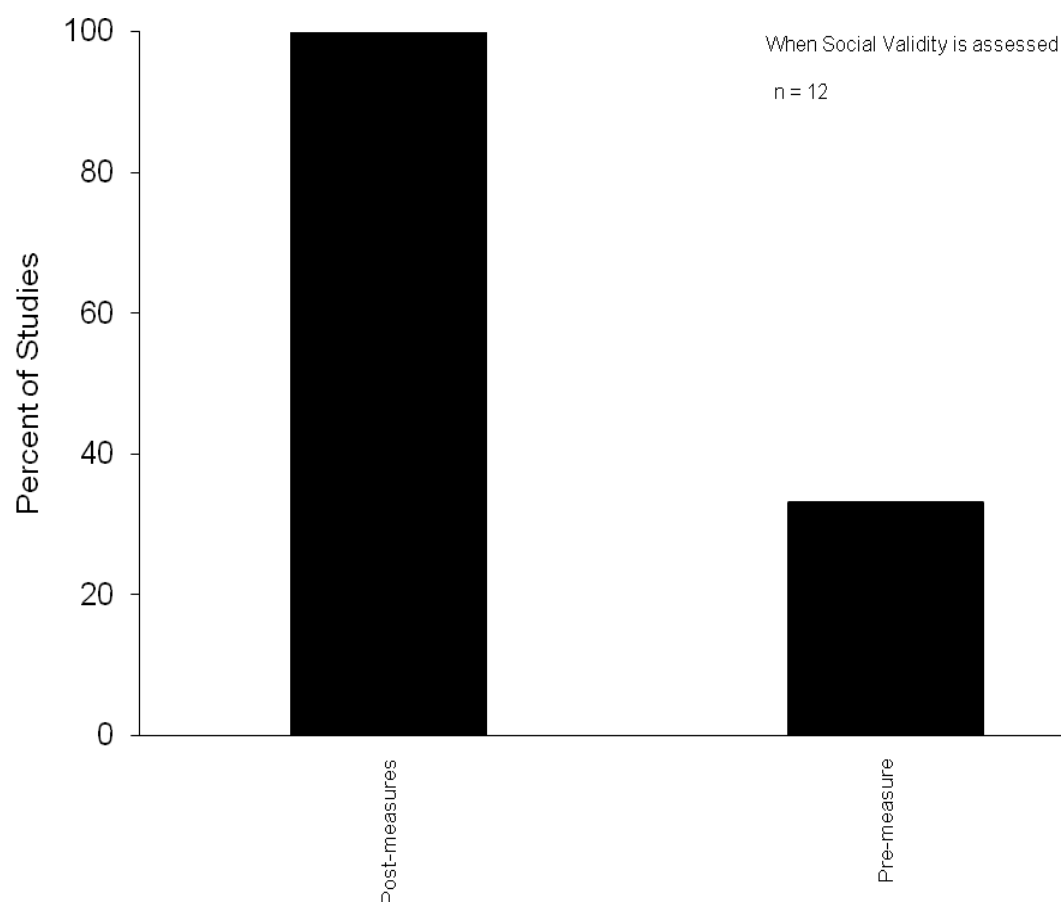


Figure 8. Respondents to Social Validity Assessments:

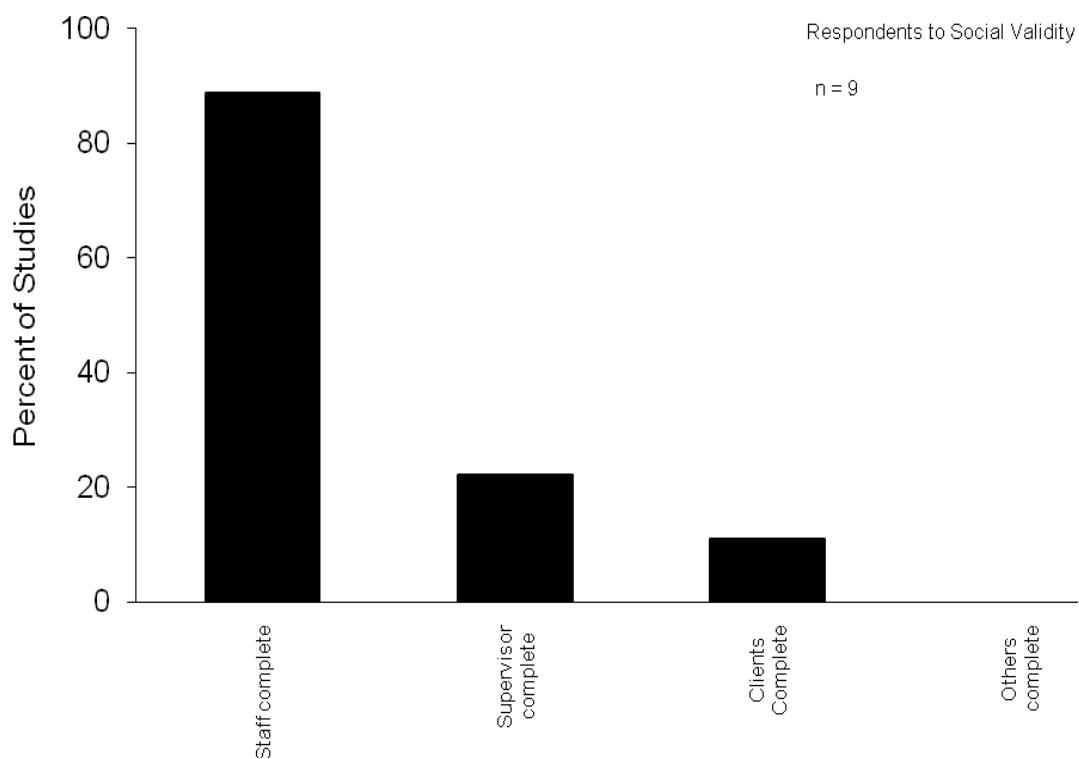
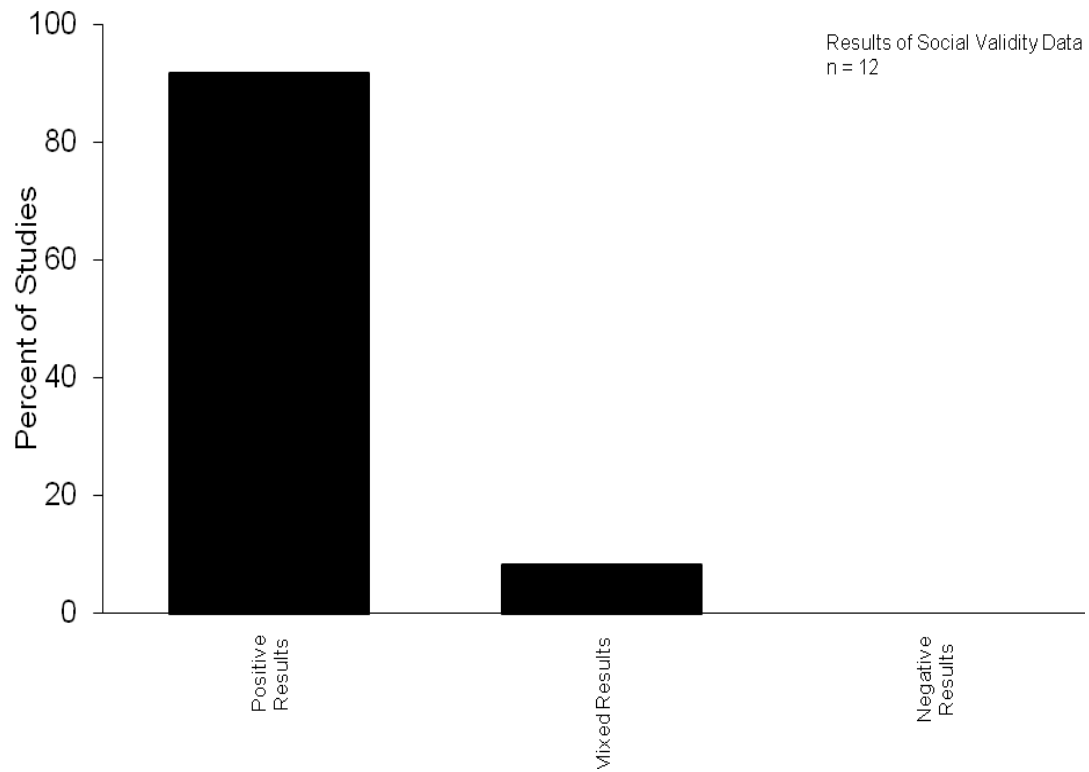


Figure 9. Results of Social Validity Data: Results obtained by social validity assessments within this meta-analysis.



CHAPTER 7

MAINTENANCE

The analysis presented here of the results obtained by researchers (both of social validity measures and of dependent variables) are very encouraging. Data obtained by this analysis suggests OBM programs in these settings are both effective and liked by participants. However, if these programs are to be useful, behavior analysts must ensure an OBM program is practical in real work settings. To ensure practicability, practitioners must ensure the success of a program for an extended period of time. To accomplish this behavior analysts need to ensure behavioral maintenance is being addressed. Like social validity, maintenance is often not the main focus of much behavior analytic research (Boyce & Geller, 2001). Yet both are crucially important in applied settings.

Maintenance is essential as all outside influence by others (e.g., consultants, students, and interns) must end at some time (Sigurdsson & Austin, 2006). Programming for maintenance assists in the transfer of control to the natural environment (Boyce & Geller, 2001). Many programs analyzed here were multi-component programs. These programs may require more time and effort than possible for a human service agency to sustain long-term. Additionally, if positive results were to continue in the absence or reduction of a program component, it is obviously advantageous to eliminate an unneeded component. For the above mentioned reasons maintenance is important to behavior analytic literature.

There appears to be some confusion on the term “maintenance” within the behavior analytic literature. To be considered maintenance some portion or all of the intervention must be removed (Cooper, Heron, & Heward 2007, p. 699). Others have described maintenance as the removal of contrived contingencies and a return to the natural environment (Boyce & Geller, 2001). Maintenance is not the continuation of a program over a lengthier period of time. While compiling this meta-analysis maintenance was perhaps the most difficult category to complete.

Amongst the current literature it may be difficult to discern when the author is presenting maintenance data or generalization data. For example, some studies may conduct classroom-type trainings and use the term “maintenance” when employees later perform the behavior in the work setting. In this instance the above example would be a demonstration of behavior generalization (i.e., the employees are demonstrating a learned behavior in a new setting). Adding further confusion, authors may use the term “follow-up” rather than “maintenance” when describing maintenance data (e.g., Fleming & Sulzer-Azaroff, 1992; Ford, 1984).

First, studies were organized by the presence or absence of maintenance measures. As described above, many authors did use the term “follow-up” when the information described met the definition for maintenance data. In these cases the presented data was considered maintenance and included in the current review. Of the studies in this review, 12 studies (41%) reported maintenance data. Studies were then categorized by the time range for which studies

collected maintenance data. The three categories used were: less than six months, six months to one year, and more than one year. The majority of studies assessed maintenance for less than six months (58%). Three out of twelve studies (25%) assessed for six months to one year. Fewer studies assessed maintenance (17%) for more than one year.

The data obtained by researchers within maintenance assessments was also reviewed. Results were divided by the following categories: positive results (i.e., behavior did maintain), mixed results, and no results. If maintenance data presented by a study was generally positive but included a high amount of variability, this study was included in the mixed results category. Six (55%) studies presenting maintenance data reported positive results. Fewer investigations (42%) found mixed results, and only one study (8%) reported a failure of behavior to maintain (i.e., Harchik et al., 1992).

Unfortunately, this meta-analysis of maintenance data found a similar result to the analysis of social validity data; relatively few studies completed this type of assessment. Perhaps the most noticeable pattern to emerge from this data is the inverse relationship between length of maintenance assessment and the number of studies assessing it for that period of time. Additionally, no study examined behavior over an extended period of time (i.e., multiple years). Over half of all studies assessing maintenance, did so for six months or less. Only two studies (of the 29 total analyzed here) examined maintenance for a year or more. One study (Parsons et al., 1989) presented data for 15 months, while the other (Parsons et al., 2004) presented data for 64 weeks. These two general findings

(the absence and relatively short time frames of maintenance assessments) are similar to the findings of past literature reviews (e.g., Reid & Whitman, 1983). When analyzing the research on behavioral staff management strategies, Reid and Whitman (1983) found less than 25% of articles provided maintenance data. Reid and Whitman also found studies generally assess maintenance for a relatively short period of time. The length of maintenance assessments found by these reviewers included 2, 4, 6, 7, 9, 11, and 19 weeks. Thus, in the Reid and Whitman analysis, no study assessed maintenance for longer than six months.

Several studies in the current meta-analysis (e.g., Durcharme & Feldman, 1992; Fleming & Sulzer-Azaroff, 1992) reported difficulty collecting maintenance data due to attrition resulting from staff turnover. Turnover has long been noted as a problem in these settings (Sturmey, 1998). Further research has shown higher turnover rates in community-based settings (e.g., Sturmey, 1998; Williams & Lloyd, 1992). Thus, problems associated with turnover may further increase as services move from institutions to smaller community-based settings. Additional studies in the present meta-analysis may have been unable to provide maintenance data due to turnover, but did not mention this problem directly. Additionally, those studies assessing maintenance within this analysis may not have included all employees initially trained due to attrition. It does seem possible that lower-performing employees would be more likely to leave an agency, either voluntarily or involuntarily. This increases the likelihood that lower-performing employees were unavailable for maintenance assessment. As a result of the underrepresentation of these lower-performing employees, the

data obtained by maintenance assessments may have been skewed in a positive direction.

Based upon the above data and discussion, it is suggested that researchers should describe maintenance data more consistently. Maintenance data can be mislabeled multiple ways. As discussed, the term “follow-up” may be used in place of “maintenance.” Furthermore, maintenance phases may be labeled with a description of the contingencies, such as “Feedback only”, or a letter designation, such as “Phase B” (e.g., Harchik et al., 1992). While these methods of labeling may accurately describe the information presented, it should also be made clear this data qualifies as a maintenance assessment.

Researchers should also clarify when no changes in an intervention are made but it is continued long-term. If the intervention being continued long term is sustainable and needed, perhaps a partial or complete withdrawal is not needed. In these instances authors could note that maintenance data, as defined by behavior analytic literature, is unnecessary.

Less than half of the studies within the present analysis included maintenance data, a disappointing finding. As discussed previously, other reviewers have reported similar findings with Reid and Whitman (1983) reporting less than 25% of studies in their review providing maintenance data. A lack of maintenance research may be partially responsible for the lack of OBM in adult residential settings (Williams & Lloyd, 1992). Agency administrators may be less likely to seek out services if they perceive OBM programs as temporarily beneficial, especially if associated costs are high. They may also feel a reliance

on outside help negates any benefits received (Sigurdsson & Austin, 2006).

Problems may also arise when an in-house behavior analyst fails to implement long-term programs. Others within an agency may hesitate to implement further programs proposed by the behavior analyst in this example.

Based upon the above data and discussion, future researchers should include maintenance assessments more frequently within their research. While more maintenance assessments in general are needed, long-term assessments appear to be especially rare. The lack of long-term demonstrations has been identified as a problem by other reviewers (e.g., Christian, 1984). In his discussion, Christian describes long-term maintenance assessments as those lasting approximately 3 to 5 years, not the more common 6 to 12 months. As mentioned previously, no study in this meta-analysis presented maintenance data for multiple years. Christian (1984) claimed that long-term maintenance assessments were lacking 26 years before the current meta-analysis was completed. Thus, this finding is especially discouraging as it appears little advancement on Christian's recommendations has occurred.

Explanations for the high prevalence of short maintenance assessments should be explored as well. One possible reason for this problem is the contingencies controlling researcher behavior favor short maintenance assessments. Researchers may be rewarded more for showing a large effect or strong functional relationship than for showing behavior maintenance (Boyce & Geller, 2001; Schwartz & Baer, 1991). Practitioners and researchers alike may be reinforced for abandoning a program once the problem has been "fixed" and

encouraged to address new concerns (Williams & Lloyd, 1992). Some researchers may have access to a setting for a limited time, such as student researchers. Thus the research ends when the student's time in the setting does. Finally, researchers may continue a successful program long after the research article describing it has been published. What appears to be a short-term demonstration may actually be a program with years of successful implementation. While long-term maintenance demonstrations are helpful, it is not recommended here that they be completed to the exclusion of short-term demonstrations. Obviously, a short-term assessment is preferable to no assessment. Short-term assessments can be beneficial in that they show maintenance is possible in the absence of at least one intervention component. How long behavior maintenance can be achieved may then be a question for further review.

Similarly to the recommendation made regarding social validity, not only should maintenance assessments be completed, more discussion of maintenance within research articles might be necessary. Discussion as to why researchers believe behavior maintained (or failed to maintain) could be useful. As long-term maintenance is often not the focus of many articles, researchers may omit useful details (Boyce & Geller, 2001). It is suggested that researchers include these details when possible, even if maintenance is not a major focus of a given study. Based upon the presented data, it is also recommended that studies where behavior failed to maintain be published. Only one study included here reported a failure of behavior to maintain (i.e., Harchik et al., 1992). The

suggestion for more “failed” maintenance is similar to the suggestions made for studies with less desirable results and social validity data (Schwartz & Baer, 1991). Behavior analysts need a rich sample of “failed” maintenance assessments to compare with their successes. In other words behavior analysis needs to “learn from its mistakes.”

Although few studies included here provided any maintenance information, one investigation (Harchik et al., 1992) did address it as part of the study’s main focus. Interestingly, this study was also the only study to report a failure of behavior to maintain. Harchik et al. (1992) systematically manipulated the independent variable with the intention of providing a clear maintenance assessment. In this study, the researchers utilized an ABAB reversal design to investigate the effects of a consulting process on several staff behaviors. These researchers found ongoing consultation was needed for behavior to maintain at desired levels. This information could be vital for any practitioner or researchers attempting to implement a similar consulting model. Further investigations with a major focus on maintenance are needed, including those with a more complete analysis of its contributing factors (Boyce & Geller, 2001). More complete analyses should include manipulations to determine how much of what program components are necessary. This information could be further dissected into the amount of intervention needed to obtain acceptable levels of behavior and the amount of needed for behavior to maintain at its’ highest levels.

Many authors (e.g., Christian, 1984; Reid, 1998; Sturmey, 1998) have reported a need for larger organizational changes in order to foster an

environment more supportive of OBM programs. Organizational changes may be especially important as they may create a supportive environment for OBM programs long-term or in the absence of the experimenter (Sigurdsson & Austin, 2006). Research on the types of organizational changes that facilitate maintenance is needed. A discussion of the various potential contingencies behavior analysts face could be a beginning. As mentioned previously, more reports of failed systems should be published, along with the problems those failed programs faced. In this way behavior analysts could learn what caused previous programs to fail and avoid the same pitfalls.

Figure 10. Results of Maintenance Assessments within the studies.

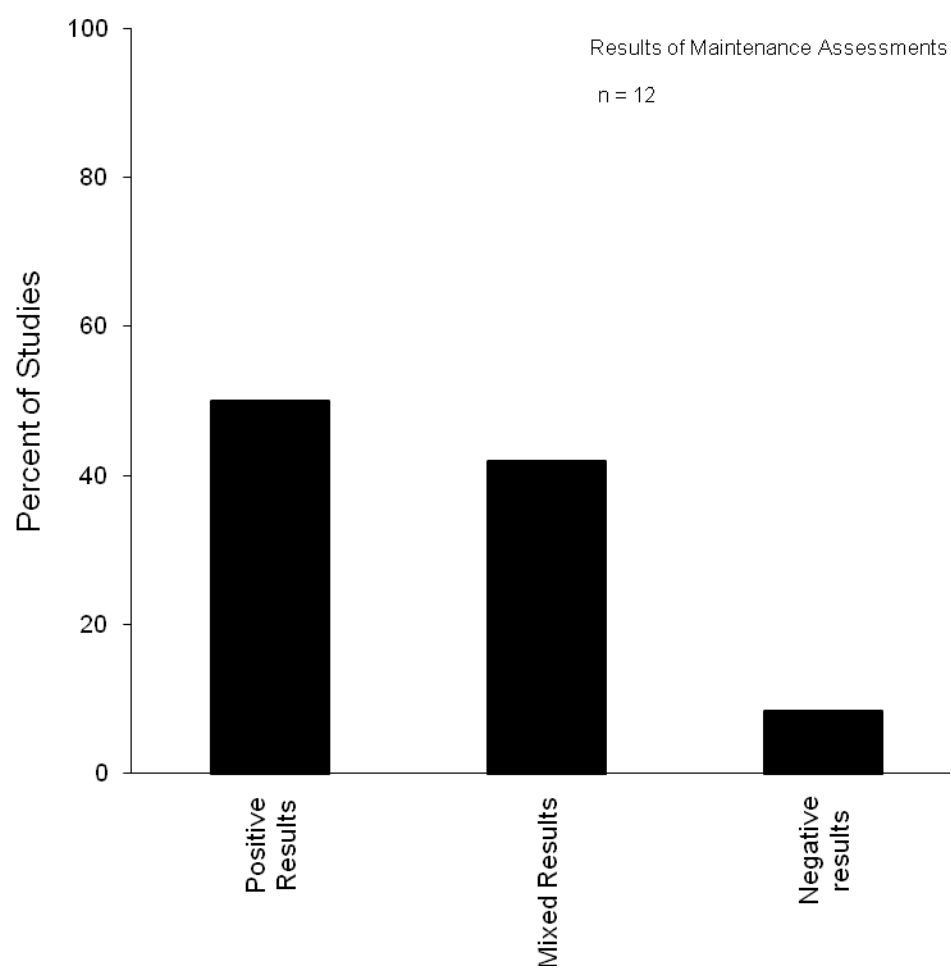
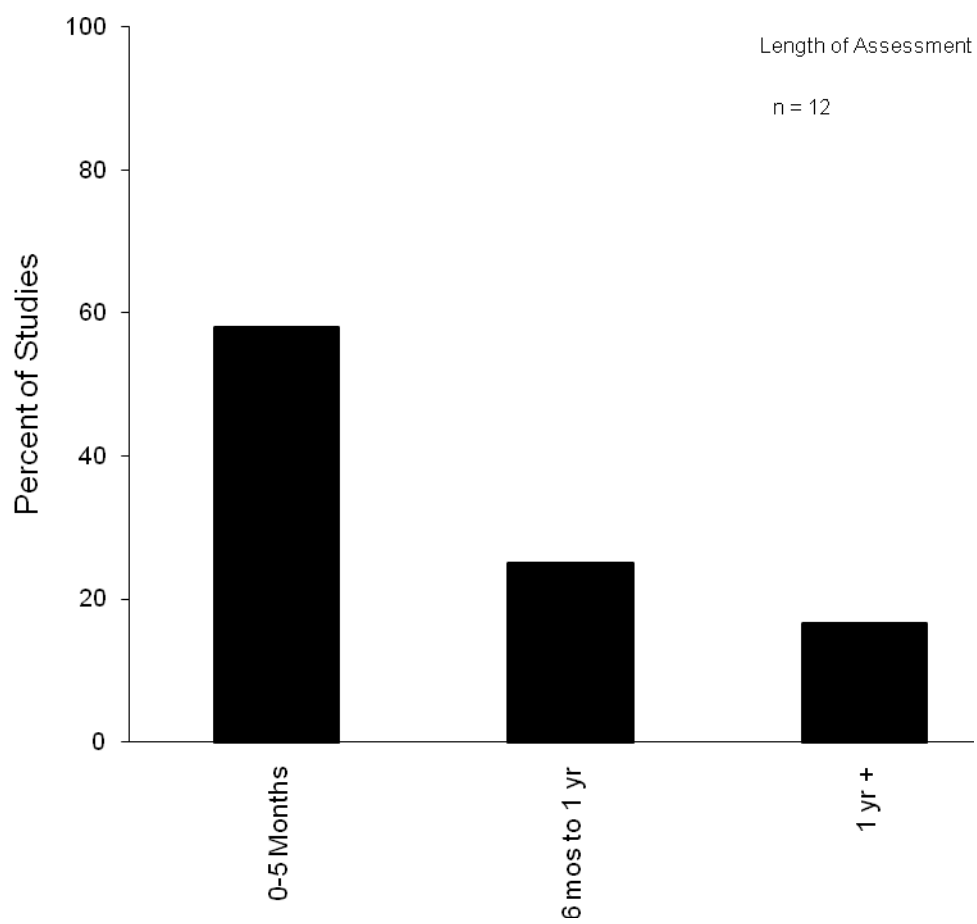


Figure 11. Length of Maintenance Assessment:



CHAPTER 8

CONCLUSION

Several overall themes emerge from this meta-analysis. Behavior analysts appear to be quite successful in utilizing OBM programs to ensure the usage of ABA techniques in adult residential settings. The majority of research analyzed presented favorable outcome data (i.e., favorable dependent variable measures, high social validity, and favorable maintenance data). However, these encouraging results have not led to the wide-spread adoption and dissemination of these programs. This paper presents some possible explanations regarding the lack of OBM in these settings. Although behavior analysts have been successful, the validity and reliability of the data presented by these studies should be questioned. This may hold especially true for social validity data, due to its subjective nature and the possibility of false positives. Additionally, the low percentage of studies presenting maintenance data is a problem. This problem should prompt behavior analysts to question the utility of many OBM programs to practitioners in real work settings. However, it should be noted that this paper examined a relatively specific sample of the ABA literature (i.e., the use of OBM to implement ABA techniques with clients in adult residential settings). Nonetheless, many issues presented here may apply to other areas of research within behavior analysis or even to the field as a whole. Further meta-analyses could be completed to replicate (or fail to replicate) the findings of this analysis. It may be helpful to see if the findings here (e.g., few studies presenting

implementer details, a high possibility of false positives amongst social validity data) are common within other samples of the ABA literature.

Several categories within the current meta-analysis were difficult to analyze. One reason for this difficulty was the lack of detail available in many articles. Authors must ensure descriptions of OBM programs are sufficiently detailed if a program is to be replicated, disseminated, or studied further. More complete information could be especially useful within two categories: maintenance and implementer characteristics. Furthermore, the lack of information presented in some articles may have skewed the data obtained by the current meta-analysis. Although “implementer characteristics” applies to all studies here, only 23 of the total 29 provided any information regarding implementers. Many studies here were excluded from further analysis within various subcategories due to a lack of sufficient details. Thus, it is recommended that authors of research articles describe their research as sufficiently as possible.

Unfortunately, space limitations imposed by the publishing journal may present a problem in providing more detail within a research article. More discussion on the amount and type of details most useful to the consumers of the research is needed. This discussion could provide a clear guide on the type of information most crucial for inclusion. Obviously, an author cannot include every possible detail in a journal article. Although details were often lacking regarding maintenance and implementer characteristics, authors generally did give complete information in other categories (e.g., results of interventions, staff

characteristics). It is suggested that authors continue to provide this information. Adding further confusion when reviewing the literature, information regarding some categories analyzed here (i.e., maintenance, implementer characteristics) was often scattered amongst different sections of an article. For example, one author may describe the program implementer in the “method” section of a journal article. Another author however, may describe the implementer in the “discussion” section while addressing concerns of possible reactivity.

While a lack of various types of details is a problem, the terms used to describe these details may also be problematic. Frequently, multiple terms were used to describe similar information among different articles. For example; “consumer satisfaction” and “acceptability measures” were used to describe social validity measures. “Follow-up” and letter designations were used to describe maintenance phases. Finally, “experimenters” and “data collectors” are two examples of the terms often used to describe program implementers. In this last example, these terms can lead to confusion regarding these individuals’ roles implementing the OBM program, and the amount of contact had with study participants. It is recommended experimenters/authors use standardized terms presented in the behavior analytic literature to avoid confusion. For example, using the term “maintenance” instead of “follow-up” when the information described meets the definition of maintenance as presented by behavioral analytic literature.

Interestingly, the data obtained amongst many different categories was similar. This is especially true of categories pertaining to program outcomes (i.e.,

results, social validity, and maintenance). The majority of studies obtained positive results when data within these categories was collected, followed by those reporting mixed results. It seems unlikely behavior analysis can achieve this level of near-perfection. These results provide further evidence of a publication bias. Publication biases have been discussed in the literature, generally in terms of the results obtained by an intervention. However, it seems a publication bias may apply to other areas of the literature, such as maintenance and social validity. As discussed, the main focus of many authors may be showing a treatment effect. This focus may be a contributing factor for the occurrence of publication biases. If a large treatment effect is not shown, researchers may decide against publication. One reason for the emphasis on treatment effects could be a result of many authors being trained as researchers, not practitioners (Schwartz & Baer, 1991).

Perhaps, the most important point to be emphasized here pertains directly to the lack of OBM in human service settings. The overall positive findings by most articles (i.e., within study results, social validity assessments and maintenance assessments) provide further evidence that obtaining desired results is not enough. If OBM is not utilized in these settings, these encouraging results simply do not matter. As discussed in the introduction, a lack of OBM in human service has been identified and discussed for a number of years. Many authors have presented hypotheses as to why OBM is lacking. In spite of this attention, the problem still exists. The average consumer or researcher may be unable to address hypotheses presented by others in a meaningful way. For

instance, one discussion piece (Williams & Lloyd, 1992) proposed behavior analysts implement their procedures within agencies willing to take 'a step ahead.' Nonetheless, how a practitioner identifies such an agency is not discussed. Further research on the lack of OBM in human service settings and concrete plans to address noted problems are essential.

A restricted scope of many OBM programs may further contribute to the absence of OBM in human service settings (Christian, 1983; Fleming & Reile, 1993; Reid 1998). Various elements of this meta-analysis do provide evidence for this problem. A restricted scope may refer to several issues including but not limited to: few staff participants, a lack of long-term studies, and relatively restricted dependent variables. Two of these areas are addressed in the present meta-analysis (i.e., number of subjects and the length of maintenance assessments). The current analysis frequently found short maintenance assessments and few staff participants within analyzed studies. It appears the criticism of OBM as offering too small a scope to many practitioners may be justified, as evidenced by the number of participants and length of time typically evaluated by these programs.

Dependent variables of included studies were not directly examined by this meta-analysis. Therefore, it is difficult to draw concrete conclusions concerning the scope of these variables. However, evidence of narrow dependent variables was present. Several of the included studies could be used as examples of investigations with relatively narrow dependent variables. One such study (Sigafoos et al., 1992) trained staff to increase the use of a time delay

and reinforcement strategies when teaching clients a sandwich making task. Another example included here, (Reid et al., 2003) trained staff to provide choices to clients regarding work tasks. It should be noted however, that the inclusion criteria of this meta-analysis may have favored studies with a narrowly defined dependent variable in some way.

As mentioned previously, only the data related to staffs' usage of ABA is analyzed here. However, several studies did provide evidence of wider-scope dependent variables. One study, Harchik et al. (1992) included a wide scope of staff behavior, including several measures of staffs' usage of ABA. This study divided the dependent variables of interest into three main categories, two of which were related to ABA. Within these categories were two to four secondary dependent variables (e.g., the number of times tokens were exchanged for backup reinforcers, the amount of time staff interacted with a client). This study, (Harchik et al., 1992) provides not only an example of a study with a wider-scope, but a study examining several facets of staff's usage of ABA. Another example of a study examining a broad range of staff behavior included Harchik et al. (2001). In this investigation researchers examined 17 staff competencies. While many of these competencies were related to staff's usage of ABA (e.g., teaching techniques, the proper use of punishment procedures) many were not (e.g., handling medication-related incidents, knowledge of corporate policies and procedures). Studies such as these may provide a model for others looking to make sweeping changes within an organization or address multiple staff behaviors.

The somewhat narrow focus of OBM in settings serving those with developmental disabilities has been noted by other reviewers (e.g., Fleming & Reile, 1993; Reid, 1998). Fleming & Reile specifically praised one of the studies in their review as a 'landmark' investigation (Parsons, Schepis, Reid, McCarn, & Green, 1987). Parsons et al. (1987) was praised for investigating behavior both long-term and including a large number of participants (i.e., across 21 classrooms) two areas found lacking in the current meta-analysis. This study (Parsons et al., 1987) was considered for inclusion within this meta-analysis as it was identified through one of the PsycInfo searches utilized. However this study was excluded on the basis of its setting (having not taken place in a residential setting for adults). Studies such as Parsons provide a model of a successful long-term, large-scale OBM program. Successful programs such as these should be replicated in different settings, including residential settings with adults. To build successful OBM programs, researchers should look to all settings in which previous OBM programs have been successful. It is recommended here that researchers look to implement wide-scale OBM programs in human service settings. However, this type of research would be very complex and labor intensive. Wide-scale research should complement smaller-scale OBM demonstrations, not replace them (Reid, 1998).

In conclusion, behavior analysts have made many strides in this area of research. While there is still much work to do, it is obvious OBM is useful in these settings and as a tool to further the advancement of ABA. The fact that many investigations were able to obtain meaningful differences in applied

situations is encouraging. It is also encouraging to find many authors addressing the lack of OBM and providing helpful discussion. While much of the current research is encouraging, holes do exist. Further research is needed within long-term wide-scale OBM programs. Research should also examine publication biases, how they may affect multiple types of data (e.g., intervention results, social validity assessments), and how they might be addressed. Further research may want to address the validity and reliability of many assessments used. This type of investigation may be especially needed within social validity assessments.

To ensure the survival of OBM, behavior analysts should continue research in these settings. Concerns listed here, both by this meta-analysis and by other analyses should be addressed. Further research should work to uncover other barriers to the widespread adoption of OBM. To do so would ensure the adoption and continuation of organizational behavior management.

REFERENCES

References marked with an asterisk indicate studies included in the meta-analysis.

- *Adams, G. L., Tallon, R. J., & Rimmel, P. (1980). A comparison of lecture versus role-playing in the training of the use of positive reinforcement. *Journal of Organizational Behavior Management*, 2, 205-212.
- *Arco, L. (1991). Effects of outcome performance feedback on maintenance of client and staff behavior in a residential setting. *Behavioral Residential Treatment*, 6, 231-247.
- *Azrin, N. H., & Pye, G. E. (1989). Staff management by behavioral contracting. *Behavioral Residential Treatment*, 4, 89-98.
- *Azrin, N. H., Jamner, J. P., & Besalel, V. A. (1989). Student learning as the basis for reinforcement to the instructor. *Behavioral Residential Treatment*, 4, 159-170.
- *Baker, D. J., Fox, J., & Albin, R. W. (1995). Staff support strategies to increase the involvement of individuals with disabilities in meal preparation. *Behavioral Interventions*, 10, 197-210.
- Boyce, T. E., & Geller, E. S. (2001). Applied behavior analysis and occupational safety: The challenge of response maintenance. *Journal of Organizational Behavior Management*, 21, 31-60.

Bucklin, B. R., Alvero, A. M., Dickinson, A. M., Austin, J., & Jackson, A. K.

(2000). Industrial-organizational psychology and organizational behavior management: An objective comparison. *Journal of Organizational Behavior Management*, 20, 27-75.

Christian, W. P. (1984). A case study in the programming and maintenance of institutional change. *Journal of Organizational Behavior Management*, 5, 99-153.

Cooper, J. O., Heron, T. E., & Heward, W. L. (2007). *Applied behavior analysis* (2nd ed.). Upper Saddle River, NJ: Pearson/Merrill-Prentice Hall.

*Durcharme, J. M., & Feldman, M. A. (1992). Comparison of staff training strategies to promote generalized teaching skills. *Journal of Applied Behavior Analysis*, 25, 165-179.

Fawcett, S. B. (1991). Social validity: A note on methodology. *Journal of Applied Behavior Analysis*, 24, 235-239.

Fisher, D. (1983). The going gets tough when we descend from the ivory tower. *Analysis and Intervention in Developmental Disabilities*, 3, 249-255.

Fleming, R. K., & Reile, P. A. (1993). A descriptive analysis of client outcomes associated with staff interventions in developmental disabilities. *Behavioral Residential Treatment*, 8, 29-43.

- *Fleming, R. K., & Sulzer-Azaroff, B. (1992). Reciprocal peer management: Improving staff instruction in a vocational training program. *Journal of Applied Behavior Analysis*, 25, 611-620.
- *Fleming, R. K., Oliver, J. R., & Bolton, D. M. (1996). Training supervisors to train staff: A case study in human service organization. *Journal of Organizational Behavior Management*, 16, 3-25.
- *Ford, J. E. (1984). A comparison of three feedback procedures for improving teaching skills. *Journal of Organizational Behavior Management*, 6, 65-77.
- *Green, C. W., Parsons, M. B., & Reid, D. H. (1993). Integrating instructional procedures into traditional congregate care situations for people with severe disabilities. *Behavioral Residential Treatment*, 8, 243-262.
- *Green, C. W., Reid, D. H., Perkins, L. I., & Gardner, S. M. (1991). Increasing habilitative services for persons with profound handicaps: An application of structural analysis to staff management. *Journal of Applied Behavior Analysis*, 24, 459-471.
- Greene, B. F., Willis, B. S., Levy, R., & Bailey, J. S. (1978). Measuring client gains from staff implemented programs. *Journal of Applied Behavior Analysis*, 11, 395-412.
- Harchik, A. E., & Campbell, A. R. (1998). Supporting people with developmental disabilities in their homes in the community. *Journal of Organizational Behavior Management*, 18, 83-101.

- *Harchik, A. E., Anderson, M., Thompson, R., Forde, K., Feinberg, L., Rivest, S., & Luiselli, J. K. (2001). Evaluation of a participatory, competency-based model of staff training in a community habilitative setting. *Behavioral Interventions, 16*, 1-13.
- Harchik, A. E., Sherman, J. A., Sheldon, J. B., & Strouse, M. C. (1992). Ongoing consultation as a method of improving performance of staff members in a group home. *Journal of Applied Behavior Analysis, 25*, 599-610.
- Hawkins, R. P. (1991). Is social validity what we are interested in? Argument for a functional approach. *Journal of Applied Behavior Analysis, 24*, 205-213.
- *Huberman, W. L., & O'Brien, R. M. (1999). Improving therapist and patient performance in chronic psychiatric group homes through goal-setting, feedback, and positive reinforcement. *Journal of Organizational Behavior Management, 19*, 13-36.
- Hyten, C. (2002). On the identity crisis in OBM. *The Behavior Analyst Today, 3*, 301-310.
- Kazdin, A. E. (1977). Assessing the clinical or applied importance of behavior change through social validation. *Behavior Modification, 1*, 427-452.
- Kennedy, C. H. (1992). Trends in the measurement of social validity. *The Behavior Analyst, 15*, 147-156.
- *Methot, L. L., Williams, W. L., Cummings, A., & Bradshaw, B. (1996). Measuring the effects of a manager-supervisor training program through the

generalized performance of managers, supervisors, front-line staff and clients in a human service setting. *Journal of Organizational Behavior Management*, 16, 3-34.

*Mozingo, D. B., Smith, T., Riordan, M. R., Reiss, M. L., & Bailey, J. S. (2006). Enhancing frequency recording by developmental disabilities treatment staff. *Journal of Applied Behavior Analysis*, 39, 253-256.

Parsons, M. B. (1998). A review of procedural acceptability in organizational behavior management. *Journal of Organizational Behavior Management*, 18, 173-190.

*Parsons, M. B., & Reid, D. H. (1993). Evaluating and improving residential treatment during group leisure situations: A program replication and refinement. *Research in Developmental Disabilities*, 14, 67-85.

*Parsons, M. B., & Reid, D. H. (1995). Training residential supervisors to provide feedback for maintaining staff teaching skills with people who have severe disabilities. *Journal of Applied Behavior Analysis*, 28, 317-322.

*Parsons, M. B., Cash, V. B., & Reid, D. H. (1989). Improving residential treatment services: Implementation and norm-referenced evaluation of a comprehensive management system. *Journal of Applied Behavior Analysis*, 22, 143-156.

*Parsons, M. B., Rollyson, J. H., & Reid, D. H. (2004). Improving day-treatment services for adults with severe disabilities: A norm-referenced application

of outcome management. *Journal of Applied Behavior Analysis*, 37, 365-377.

Parsons, M. B., Schepis, M. M., Reid, D. H., McCarn, J. E., & Green, C. W. (1987). Expanding the impact of behavioral staff management: A large-scale, long-term application in schools serving severely handicapped students. *Journal of Applied Behavior Analysis*, 20, 139-150.

*Realon, R. E., Lewallen, J. D., & Wheeler, A. J. (1983). Verbal feedback vs. verbal feedback plus praise: The effects on direct care staff's training behaviors. *Mental Retardation*, 21, 209-212.

Reid, D. H. (1998). Introduction. *Journal of Organizational Behavior Management*, 18, 1-5.

Reid, D. H., & Parsons, M. B. (1996). A comparison of staff acceptability of immediate versus delayed verbal feedback in staff training. *Journal of Organizational Behavior Management*, 16, 35-47.

Reid, D. H., & Whitman, T. L. (1983). Behavioral staff management in institutions: A critical review of effectiveness and acceptability. *Analysis and Intervention in Developmental Disabilities*, 3, 131-149.

*Reid, D. H., Green, C. W., & Parsons, M. B. (2003). An outcome management program for extending advances in choice research into choice opportunities for supported workers with severe multiple disabilities. *Journal of Applied Behavior Analysis*, 36, 575-578.

*Richman, G. S., Riordan, M. R., Reiss, M. L., Pyles, D. A., & Bailey, J. S.

(1988). The effects of self-monitoring and supervisor feedback on staff performance in a residential setting. *Journal of Applied Behavior Analysis*, 21, 401-409.

Riley, A. W., & Frederiksen, L. W. (1984). Organizational behavior management in human service settings: Problems and prospects. *Journal of Organizational Behavior Management*, 5, 3-16.

Schell, R. M. (1998). Organizational behavior management: Applications with professional staff. *Journal of Organizational Behavior Management*, 18, 157-171.

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

*Shore, B. A., Iwata, B. A., Vollmer, T. R., Lerman, D. C., & Zarcone, J. R.

(1995). Pyramidal staff training in the extension of treatment for severe behavior disorders. *Journal of Applied Behavior Analysis*, 28, 323-332.

*Sigafoos, J., Roberts, D., Couzens, D., & Caycho, L. (1992). Improving instruction for adults with developmental disabilities: Evaluation of a staff training package. *Journal of Applied Behavior Analysis*, 7, 283-297.

- Sigurdsson, S. O., & Austin, J. (2006). Institutionalization and response maintenance in organizational behavior management. *Journal of Organizational Behavior Management*, 26, 41-77.
- Sturmey, P. (1998). History and contribution of organizational behavior management to service for persons with developmental disabilities. *Journal of Organizational Behavior Management*, 18, 7-32.
- Sturmey, P. (2010, May 29). *Old friends: Organizational behavior management and developmental disabilities*. Speech presented at Thirty-Sixth Annual Convention of the Association of Behavior Analysis International in Henry B. Gonzalez Convention Center, San Antonio, TX.
- *Williams, W. L., & Lloyd, M. B. (1992). The necessity of managerial arrangements for the regular implementation of behavior analysis skills by supervisors and front line staff. *Developmental Disabilities Bulletin*, 20, 31-61.
- *Williams, W. L., Di Vittoria, T., & Hausherr, L. (2003). A description and extension of a human services management model. *Journal of Organizational Behavior Management*, 22, 47-71.
- Wolf, M. M. (1978). Social validity: The case for subjective measurement or how applied behavior analysis is finding its heart. *Journal of Applied Behavior Analysis*, 11, 203-214.

VITA

Graduate School
Southern Illinois University

Katherine M. Stegman

Date of Birth: September 2, 1983

105 South Forest Street Apt. # 2, Carbondale, Illinois 62901

305 Moody Street, Spearville, KS 67876

kmstegman@gmail.com

University of Kansas

Bachelor of General Studies, Applied Behavior Science, May 2006

Research Paper Title:

The use of organizational behavior management in training applied
behavior analysis methods in residential settings: A review

Major Professor: Dr. Nicole A. Heal