University of Nevada, Reno

Treatment Fidelity in the *Journal of Applied Behavior Analysis:* 1991-2010

A thesis submitted in partial fulfillment
of the requirements for the degree of

Bachelor of Science in Neuroscience and the Honors Program

by

Isha K. Patel

Dr. Patrick Ghezzi, Thesis Advisor

May, 2014

UNIVERSITY
OF NEVADA   **THE HONORS PROGRAM**
RENO

We recommend that the thesis
prepared under our supervision by

ISHA K. PATEL

entitled

Treatment Fidelity in the *Journal of Applied Behavior Analysis:* 1991-2010

be accepted in partial fulfillment of the
requirements for the degree of

BACHELOR OF SCIENCE, NEUROSCIENCE

_____
Patrick Ghezzi, Ph.D., Thesis Advisor

_____
Tamara Valentine, Ph. D., Director, **Honors Program**

May, 2014

Abstract

Treatment integrity is a critical component in ensuring that any protocol is implemented as it is planned. Unfortunately it is not a variable that has been recorded very often in the field of Behavior Analysis. Protocols that are not implemented with high integrity threaten the validity and outcome of the protocol. This review replicated the study of Gresham, Gansle, and Noell (1993) and coded all of the articles with children as subjects between the years of 1991 and 2010 and published in the *Journal of Applied Behavioral Analysis.* The protocol was modified slightly to allow for the incorporation of the variables: location, article type, presence of disabilities. The results showed that only 33.3% of all of the articles that fit the inclusion criteria for this study reported data on treatment integrity. The shocking lack of growth in awareness for the importance of treatment integrity in the years after Gresham's (1993) study has huge implications for the state of validity in Behavior Analysis. Treatment integrity is an important factor in Applied Behavior Analysis because protocols that are implemented with high integrity have a more effective outcome than studies which do not provide data on treatment integrity. This study may help further awareness for the importance of the collection of treatment integrity data in protocols implemented in Behavioral Analysis, but also across all other fields of treatment.

Acknowledgements

I would like to express my sincerest gratitude to Dr. Patrick Ghezzi, Ph.D. and to Emily

Skorzanka for guiding and supporting me through this project and helping me research a very

important, under acknowledged aspect of Applied Behavior Analysis. I would also like to

thank the University of Nevada-Reno and the Honors Program for providing me with the

opportunity to delve into such a project. Lastly, I would like to thank my friends and family

for their endless support and encouragement.

Table of Contents

List of Figures

List of Tables

Introduction

Treatment integrity, also called treatment fidelity, or procedural integrity is defined as the degree to which a treatment is implemented as planned (Fiske, 2008). It is a factor that is important across all fields, but especially so in medical and behavioral sciences. Without integrity, there would be no consistency within or between treatment protocols or among those individuals implementing treatments.

Applied Behavior Analysis (ABA) uses learning techniques in order to cause a change in behavior. The term treatment integrity, is a term specific to the field of ABA. As defined by Gresham, Gansle, and Noell , treatment integrity is, " The degree to which an independent variable is implemented as intended…" (1993). Treatment integrity is important for a variety of reasons that will be covered below.

McIntyre, Gresham, DiGennaro, and Reed (2007) explained the importance of treatment integrity and stated that incorrectly following protocols could result in poor outcomes. For example, if a researcher was asked to follow a protocol that has 4 parts and the researcher forgot to implement one of the parts or implemented them in the wrong order, the procedure would be thought to have poor integrity. Poor integrity would jeopardize the reliability of the procedure, not to mention, the outcome. Another example of procedural integrity is in medicine. Procedural fidelity is an especially important factor in medicine because it is imperative to ensure that all procedures are performed as they were intended to be performed and also executed consistently by all treatment providers. Without treatment integrity, a patient could lose the wrong limb, have surgery on the wrong organ, and could

come away from the hospital in worse shape than when the patient initially reached there. The examples show just how important it is to have integrity present in all fields.

Although applicable to everything from education to mechanics, treatment integrity is a variable that is especially important in the field of behavioral science where the protocols followed with high integrity produce the strongest outcomes (Fiske, 2008). However, treatment integrity is a factor that is often overlooked by behavioral researchers. Gresham, Gansle, and Noell (1993) found that only 16% of studies pertaining to Applied Behavior Analysis (ABA) and children done between 1980 and 1990 reported data on treatment integrity. Out of the 158 studies that were included in this study, this is infact a very small percentage of articles. Gresham, Gansle, and Noell (1993) analyzes articles pertaining to ABA with children (defined as individuals under 19 years of age) and published in the *Journal of Applied Behavioral Analysis (JABA)* between the years of 1980 and 1990. These articles were coded to check for the reporting of treatment integrity data in the studies that met their inclusion criteria.

The purpose of the current study was to replicate the protocol established by Gresham el at (1993) and examine the trends of reporting treatment integrity in research studies with children between 1991 and 2010 and published in the *JABA*. The study also analyzes the correlation between treatment integrity and the location of the studies to determine whether there was a tendancy to over or under report integrity data based on setting. After Gresham, Gansle and Noell's study (1993), the aim of this study is to analyze the trends in the collection of treatment integrity data in Applied Behavior Analysis, as well as to call attention to the importance of treatment integrity in ABA and to urge researchers to collect and report integrity data.

Literature Review

Treatment integrity is the extent to which a protocol has been implemented as it was intended to be implemented (Fiske, 2008). The lack of information about treatment integrity in a study leads to difficulties in making correct conclusions about the data (Sanetti, 2009). Treatment integrity is crucial to all fields of research because it ensures that a particular procedure is being executed by the means it was intended to be implemented. Without treatment integrity, not only are we unable to draw accurate conclusions, but we also lose a factor of validity to the research (Wheeler, Mayton, Carter, Chitiyo, Mendez, & Huang, 2009). Treatment integrity is the essential factor in ensuring that the best desired outcome is reached via the proper implementation of the protocol. The concept of treatment integrity applies across all fields and provides many implications as far as the safety and validity of human subject procedures and research.

Medicine is a key example of a field that revolves around treatment integrity. Without safety protocols in place in medicine, there would be no consistency between treatment providers, which is a startling realization. In 2008, the World Health Organization started an initiative called the "Safe Surgery Saves Lives" campaign, which developed a checklist for integrity during invasive procedures (Marjot, Maruthappu, & Shalhoub, 2013). This checklist was well received by the operating room staff and reduced morbidity and mortality, but unfortunately it did not generalize well outside of the operating room because the procedures were not followed by medical staff in other medical settings (Marjot et al., 2013). During invasive procedures conducted in settings outside of the operating room (such as labs and clinics), 10% of patients experienced adverse reactions, 40% of which were preventable. The conclusion is that outside of the operating room; a checklist is seldom used,

leading to preventable complications (Marjot et al., 2013). This is an example of treatment integrity generalized to fields outside of Behavior Analysis because it shows how a procedure with high integrity leads to a reduced risk of adverse reactions. In the medical example, when the procedure was not followed with high integrity, the result was adverse reactions to an improperly implemented procedure.

The medical example above shows how critical treatment integrity is to the correct implementation of a procedure. In the example, the presence of treatment integrity checks in the operating room and the lack of treatment integrity checks in other settings were the difference between the best positive outcome and complications due to incorrect implementation of the protocol. Fortunately, hospitals and clinics have many procedural integrity checks in place in order to ensure correct implementation of procedures.

Similar to medicine, treatment integrity plays a crucial role in the field of ABA. This is because treatment integrity strengthens the effects of intervention, which means that higher integrity will produce more positive results (Fiske, 2008). It is important to have high integrity when implementing behavioral protocols because having low integrity and a lack of consistency between treatments or between researchers will most likely lead to negative results. Weiss states that it is important to check integrity on a regular basis in order to ensure the highest quality of treatment, and therefore the most positive outcome possible (2005). Regular integrity checks are important to conduct because they ensure that the protocol is constantly being implemented with high integrity.

Treatment integrity reporting is often a neglected area of research even though researchers have brought attention to the importance of collecting treatment integrity data during ABA procedures. Fryling, Wallace, and Yassine stated that there is a correlation

between higher treatment integrity and improved treatment (2012). In 1982, Peterson et al found that a large number of cases between the years of 1968 and 1980 did not provide any measure of treatment integrity. Unfortunately this is not a fact that has changed much since that time.

In a study conducted by Wheeler et at (2006), 18% of the studies conducted with children with autism provided a measure for treatment integrity. This is a worrisome result as it shows that there was not much evidence for consistency in the implementation of protocols in the field. Also, it is surprising to find such a lack of data on treatment integrity when studies by Peterson et al (1982), Gresham, Gansle, and Noell (1993) and Wheeler et al (2006), that all report a measure of integrity under 20% of the total number of studies included in their reports. More surprising is that there are a number of methods for collecting treatment integrity data, all of which are fairly simple, and include methods such as self-reporting, feedback from observers, videotaping of the session, and direct observation (Lane, Bocian, MacMillan, & Gresham, 2004; Weiss, 2005). Still, there is such a lack of treatment integrity measures reported in ABA. Unfortunately, treatment integrity is not reported as commonly as it should be (McIntyre et al., 2007) and has especially been overlooked in the field of ABA (McIntyre et al., 2007; Peterson, Homer, & Wonderlich, 1982; Wheeler, Baggett, Fox, & Blevins, 2006). Wheeler et al. makes the claim that, "…the field needs to be more steadfast in promoting the inclusion of treatment integrity measures in behavioral treatment measures…" (2009). Weiss states that treatment integrity should be evaluated on an ongoing basis to ensure that, over time, trained treatment providers maintain a particular level of integrity (2005).

Method

All of the articles being reviewed were published in the *Journal of Applied Behavior Analysis (JABA)* between the years of 1991-2010. This study is a replication and continuation of the study done by Gresham, Gansle, and Noell published in 1993. The methodology for this study was taken from the protocol provided in the Gresham study and modified in order to analyze different variables that could be of significance.

Criteria for Review

Studies that were included in this review were based on the following criteria: Both research and brief reports were included in the review. Brief reports were defined as three or fewer pages, and a research report was defined as anything greater than three pages in length. The population studied was children, defined as those individuals 19 years or younger; and the date of publication in JABA was 1991-2010.

Coding Procedure

The studies that fit the criteria were coded on the basis of these 5 variables: (1) age; (2) location of the study; (3) if the participants had disabilities (which included all mental and intellectual diagnoses); (4) type of report (brief or research); (5) and assessment of treatment integrity. Age was recorded as it was reported in the studies. Location data was recorded based on four categories: Home, school, clinical (which included all medical and non medical based treatment facilities), and other (which included community sessions, sports practices off of the school premises and when a study was conducted in more than one location). Disabilities were coded as 'yes' or 'no' and included all physical and intellectual diagnoses. The assessment of treatment integrity was coded in three categories, (1) Yes; (2)

no; and (3) monitored; following the coding strategy used by Gresham, Gansle, and Noell (1993). A checklist of the coding procedure can be found in Appendix 1.

Treatment integrity

If a study provided information on integrity and reported numbers or percentages of integrity, it was coded as yes. Studies that provided protocols and information on how and to what extent the integrity protocols were implemented were also coded as yes. If a study provided no mention of integrity or provided no results on integrity, it was coded as no. If a study mentioned integrity but did not report the results of integrity checks, or said that the integrity data are provided elsewhere, it was coded as monitored.

Results

From 1991-2010, 652 articles met the criteria for inclusion for this study. On average, there were 32 articles children as the subjects published each year between 1991 and 2010, with the least being 17 studies in 1991 and the most being 45 studies in 2010. Of the studies that were published during this time, 33.2% (N=213) reported data on treatment integrity. During this time 38.3% (N=246) did not provide measure of treatment integrity. The remaining 30.1% (N=193) of the studies fell into the monitored category, which meant that treatment integrity was mentioned in the study but no data were provided.

The data are included in Tables 1 and 2. Table 2 shows the data reported by Gresham, Gansle, and Noell (1993) and Table 1 shows data from the current study. The trends in integrity data from 1980-2010 can be seen in Figures 1 and 2 which graphically represent the data of Gresham, Gansle, and Noell continuing into the data from this study. Figures 3 and 4 show data on treatment integrity in relation to the location of the studies. Figure 5 shows the cumulative frequency of the studies with and without treatment integrity

reported. Figures 6 and 7 show the data collected by Gresham and his colleagues (Table 2)

in a graphical manner.

Table 1. Table presented in Gresham , Gansle, and Noell (1993). All of the data is presented as a percentage of the number of studies for each year. (The raw numbers were not provided in the Gresham, Gansle, Noell study (1993).

Year	yes	No	Modified	Combined
1980	18.20%	72.70%	9.10%	81.80%
1981	17.60%	76.60%	5.90%	82.50%
1982	11.10%	83.30%	5.60%	88.90%
1983	37.50%	50.00%	12.50%	62.50%
1984	13.30%	73.30%	13.30%	86.60%
1985	0.00%	100.00%	0.00%	100.00%
1986	15.00%	75.00%	10.00%	85.00%
1987	7.70%	61.50%	30.80%	92.30%
1988	25.00%	75.00%	0.00%	75.00%
1989	10.00%	80.00%	20.00%	100.00%
1990	13.30%	86.70%	0.00%	86.70%

Note: The combined no/monitored category is also present in this table in the starred no

category at the end. This table is the same table that was provided in the Gresham, Gansle,

Noell study (1993) with the addition of the combined no/monitored category at the end. It

shows the percentages of treatment integrity data that was reported in the articles included in

their study.

Table 2. Data in percentages and numbers for the current study: 1991-2010 with "yes", "no", "monitored", and combined no/monitored categories. This table shows the number of studies in each category: Yes, No, and Monitored, as well as the percentages of the total number of studies for each year.

Year	Yes	% Yes	No	% No	Monitored	% Monitored	Combined No*	% Combined No
1991	13	76.5%	1	5.9%	3	17.6%	4	23.5%
1992	19	61.3%	3	9.7%	9	29.0%	12	38.7%
1993	12	50.0%	7	29.2%	5	20.8%	12	50.0%
1994	17	51.5%	4	12.1%	12	36.4%	16	48.5%
1995	9	36.0%	6	24.0%	10	40.0%	16	64.0%
1996	9	26.5%	13	38.2%	12	35.3%	25	73.5%
1997	16	43.3%	12	32.4%	9	24.3%	21	56.7%
1998	8	22.8%	15	42.8%	12	34.4%	27	77.2%

1999	8	38.1%	4	19.1%	9	42.8%	13	61.9%
2000	9	26.5%	12	35.3%	13	38.2%	25	73.5%
2001	6	18.8%	13	40.6%	13	40.6%	26	81.2%
2002	7	25.9%	7	25.9%	13	48.2%	20	74.1%
2003	8	29.6%	9	33.3%	10	37.1%	19	70.4%
2004	5	15.2%	13	39.4%	15	45.4%	28	84.8%
2005	6	17.6%	22	64.7%	6	17.7%	28	82.4%
2006	7	21.9%	21	65.6%	4	12.5%	25	78.1%
2007	12	27.3%	24	54.5%	8	18.2%	32	72.7%
2008	11	30.4%	16	47.6%	8	22.0%	24	69.6%
2009	16	30.7%	24	46.2%	12	23.1%	36	69.3%
2010	15	33.4%	20	44.4%	10	22.2%	30	66.6%

Note: This table reports the data from the replication. It reports the number of studies per year that were recorded in each category: Yes, no, and monitored. It also presents the data as percentages of the number of studies per year. The last two columns present the number of the studies and the percentage of total studies per year, respectively, with a combined no/monitored category.
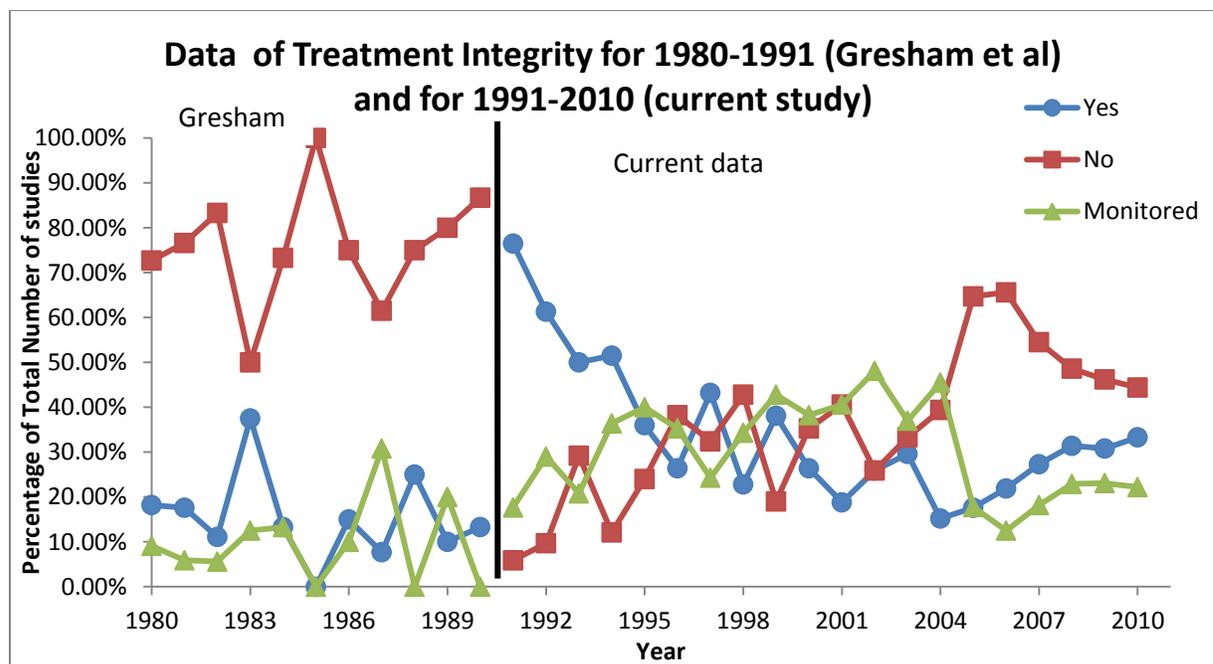

Figure 1. Continuation graph presented with a phase line between the Gresham data and the new data. The portion of this figure before the phase change line shows a graphical representation of the percentage data reported as a table in the original Gresham , Gansle, and

Noell study (1993). The "yes" line represents the percentage of the total number of studies that reported integrity data. The "no" line represents the percentage of the total number of studies that did not provide any information on treatment integrity in their study. The "monitored" line represents the percentage of the total number of studies that were considered to be in the monitored category. Studies in this category provided little information on integrity or had integrity data/protocols that were accessible from sources outside of the study (e.g. from the first author). The portion of the graph following the phase change line represents the data collected in this study, according to the same classifications.

When the "no" and "monitored" categories were combined into a broader "no"

category, 33.2% of the studies reported integrity data while 68.4% of the studies did not.

Figure 2 shows this new categorization of the Gresham data followed by the data from this
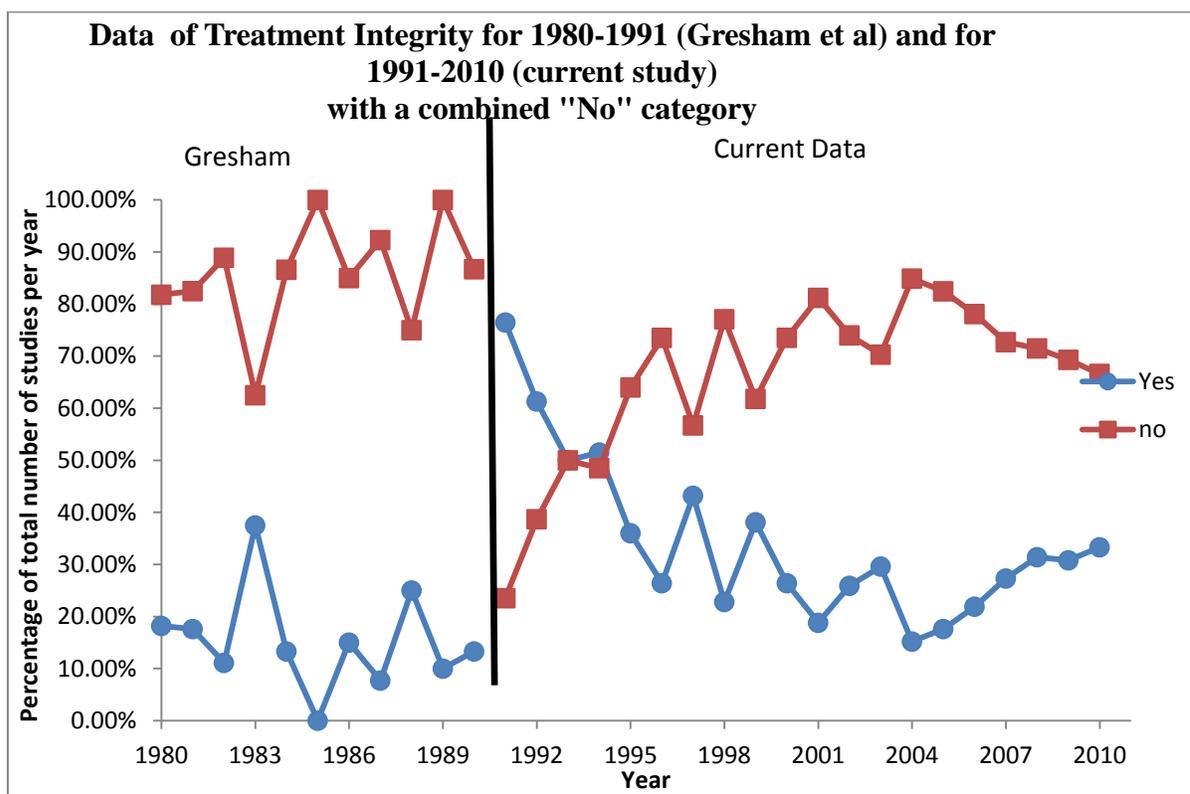
study.

Figure 2. Continuation graph presented with a phase line between the Gresham data and the new data where the "monitored" and "no" categories are both represented by the "no" line. The portion of data before the phase change line presents the data reported by Gresham as a percentage of the total number of studies, but shows the data for the "no" and "monitored" combined into one "no" category. The portion after the phase change line presents the data

from this study in percentages of the total number of studies. The "monitored" and "no" categories were combined into one "no" category for this as well.

Both Figure 1 and Figure 2 show a spike in the reporting of integrity data following Gresham's study, with an immediate decline in the years following with a slight upward trend in the most recent years. One potential reason for this spike in the reporting of treatment integrity between the years of 1991-1993 could be that treatment integrity could have been a topic at an ABA conference during this time.

Location data showed that studies conducted in a school or classroom setting most often collected and reported integrity data. Of the studies conducted in a school setting 41.4% reported integrity data. There were integrity data reported for 24.6% of studies conducted in a clinical setting, 32.8% of studies conducted in a home setting, and 28% of studies conducted in other or a combination of settings. For 43.8% of clinical studies, 43.1% of home studies, 30.8% of school studies, and 40.2% of other settings, there was no integrity data or protocol reported. The remaining 31.7% of studies in the clinical setting, 24.1% of studies in the home setting, 27.8% of studies in the school setting, and 31.7% of studies in other settings or a combination of settings fell into the monitored category where integrity was mentioned but there was no data reported on it (Figure 3). When the "no" and the "monitored" categories were combined into a broader "no" category, 75.4% of clinical studies, 67.2% of home studies, 58.6% of school studies, and 72% of studies in other or a combination of locations were considered to not provide integrity data (Figure 4). Location data was not provided by Gresham , Gansle, and Noell (1993) so the figures only include data from the current study.
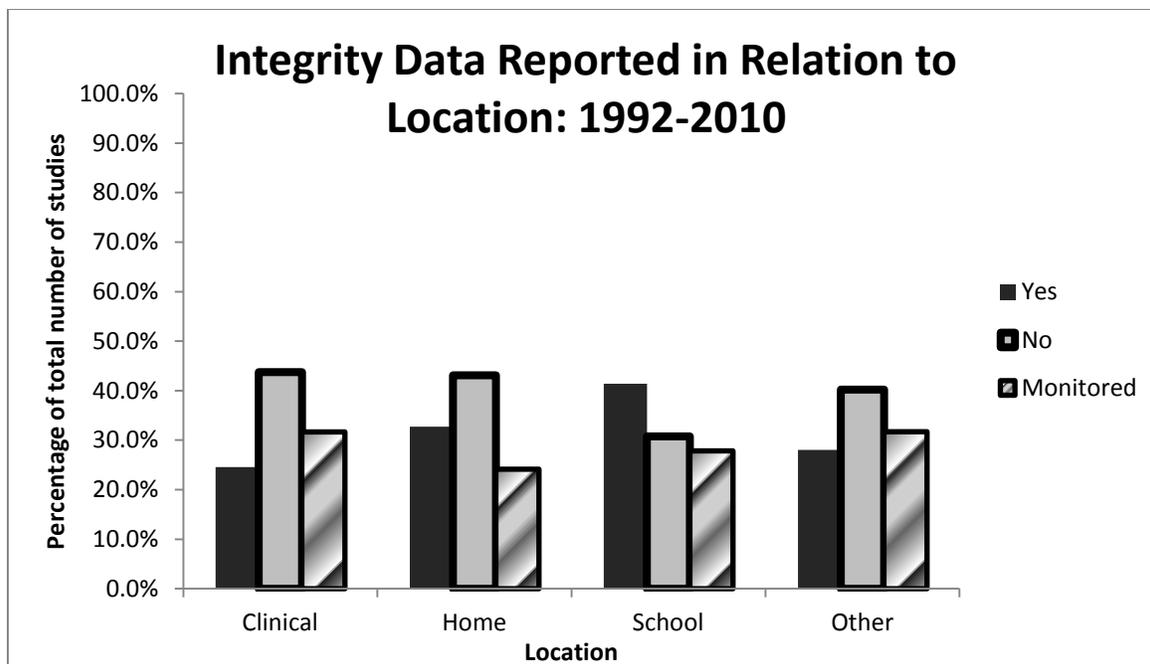
Figure 3. Graph of integrity data and the locations where the studies were conducted. The graph shows the percentage of the total number of studies presented in terms of location. The location data was not collected in the initial Gresham study but was collected for the purposes of this study. This graph shows how the three categories of integrity were represented across various locations.
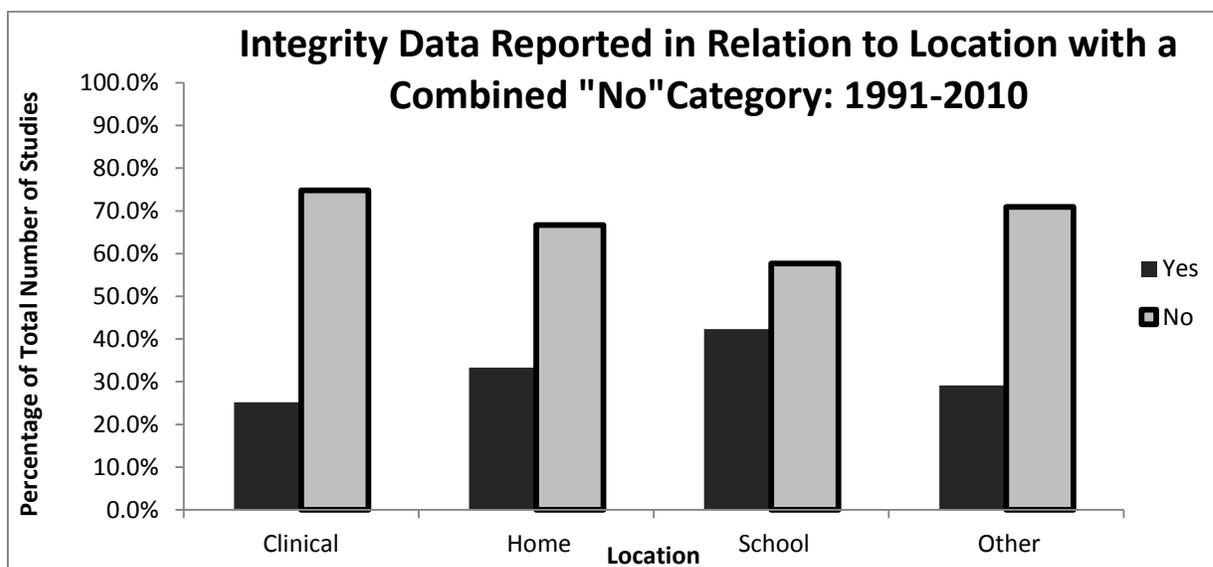

Figure 4. Graph of integrity data and the locations where the studies were conducted with a combined no category. This graph represents the percentage of the total number of studies in terms of the location. In this graph, the "no" and "monitored" categories are combined.

The data from the current study are also shown as a cumulative frequency in Figure 5. Cumulative frequency graphs show the rate at which there is change in a category. The results for this study showed that there was a much more rapid change in the "no" category than there was in the "yes" category. This graph showed a much steeper rise in the "no" category, with a slope of m=22.5 while the "yes" category had a much shallower rise with a slope of m=9.21. The data for this graph was calculated by adding each successive year worth of data together in order to achieve a graph that depicts rate. The slope was calculated by using a line of best fit though Excel.
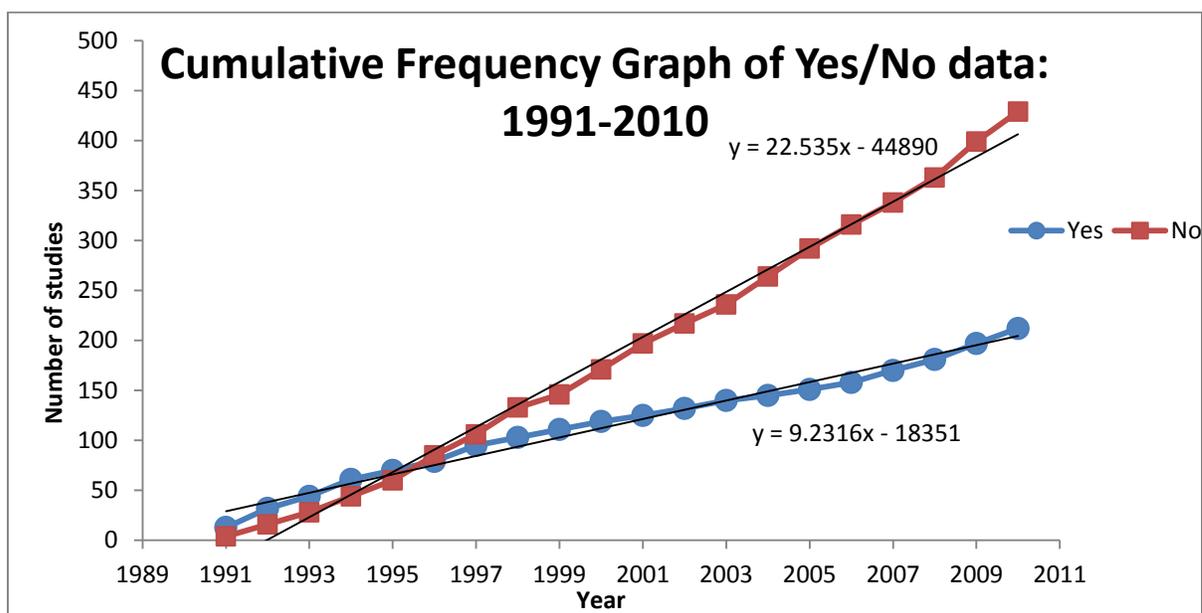
Figure 5. Cumulative Frequency graph representing the data in two categories: Yes and no. This graph presents the data from the current study in the form of cumulative frequency graph which allows the data to be seen as change over time. The data is presented as "yes" and "no", where "yes" is made up of studies that reported treatment integrity and "no" consists of all of the studies in the original "no" and "monitored" categories. The graph shows the slopes of these data and represents the rate of the occurrence of these conditions.

Discussion

The study by Gresham , Gansle, and Noell (1993) brought to light the importance of treatment integrity in procedures done with children and showed how an astonishing number

of researchers have not been conducting integrity checks to ensure that their protocols were being implemented as intended. The lack of procedural integrity checks could mean that protocols could have been implemented incorrectly, leading potentially, to a negative outcome. The results of the current study show that despite Gresham's efforts, treatment integrity data are still not consistently recorded in studies conducted with children, which implies that the importance of treatment integrity may still be a factor that many researchers do not fully understand the importance of or see as an optional variable when implementing a protocol.

Following the Gresham, Gansle, and Noell (1993) study which could have been due similar publication on the importance of treatment integrity to have been published around this time. Immediately after the publication of the Gresham, Gansle, and Noell (1993) study, we see a large decline in the reporting of integrity and large upward trend in the "no" category. This is a surprising trend, especially when considering that the desired and anticipated aftereffects of Gresham's study would have been to see a steady upward trend in the reporting of integrity data in studies conducted with children in *JABA*. This shows that the Gresham study did not have the lasting results they had hoped for but instead may have had a temporary effect in the years surrounding the conduction of the study. The Gresham study was first received by *JABA* in 1992, and although published in late 1993, it is possible that the results of the study could have been released into the scientific community before this time. It is also possible that treatment integrity could have been a topic at ABA conferences during this time frame which could explain the short increase in the reporting of treatment integrity data.

        The results also showed that despite Gresham's study (1993), there has not been a

change in the reporting of integrity data in *JABA* articles since 1990. The immediate increase

in reporting of integrity data was followed by a decrease that eventually leveled off to a level

close to the levels reported in Gresham's study , although the percentage of "no" data

remained lower in this study than in the original study. This could largely have to do with

the number of journal articles included in each study. The Gresham study included 158

articles that fit the inclusion criteria while this study included 653 articles. Just the vast

difference in the numbers of studies could have made a difference on the overall data. Also,

the difference in number of articles included from each year could have affected the data.

The percentages for each year were reported out of the total number of articles that met the

inclusion criteria for that year. There were fewer articles that met the inclusion criteria in the

years where the "yes" data spiked than there were in the years towards the end of the range.

Similarly, there were fewer articles per year published in the Gresham, Gansle, and Noell

(1993) study than in this study, with an average of 14 articles per year in Gresham and 33

articles per year in this study.

        The data reported by Gresham, Gansle, and Noell (1993) was graphed side by side

with the data from this study in a category combining "no" and "monitored" into one broad

"no" category. The effect of the combined category can be seen in the figures presented in

the results (Figure 1; Figure 2). This was not a factor that was analyzed in the study done by

Gresham, Gansle, and Noell (1993) but was added in this study because the importance of

the "monitored" category was questionable. The monitored category was defined as those

articles that mentioned integrity but did not include a detailed protocol or numerical values to

support their mention of integrity. The "monitored" category itself provided an issue to

treatment integrity because the studies coded as "monitored" touched on the concept of

integrity but with stated that the information could be found in a different location or claimed

they had high integrity but provided no proof. An argument can be made for the fact that

making an empty claim without solid evidence is a breach of treatment integrity because it is

impossible to know if the researchers implemented all, some, or none of the aspects of their

protocol. We analyzed the monitored category as a branch of the no category because a lack

of quantitative evidence to prove that a particular protocol has been implemented as intended

is comparable to a lack of integrity collection entirely. Gresham, Gansle, and Noell (1993)

used the monitored category as a way to separate out studies which claimed to include

procedural integrity checks but had not included the data recorded on this variable. Figure 5

shows a graphical representation of the original Gresham data that was reported in the study.

This figure shows the "yes" and the "no" as two distinct figures in the data, with no overlap.

Figure 6 shows how the "yes" and the "no" categories remained two distinct categories with

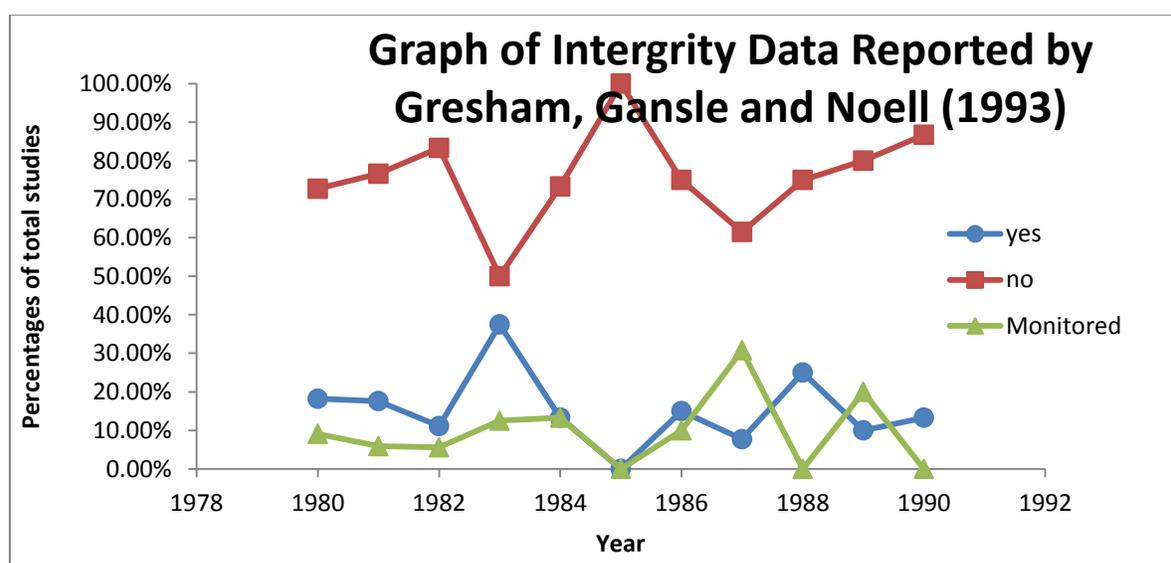"yes" still close to baseline and "no" closer to 100%.

Figure 6. Graph of the data reported by Gresham , Gansle, and Noell (1993). This figure shows a graphical representation of the percentage data reported as a table (Table 2) in the original Gresham study. The "yes" is the category including all articles within the inclusion criteria which provided treatment integrity data. The studies in the "no" category made no mention of integrity data and the "monitored" category included all articles within the inclusion criteria that mentioned integrity but provided mo empirical evidence of it.
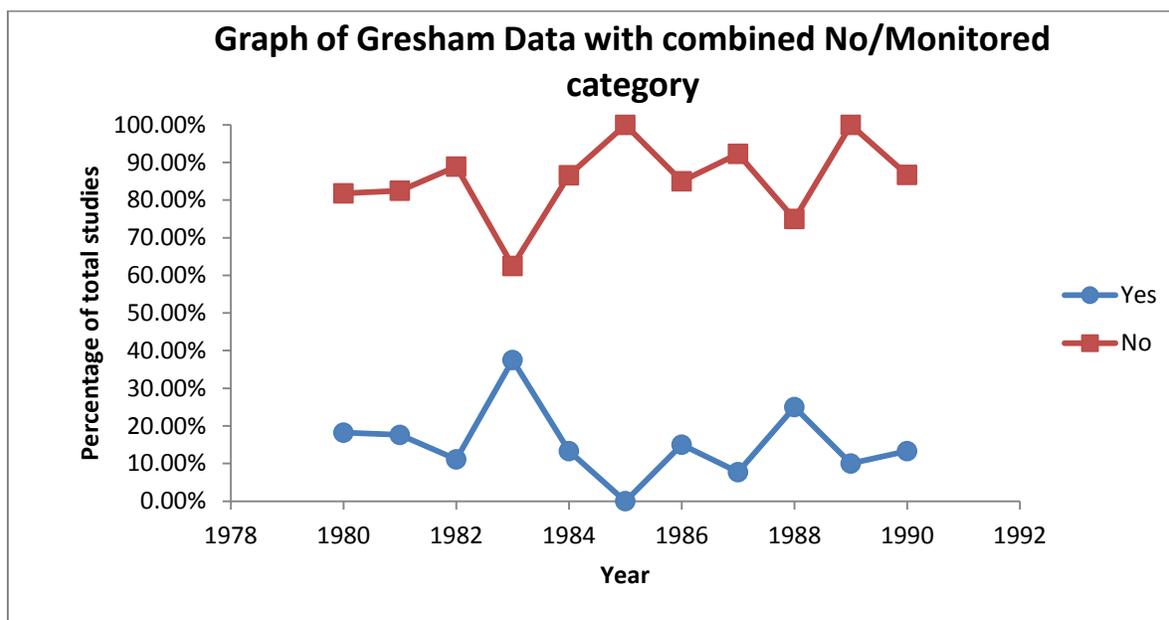

Figure 7. Graph of the data reported by Gresham , Gansle, and Noell (1993) with the no/monitored combined into the "no" category. This figure shows a graphical representation of the percentage data reported as a table in the original Gresham study. The blue line represents the percentage of the total number of studies that reported integrity data. The red line represents the percentage of the total number of studies that did not provide any information on treatment integrity in their study or that mentioned integrity data but did not provide any

There could be many reasons for this including changes in editorial policies or an increase in procedural safeguards such as interobserver agreement. A change in budget or editorial policies could be one reason for the spike in integrity data followed by the increase

in no integrity data. If *JABA* changed their publishing policies or lowered their budget, it could have potentially caused authors to shorten the length of their journal articles. An increase in procedural safeguards could also have led to an increase in the lack of integrity data provided in studies because researchers could have increased the number of procedural safeguards in order to imply integrity but not actually measure integrity itself. Almost all of the studies published in *JABA* have a section dedicated to inter observer reliability which implies a level integrity assessment in these studies despite the fact this it is not an overall integrity check for the protocol.

The Gresham, Gansle, and Noell study (1993) seemed to have some effect in the years surrounding its publication. Unfortunately, this trend has not spread across the years. Interestingly enough, there has been a gradual increase in the "monitored" category which could have implications about the long term effects of Gresham's study. This gradual increase could mean that researchers are becoming more aware of the importance of including procedural integrity checks within a study. Although these studies did not provide empirical evidence for the integrity measured from their study, the rise here shows that researchers have acknowledged the importance of at least noting integrity. The "yes" category is also gradually rising which shows that many years after the publication of Gresham's study, the importance of collecting and reporting integrity data in behavioral studies conducted with children has been noted. This rise in the "yes" category coincides with a gradual decrease in the "no" category. This combination holds potential for the future and suggests that more and more, the importance of reporting integrity data is being recognized.

Location data also provided interesting information. The data showed that in all home, clinical, and other locations, integrity data was only reported for an average of 28.5% of studies while in the school category 41.4% of studies reported integrity data. This could be because there are more safeguards in place in a school setting because there are higher levels of distracters when adhering to a protocol in a classroom. It could also be possible that that there is more focus given to studying treatment integrity as it pertains to school based treatment protocols, which could be the reason for the school setting to be the only category to have a higher level of treatment integrity data present. For example, the study by Fiske (2008) showed the importance of treatment integrity in behavioral protocols implemented by teachers and showed a correlation between high integrity and positive outcome. Also, Coding, Livanis, Pace, and Vaca (2008) which studied the importance of treatment integrity in providing feedback in the school setting and showed that performance feedback to teachers led to high integrity in the academic setting. It is possible that there was more data collection on to integrity in the school setting because researchers need to observe the teachers who conducted the sessions to ensure they were correctly following the protocol whereas in most other settings, the experimenters themselves are the ones conducting sessions.

With the combined no/monitored category, the results showed that there were still more studies that did not report integrity data than there were that did. This shows that location does not have a very strong correlation with the collection of integrity data and shows that this is a factor in experimenting which needs to have more attention brought to it.

Although the presence of disabilities was a variable that as recorded in this study, it was not a variable that showed much significance or correlation to the presence of treatment

integrity data presented in a study. One thing to be noted though, was the significant

overrepresentation of individuals with disabilities in the data set. Out of the 652 studies that

were represented in this study, 491 of the studies included participants with either physical or

intellectual disabilities. The overrepresentation of participants with disabilities is a factor

that should be further analyzed so that a reason for the over representation can be determined.

One possible explanation could be that ABA treatments often focus on individuals with

disabilities, so there is more representation of individuals with disabilities across publications

in *JABA*.

Conclusion

Limitations

A potential limitation in this study could be the number of articles included in this

review. The criteria for this study was based on the protocol developed by Gresham, Gansle,

and Noell (1993) and focused solely on studies based on children that were published in

JABA between the years of 1991 and 2010. This study provides a small look at integrity data

but perhaps by expanding the criteria to cover all demographics and encompass other

journals, more thorough data on the collection of treatment integrity in ABA could be

reported. Also, by examining treatment integrity outside of the scope of *JABA*, more data on

treatment integrity could be recorded and more information on treatment integrity would be

known.

Furthermore, a major limitation in this study was the lack of integrity data collection,

both in this research, as well as in the Gresham, Gansle, and Noelle study which this research

was mirrored after. It is interesting that the Gresham, Gansle, and Noelle study focuses on

the collection of treatment integrity data, yet it does not provide detail on how integrity was maintained in their study.

Implications

The findings of this study may contribute to the realization that there is not nearly a strong enough focus on the collection of treatment integrity. The lack of treatment integrity data in the majority of studies done with children implies that there may not be a control measure within the protocols of these studies. This means that some of these studies may not be being implemented as planned by the researchers.

Treatment integrity is a check of the independent variable. When confidence is low on how the independent variable was presented, then confidence in how the independent variable affects the dependent variable is also low. This means that when the integrity data is not consistently reported, there is no strong evidence for determining the relationship between the independent and dependent variables. Treatment integrity also plays a critical role in the replication of a study. Researchers who pay attention to the treatment integrity must pay attention to how their independent variables are presented, which in turn forces them to specify exactly how their procedures should be presented. This is very helpful for other researchers who wish to replicate that study.

The results of this study are important to Behavioral Analysis specifically, but also can be generalized to any other field of work. Having high integrity is critical for achieving the desired outcome, which is a concept that can apply to any protocol in any area of study. All researchers and treatment providers should provide the best treatments they can, and in order to do so, it is important that treatment integrity data be collected in order to ensure that the treatment is delivered as designed. Another point that can be made is that the recipients of

ABA treatment are consumers and consumers have a right to effective treatment. The treatment that is provided to a consumer is only as good as how effective the treatment is. This poses another limitation with the lack of treatment integrity data because consumers have a right of high integrity, highly effective treatments.

This study shows how severe the lack of focus on treatment integrity in Behavioral Analysis is, even after the publication of the study by Gresham, Gansle, and Noell (1993) and calls to attention the importance for treatment professionals and researchers alike, to support high levels of desired outcomes with high treatment integrity and the collection of this data during all studies.

The results of this study are a call to action for researchers everywhere. The lack of treatment integrity data being reported is alarming because the implication here is that researchers are following a procedure as published but without the proof that treatment integrity data collection provides, there is no way to be 100% certain. Furthermore, the lack of integrity data makes it difficult to replicate a study because it becomes impossible to ascertain if the protocol is being implemented correctly when the original study lack data on integrity.

Future Direction

The results from this study may lend to the development of future studies. A major future direction could be in the reporting of treatment integrity data in all studies conducted with children. This study focused on articles about children specifically published in *JABA,* but there could be important implications from expanding this study to other journals applicable to ABA and also from analyzing articles pertaining to ABA with adults.

Second, by addressing the correlation between the location of a study and the presence of treatment integrity reporting, a further analysis could be done in order to determine the relationship between setting and treatment integrity. The present study found that the only location to be associated with a larger number of studies reporting treatment integrity data was the school setting. One reason for this could be the fact that there is researcher supervision in the implementation of protocols by teachers. In this case, because the teachers are also participants in the experimental study, researchers record integrity data on the teachers. In the future, further analysis could examine the relationship between treatment integrity and location.

Finally, in regards to how treatment integrity is reported outside of the field of ABA, future studies could demonstrate the extent to which protocols in other fields are being conducted as planned. Treatment integrity is such a critical component for the proper implementation of protocols across all fields and it is crucial that the importance of treatment integrity becomes generalized across all fields of work. By assessing studies and protocols published in other areas for treatment integrity, better techniques for the evaluation of integrity can be developed. Future studies may develop a successful method for the collection and assessment of treatment integrity, which could lead to larger numbers of studies across all journals and fields of work providing data on integrity. In the future, it would be helpful to create more awareness for the importance of treatment integrity in research and establish prerequisites for publication in journals in which integrity data is a requirement for publication.

Another factor to research in the future could be the publication requirements and change in editors in *JABA*. There may be the possibility that publication guidelines on length

or on material could have prevented researchers from publishing specific data on treatment integrity. This research could help determine if there is a lack of focus on the importance of treatment integrity or if stringent publication guidelines cause researchers to remove treatment integrity data from the final published piece.

Appendix

Checklist for data collection:

_____ Make sure article is from *Journal of Applied Behavior Analysis* between the years of 1991-2010

_____ Open articles and first look at the age of participants in the study. If the age falls between the ages of 0-19 years, the study is to be included in the research collection

_____ Record the age of the participant in the data collection spreadsheet

_____Next, record if the participants had disabilities (Physical or mental)

____Then look at the setting that the study was conducted in (Either school, home, clinical, or other if it was conducted elsewhere or in a combination of these settings)

____ Search the article for procedural/treatment integrity/fidelity. If none of these words come up, read the article for mentions of integrity worded other ways, such as "Procedure was implemented with a high accuracy"

_____If the article provided quantitative evidence for treatment integrity such as the percentage with which the integrity was implemented---code the article as "Yes"

_____If the article said that they collected treatment integrity or reported a non quantifiable measure of integrity (ie the protocol was implemented with high integrity) but did not provide any quantitative support for it---code the article as "Monitored"

_____If the article did not mention treatment integrity or stated that integrity was not collected----code the article as "no"

____ Lastly, record the type of article. If it was 3 or less pages code it as brief and if it was >3 pages code it as a full research report

References

Fiske, K. E. (2008). Treatment integrity of school-based behavior analytic interventions: A review of the research. *Behavior Analysis on Practice*, 1, 19-25.

Fryling, M. J., Wallace, M. D., Yassine, J. N., & Lerman, D. (2012). Impact of treatment integrity on intervention effectiveness. *Journal of Applied Behavior Analysis*, *45*(2), 449-453.

Gresham, F.M., Gansle, K.A., & Noell, G.H. (1993). Treatment integrity in applied behavior analysis with children. *Journal of Applied Behavior Analysis*, 26, 257-263.

Lane, K. L., Bocian, K. M., MacMillan, D. L., & Gresham, F. M. (2004). Treatment integrity: An essential—but often forgotten—component of school-based interventions. *Preventing School Failure: Alternative education for children and youth*, *48*, 36-43. Retrieved from: http://www.wce.wwu.edu/Depts/SPED/Forms/Kens%20Readings/Instruction/Instruct%20Treatment%20Integrety%20Lane%202004.pdf

Marjot, T., Maruthappu, M., & Shalhoub, J. (2013). Checklists for invasive procedures. *New England Journal of Medicine*, *368*(3), 293-294. Retrieved from: http://www.nejm.org/doi/full/10.1056/NEJMc1213329

McIntyre, L. L., Gresham, F. M., DiGenarro, F. D., & Reed, D. D. (2007). Treatment integrity of school-based interventions with children in the Journal of Applied Behavior Analysis 1991-2005. *Journal of Applied behavior Analysis*, *40*(4), 659-672.

Peterson, L., Homer, A. L., & Wonderlich, S. A. (1982). The integrity of independent variables in behavior analysis. *Journal of Applied behavior Analysis*, *15*(4), 477-492.

Sanetti, L. R. (2009). Toward developing a science of treatment integrity: Introduction to the

special series. *School of Psychology Review, 38*(4), 445. Retrieved from:

http://web.a.ebscohost.com/ehost/pdfviewer/pdfviewer?sid=a00e2dd3-0fd6-4f02-bc62-04e3276cd461%40sessionmgr4004&vid=2&hid=4106

Weiss, M. J. (2005). Comprehensive ABA programs: integrating and evaluating the implementation of varied instructional approaches. *The Behavior Analyst Today*. Retrieved from:

http://www.thefreelibrary.com/Comprehensive+ABA+programs%3A+integrating+and+evaluating+the...-a0170115016

Wheeler, J. J., Baggett, B. A., Fox, J., & Blevins, L. (2006). Treatment Integrity: A review of intervention studies conducted with children with Autism. *Focus on Autism and Other Developmental Disabilities*, *21*(1), 45-54.

Wheeler, J. J., Mayton, M. R., Carter, S. L., Chitiyo, M., Mendez, A. L., & Huang, A. (2009). An assessment of treatment integrity in behavioral intervention studies conducted with persons with mental retardation. *Education and Training in Developmental Disabilities*, *44*(2), 187-195. Retrieved from:

http://daddcec.org/Portals/0/CEC/Autism_Disabilities/Research/Publications/Education_Training_Development_Disabilities/2009v44_Journals/ETDD_200906v44n2p187-195_An_Assessment_Treatment_Integrity_Behavioral_Intervention.pdf.