Summer 2016

Evaluating Current Practices in Brief Experimental Analysis

Alex P. Isbill
Western Kentucky University, alex.isbill672@topper.wku.edu

EVALUATING CURRENT PRACTICES IN BRIEF EXPERIMENTAL ANALYSIS

A Specialist Project
Presented to
The Faculty of the Department of Psychology
Western Kentucky University
Bowling Green, Kentucky

In Partial Fulfillment
Of the Requirements for the Degree
Specialist in Education

By
Alex P. Isbill

August 2016

EVALUATING CURRENT PRACTICES IN BRIEF EXPERIMENTAL ANALYSIS

Date Recommended 7-7-2016

Elizabeth L. Jones

Elizabeth Jones, Director of Specialist Project

Carl Myers

Carl Myers

Samuel Kim

Samuel Kim

Dean, The Graduate School

7/28/16

Date

ACKNOWLEDGEMENT

I would like to thank Dr. Myers and Dr. Kim for being members of my specialist project committee and supporting me during my time at WKU. And I would like to thank Dr. Jones for everything you have done for me, you pushed me to my limits and offered guidance when I needed it.

I would like to thank my cohort for their unwavering support throughout our time together. You guys rock, and I’m glad I got to share my graduate experience with you all.

Additionally, I would like to thank my family for their unconditional love and support. My parents, Pat and Jane Isbill, have provided constant support and encouragement over the course of my entire academic career, and I can never thank you enough. To my brother Kaleb and sister Abby, thanks for existing. And being cool, I guess. And finally, to my partner in crime, a big thank you to Sarah for being there when I needed you, and being patient with my busy schedule.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

EVALUATING CURRENT PRACTICES IN BRIEF EXPERIMENTAL ANALYSIS

Alex Isbill August 2016 73 pages

Directed by: Elizabeth Jones, Carl Myers, Samuel Kim

Department of Psychology Western Kentucky University

Brief experimental analysis (BEA) has shown to be an effective method of rapidly testing the relative effects of two or more interventions in order to determine interventions that best supports a student's learning. Little research has been found in regards to the consistency of methods across studies. A meta-analysis in 2008 by Burns and Wagner looked at BEAs that assessed oral reading fluency and provided recommendations for future practice. This study investigates the methods, procedures, and outcomes in BEA studies from 1994 to 2016. The findings of this study are compared to Burns and Wagner's (2008) recommendations from their meta-analysis, as well as used to discuss the shifts and consistencies found in BEA methodology over the past 23 years. There is not sufficient evidence that Burns and Wagner's (2008) recommendations have greatly impacted the process of BEA, but there have been changes in predominant methodological components of BEA such as the explicit use of conceptual models, methods of assessing interventions, and the emergence of a problem solving model to inform intervention selection. A general increase in the publication rate and a shift to publication in school psychology journals over behavioral journals was also noted. BEA outcomes continue to support its utility for informing instruction.

Introduction

Functional analysis involves the application of behavior analysis procedures "to identify important environmental variables that develop and maintain behavior. The clinical goal of functional analysis is to effectively identify targets of intervention that are alterable, in order that appropriate treatments may be rapidly implemented and evaluated" (Loverich, O'Donohue, & O'Donohue, 2010, p. 687). In the field of education, functional analysis has become a tool for addressing problematic behaviors exhibited by students. The functional analysis approach has been formalized procedurally and called a functional behavior analysis (FBA). FBA allows one to identify the antecedents and consequences to the behaviors of concern, which allows for identification of a behavior's function for the individual. Once the function of a behavior is identified, interventions can be selected that are effective in addressing the behavior's function as interventions are function specific. A functional analysis approach has been extended to the identification of the most effective instructional contingencies for promoting student learning, which is called brief experimental analysis (BEA). BEA uses a single subject design to quickly test out and compare performance on a number of academic interventions that are briefly applied. Daly, Witt, Martens and Dool (1997) noted that BEA is efficient in identifying instructional adjustments and possibly decreasing the need for more intensive interventions.

Federal legislation such as No Child Left Behind and Individuals with Disabilities Education Act creates the need to directly link assessment to intervention (i.e., assessment that informs intervention). Instead of depending on categorical diagnoses, public education is moving toward using an approach that identifies how students respond to specific interventions (Wilber & Cushman, 2006). Therefore, the problem-solving mindset of functional analysis and its offshoot, BEA, became a useful tool (Daly, Martens, Hamler, Dool, & Eckert, 1999; McComas et al., 1996). BEA is a procedure that can be used to assist in the identification and selection of interventions that are likely to be effective for individual students (Wilber & Cushman, 2006).

The first formal definition of BEA (Daly et al., 1997) and the first utilization of BEA (Daly, Martens, Dool, & Hintze, 1998) were established in the late 1990's. Daly et al. (1998) compared how various reading interventions affected oral reading fluency (ORF) for individual students. The authors sought to discover effective contingencies (interventions) that would support student learning to produce gains in academic performance. Since the Daly et al. study (1998), BEA has continued to be used to identify instructional contingencies in the area of reading, as well as in math and written expression. The focus of BEA is on making decisions about how to approach problems when students are unresponsive to regular classroom instruction (Martens & Gertz, 2009). Among both educators and researchers, BEA has gained popularity as a useful tool to compare multiple interventions to see which one is the most effective, so that the intervention with the highest chance of success will be implemented.

A recent meta-analysis conducted by Burns and Wagner (2008) provided strategies for comparing results from a BEA and suggested three criteria for identifying effective interventions in the area of oral reading fluency (i.e., a raw score increase of 30 or more points, percentage of non-overlapping data ≥ 80%, and an effect size ≥ 2.80). They were able to identify the interventions for improving oral reading fluency that were most frequently successful and that are most frequently compared within BEAs. Such information is vital for informing research and ultimately the use of BEA in educational settings.

BEA can quickly identify the environmental contingencies (i.e., academic interventions) needed to support student learning (Daly, Andersen, Gortmaker, & Turner, 2006). However, as noted by Burns and Wagner (2008) and others (e.g., Martens & Gertz, 2009), inconsistency in the implementation of BEA has not resulted in a recommended or preferred model for implementation. Inconsistencies across studies has been noted in the type of single case design of the BEA (e.g., multiple baseline, alternating treatment, mini-withdrawal; Martens, Eckert, Bradley, & Ardoin, 1999), the inclusion of and type of extended analysis (Baranek, Fienup, & Pace, 2011), how interventions are sequenced and/or combined (Daly et al., 2006), the type of assessment materials (e.g., high content reading passage overlap versus instructional level reading passages; Burns & Wagner, 2008), and conceptual models used to generate hypotheses as to the reasons for poor academic performance. Use of BEA is emerging in other academic areas such as math (Carson & Eckert, 2003) and written expression (Burns, Ganuza, & London, 2009).

The purpose of the present investigation is to identify published studies since 2008, when Burns and Wagner published their meta-analysis, to assess what is currently known about BEA. Such a review can serve to determine the current practice in BEA implementation for comparison to the published studies prior to 2008. The outcomes can serve to inform use of BEA in educational settings and provide recommendations for further research.

Literature Review

Brief Experimental Analysis

A crucial element in student intervention is the implementation of a data-based instructional contingency (or contingencies) that can aid in performance. One strategy advocated by researchers in the field of school psychology is a procedure known as brief experimental analysis (BEA; Daly et al., 1997). A BEA procedure provides an efficient means of quickly comparing a student's response to multiple interventions with the purpose of finding the most effective one (Daly et al., 1997). BEA uses single-subject designs to test multiple variables (interventions) back to back, over a short time frame, with fewer data points per variable obtained than in an experimental study. While there may be less data collected on each variable, the data are sufficient to be able to compare different interventions to see which one has the most influence on academic performance.

The origins of BEA go back to its applied behavioral analysis roots (Martens & Gertz, 2009). Establishing experimental control within a study allows for the analyst to determine a causal relationship. In a standard experimental analysis, data are collected over an extended time period on specific variables to observe long-term impact on dependent variables. However, most schools do not have the resources or the time to execute a full experimental analysis that generally requires long-term data collection. BEA was developed to provide systematically obtained data for use in making instructional decisions regarding the intervention that is most effective for a student.

While BEA was originally proposed as a simple conceptual framework that applies functional analysis procedures to academic skills (Daly et al., 1997), its current use has been promoted within school psychology because it can also directly inform instruction or intervention. The procedure is designed to describe and formulate hypotheses about the relationship between environmental stimuli and behavior. A functional analysis of behavior would involve direct assessment of the behavior while potentially maintaining variables are systematically manipulated. A brief model for conducting an analysis of the functions of behavior is the functional behavior assessment (FBA), which is an abbreviated experimental analysis procedure that is currently used within educational setting to identify the cause or function of a child's behavior. BEA utilizes similar techniques in order to identify the causal relationships that will support academic performance.

Daly et al. (1998) were the first to use a brief experimental analysis to test out variables that enhance academic performance in the area of reading fluency. They used the five reasons for academic failure conceptual model (i.e., doesn't want to do it, insufficient practice, insufficient feedback, insufficient modeling, or too difficult) to formulate hypotheses for academic failure. BEA uses single case experimental design to evaluate the functional relations as they relate to the components of the conceptual model, which allows for the efficient comparison of multiple interventions prior to full scale implementation (Jones & Wickstrom, 2002).

The No Child Left Behind (NCLB 2001) legislation of 2001 and the reauthorization of federal special education law in 2004 (Individuals with Disabilities Education Improvement Act; IDEA 2004) promote a preventative approach to instructional supports rather than an intervention approach evident in prior legislation. Prevention focused frameworks like the Response To Intervention (RTI) model advocated in current legislation are reliant on assessment data. In particular, assessment data that directly informs instruction (Ysseldyke & Reschly, 2014). A major assumption of the RTI approach is that one cannot predict intervention effectiveness or which intervention will work best by the characteristics of the child (Petursdottir et al., 2009). School psychologists, in particular, can utilize BEAs as a tool for obtaining data on several instructional interventions used with a student that can serve to inform instructional decisions regarding the selection of contingencies that promote the greatest change in performance (Hosp & Ardoin, 2008). Since the onset of legislation advocating for linking assessment to intervention, the body of evidence based interventions has grown substantially. Academic interventions are highly accessible, as noted in the number of online clearinghouses for identifying, evaluating, and cataloging their effectiveness, such as the What Works Clearing House (http://ies.ed.gov/ncee/wwc/), Doing What Works (http://www2.ed.gov/nclb/methods/whatworks/edpicks.jhtml), Evidence Based Interventions Network (http://ebi.missouri.edu/), and Florida Reading Research Center (http://www.fcrr.org/). And there are multiple effective interventions for specific types of problems, such as reading fluency. The multitude of interventions, ranging from simple to complex, creates a problem: how do you select which intervention or combination of interventions is best suited for a particular student? BEA is a tool that can inform decisions as to which instructional contingencies are the most effective for a particular child.

Curriculum-based measurement (CBM) is an assessment tool that is best defined in terms of its attributes (Hosp, Hosp, & Howell, 2007). The assessment materials are closely aligned with content that is taught. They are reliable and valid, can reflect small changes in performance, and can be repeatedly given for multiple sessions. CBM is used in the continuous monitoring framework of Response to Intervention (RTI), its popularity gained from the expectations of NCLB and IDEA legislation. It is used within BEA to establish a current level of performance in a particular academic area. It is also used to assess the impact of an intervention via growth in skills due to instructional contingencies that target the specific academic performance deficit. BEA has mainly been employed to identify interventions in the area of reading and, within reading, most commonly to increase oral reading fluency (ORF). One reason for this specific focus is that learning to read is an important academic skill in which ORF has been identified as a meaningful measure for the construct of reading, as well as is a prerequisite to reading comprehension (Hosp et al., 2007). When children struggle to decipher words, their decoding interferes with comprehension and hinders their ability to explain what they have read (Daly et al., 2006).

Design of BEA

Developing hypotheses for poor academic performance. The selection of interventions used within a BEA is strategic, as it is based on a conceptual system that identifies the underlying causes for poor academic performance. The most frequently used conceptual frameworks for selecting relevant instructional contingencies have been the instructional hierarchy (Haring, Lovitt, Eaton, & Hansen, 1978), skill versus performance deficit (Lentz, 1988), and five hypotheses for student difficulty (Daly et al., 1997). The conceptualization of the academic failure allows for concise reasoning for the selection and implementation of specific instructional contingencies, as well as provide a rationale for why an intervention may or may not have been effective.

In the instructional hierarchy approach (Haring et al., 1978), the type of intervention or instruction should link the student's performance with one of four phases of learning. The four stages of learning are: acquisition (beginning to learn but not accurate or fluent), fluency (more accurate but slow), generalization (accurate and fluent but needing to extend range of use), and adaptation (accurate, fluent, and generalized with need to adapt/modify skill for specific tasks). Each phase of learning in turn has unique instructional needs. If a student is at the acquisition stage, for example, explicit instruction, modeling and immediate corrective feedback are most effective at improving performance. At the fluency stage effective strategies include independent practice, and immediate feedback on the speed of responding. These two stages (acquisition and fluency) are primarily utilized within BEA, which is why the instructional hierarchy conceptual model is evidenced frequently in identifying instructional contingencies affecting academic behavior. However, BEA has been used when student performance is at other levels of the hierarchy such as interventions for adaptation and generalization as well (Eckert, Ardoin, Daisey, & Scarola, 2000; VanAuken, Chafouleas, Bradley, & Martens, 2002; Wilber & Cushman, 2006).

The five hypotheses for student difficulty, as outlined by Daly et al. (1997), are: they do not want to do it (insufficient motivation), they have not spent enough time doing it (insufficient practice and insufficient active student responses to the curriculum), they have not had enough help to do it (insufficient prompting and feedback), they have not had to do it that way before (insufficient modeling or curricular materials), or it is too hard (materials are too difficult). Just as with the instructional hierarchy, each of the hypotheses represents a conceptual cause for poor student performance and each hypothesis is tied to specific instructional contingencies. If particular instructional contingencies address the hypothesized reason, the student's performance will improve. For example, if repeated reading is shown to be most successful, it indicates that the student has not had enough practice. Another example is that a successful response to a performance feedback intervention indicates that they need more explicit help and feedback from instructors. BEA allows for the comparison of instructional contingencies across all five hypotheses, as needed, in order to determine the cause of academic failure.

Lentz (1988) proposed a third conceptual model, the skill versus performance deficit. This model examines if the student's performance is because he or she does not know how to do it, or if the student chooses not to do it for another reason (e.g., anxiety, anger, frustration, and boredom). It is also possible that a student experiences both a skill and a performance deficit. The distinction of skill and/or a performance deficit results in interventions that are generally either focused on skill building or motivation. It is important to determine if the student has performance, skill, or a combination of skill and performance deficits, as the nature of the intervention varies accordingly. Many studies use this method to identify if performance deficits are due to motivation (Duhon et al., 2004; Jones & Wickstrom, 2002). If performance is identified as the reason for academic failure, the instructional materials do not need to change, but performance contingencies need to be added. When there are both skill and performance deficits, it is important to change the instruction as well as set up reinforcement contingencies for performance.

In a straightforward BEA, there are fundamental steps that formulate the process. First, a child is identified as having a problem. A student's performance data (e.g., classroom activities, homework, screening assessments) serve to determine the nature of the problem more specifically. Developing a BEA begins with analyzing the data using knowledge of the hierarchy of skill development in the academic content area (e.g., math, reading), strategies such as task analysis to identify prerequisite skills, and other relevant information about the student's performance. Based on that analysis, the conceptual models are consulted to develop a hypothesis or hypotheses about the child's performance, then instructional contingencies are chosen based on which ones are known to be effective or empirically supported in increasing performance within the area in which the student evidences deficiencies. Next, baseline data is collected utilizing CBM procedures. Then, each selected instructional contingency is applied for a short period of time in order to establish a clear pattern of effectiveness, or lack thereof (Daly et al., 1999).

Methodology. There are three basic approaches that have been used within BEA procedures for analyzing the problem of reading fluency (Daly et al., 2006). The first method is applying each intervention one-by-one until a clear pattern of improvement is established. A new baseline is then created in order to reintroduce the most successful intervention and confirm the results via replication. This one-by-one method is ideal in terms of focusing on one instructional contingency at a time, but it can take a number of sessions for the identification of an effective instructional component (e.g., only one instructional contingency such as learning passage preview or repeated reading). While this structure can provide relevant information, it only gives information for the individual instructional components one at a time, which does not generally encapsulate the extent of an intervention a teacher would implement. Another method of implementing BEA focuses on combining components by adding instructional contingencies sequentially so that they form a comprehensive package. Performance differences between the various combinations of instructional contingencies allow for systematically evaluating each new combination as a separate treatment package. The third method of implementing interventions is technically the opposite of the prior method, as comprehensive contingency packages are disassembled into smaller groups of components in order to identify the fewest instructional contingencies that produce the greatest results.

Instructional contingencies that are commonly utilized when addressing ORF are: contingent reward, performance feedback, student passage preview, repeated reading, and phrase drill (Burns & Wagner, 2008). The conceptual frameworks (i.e., instructional hierarchy, five hypotheses, and skill vs. performance) often provide the rationale for the particular instructional contingencies utilized. Interventions often times match with certain hypotheses (e.g., repeated reading for insufficient practice), which helps define which intervention(s) in a BEA should be implemented for the greatest student success.

Single-case designs used in BEA. Typical single-subject designs are reversal, alternating treatments, multiple baseline, or changing criterion design (Plavnick & Ferreri, 2013). However, the single-subject designs that have been primarily utilized in a BEA are mini-withdrawal (i.e., short reversals) and alternating treatments (Martens et al., 1999). Multiple-baseline designs, which stagger the time of treatment conditions across subjects have been sparingly employed in BEA studies (Noell et al., 1998).

The alternating treatments design is not dependent on the prior condition producing a steady response before applying the next intervention. Conditions are alternated according to a schedule outlined prior to the implementation. A strength of this design is that it can efficiently identify a particular intervention out of a group that has a high probability of effectiveness (Plavnick & Ferreri, 2013). One potential weakness in an alternating treatment design is the threat of multiple treatment interference, especially if multiple interventions are tried within a single day.

The mini-withdrawal is an alternating treatments design that consists of testing several instructional contingencies with a brief return to baseline, or least ineffective treatment, followed by a session of the most effective treatment (Harding, Wacker, Cooper, Millard, & Jensen-Kovalan, 1994). The purpose of the short reversal to baseline or to the least effective treatment is to assess if the student's performance drops significantly in the absence of the effective intervention. One can be confident in the effect of the treatment if the performance mirrors the withdrawal and reintroduction of the effective treatment, ruling out major influence from other variables that may have caused the changes. This replication provides evidence that a functional relationship exists between the treatment and the behavior.

Dependent measure(s). When it comes to measuring an intervention's success, or the dependent variable, CBM serves as a valid, reliable, and efficient method for obtaining performance data (Hosp et al., 2007). CBM procedures allow for time efficient, curriculum aligned assessments that are simple to administer and can adequately reflect small changes in performance. Within BEA, CBM is a versatile tool that can establish a current level of performance in a particular academic area, as well as to assess the impact of an intervention via growth in skills. CBM is used within the domains of reading, writing, and math. The crux of CBM is that it allows for a measurement of short-term growth of a student's performance when exposed to differences in instruction (or interventions) in a way that allows for an accurate estimation of growth (Deno, 2003). The use of global outcome and skills based CBM measures are used in BEA. CBM employs standardized administration and scoring guidelines that allow for consistency in scoring and interrater reliability.

Analysis of dependent measures. There is a history of inconsistency in researchers' criteria for success (Burns & Wagner, 2008). Various methods (i.e., post hoc visual analysis of graphed data, differentiation between performances on interventions, comparing the results of the most effective intervention to the least effective, calculating and comparing effect sizes, and a criterion equal to or more than 20% of an increase over baseline data) have been employed. The issue with these methods is that there is no way to compare directly results from study to study when different methods of data analysis are used. For example, some studies do not provide graphs or data, and other studies provide only one type of data. It is also inappropriate to compare outcome measures, such as percentage gain to PND. Therefore, research results from the studies that utilize BEA are often difficult to compare.

Meta-Analysis of BEA of Oral Reading Fluency Interventions

As previously stated, the purpose of BEA is to compare multiple interventions to see which one is the most effective, and then to implement the intervention that has the highest chance of success. However, among the published studies using BEA, variability in research designs, methods of evaluating outcomes, and methods of selecting interventions make comparison between these studies difficult. Burns and Wagner (2008) published a meta-analysis comparing the results from published studies that used BEA to identify interventions for ORF. Their goal was to better understand (a) what magnitude of treatment effect constituted as “effective,” (b) what effects were associated with interventions used within a BEA, and (c) whether or not effects were moderated by the type of reading passage used to assess ORF. This meta-analysis is the first comparative analysis of BEA studies, providing important information regarding the common factors amidst BEA studies analyzing ORF.

The studies included in this meta-analysis were selected using seven criteria that allowed for a more uniform comparison. The guidelines focused on studies with ORF as the dependent variable that had multiple interventions evaluating elementary school students with alternate form passages used across conditions. As long as the study had data to determine percentage of nonoverlapping data (PND) and/or effect sizes, it could be compared to other studies (Burns & Wagner, 2008). Using this criteria, 13 studies were ultimately chosen for the analysis. In addition, the data were further categorized by case as to the interventions used and the type of reading passage employed (High Content Overlap or HCO vs. instructional level). The interventions were categorized by passage types (HCO, instructional) because passage difficulty has been noted to impact reading fluency (Burns & Wagner, 2008).

Burns and Wagner (2008) addressed the first research question as to the magnitude of the treatment needed to be effective with three analyses: one that involved percentage of non-overlapping data (PND), one that used a no-assumptions effect size, and one that computed change in median fluency scores. When these analyses were not provided by the study authors, the raw data provided was used to calculate PND, effect size and percent change. PND is accomplished by taking the data and calculating the percentage of data, between baseline and intervention, that does not overlap. A general rule of thumb is that 90% PND is categorized as highly effective, 70% is moderately effective, 50% is minimally effective, and less than 50% is considered ineffective (Banda & Therrien, 2008). PND was computed for the intervention identified as most effective as compared to the other intervention conditions. When outcome data overlap with baseline data by 50%, the median score for each condition was calculated and used for comparison. These data were computed for each study and averaged across the studies. The outcome was 0% to 100% with a median of 81.83%.

Another method of evaluating intervention success is looking at a no assumptions effect size (Busk & Serlin, 1992). The no-assumptions effect size was computed by comparing the mean of the data for the most effective intervention and the mean of the data for each of the lesser effective interventions divided by the standard deviation of the less effective interventions. The range of the Cohen's *d* effect sizes was 0.25 to 12.02 with a mean of 2.86, indicating a large effect. However, the guidelines used by Burns and Wagner (2008) were intended for group designs, not single-subject research. Also, the rationale for using the no assumptions effect size is that there is no assumption of population distribution or equality of variances and covariance (Busk & Serlin, 1992), and allows for a comparison among studies in a meta-analysis. Therefore, Burns and Wagner (2008) recommended the results be interpreted with caution.

The last analysis calculated the actual percentage of change in the CBM ORF raw scores from baseline to intervention. The change in ORF raw scores was from 5 to 120 WRCM (words read correctly per minute) with the median intervention points ranging from 17 to 170 with a mean of 71.58. The mean change in fluency was noted to be 30.19 WRCM.

Burns and Wagner's (2008) second research question identified the effectiveness of each of the interventions attempted with BEA. Of the 18 interventions used, they identified, or computed, the mean and median PNDs and found that all but three of the interventions had median PNDs of 100%. Mean PNDs above 80% were found for 11 of the 18 interventions and intervention combinations. Based on this, Burns and Wagner (2008) concluded that the combination of listening passage, repeated reading, and performance feedback with and without incentives led to the largest PND.

The third research question was handled by grouping studies according to the type of passage used for the intervention and to determine the ORF outcome: HCO passages and instructional level passages. HCO passages consists of roughly 80% of the same words that were used in an instructional passage (Daly et al., 1998). The purpose of HCO is to measure the generalization of skills that are acquired during interventions in which instructional passages are utilized (Daly, Bonfiglio, Mattson, Persampieri, & Foremann-Yates, 2005). Instructional level passages are found using multiple grade level probes in order to find where a student meets "instructional level" (which comes before "mastery level"). Mean effect sizes were computed separately for each group. The HCO passages evidenced much less variability (*SD* = 0.96 for HCO versus *SD* = 2.47 for instructional level). However, the instructional level passages had much larger effect size (*d* = 2.86 for Instructional passages versus *d* = 1.68 for HCO).

Based on the findings of this meta-analysis, Burns and Wagner (2008) concluded that an average no assumptions effect size of 2.80 and a PND of 80% could serve as a potential criterion for identifying an intervention as effective. They translated these data to make the recommendation of an increase in 30 WRCM and/or 73% above baseline as indication of an effective intervention. Burns and Wagner state that these criteria are more of a suggestion for future research and implementation of BEAs than an absolute requirement. They also state that an increase in 30 WRCM could be used to compare two interventions conducted within one session, but do not offer much insight into long-term performance. Overall, Burns and Wagner caution that these suggested criteria come from an analysis of only 13 articles, and should not be automatically taken as the gold standard for BEA. However, it does provide some guidelines for use for researchers and practitioners when evaluating the outcome of a BEA.

While the outcomes addressing the hypotheses explored by Burns and Wagner (2008) are important to note, they highlight four limitations to be considered when interpreting the results: (a) the meta-analysis was limited to BEAs for ORF; (b) some studies only looked at one intervention and others compared two or more; (c) only studies that used instructional-level and HCO passages were used to answer the first two research questions; (d) there was not a consistent methodology among studies.

Burns and Wagner's (2008) meta-analysis offered some general criteria and direction for selecting effective interventions within a BEA addressing ORF, as well as provided some general observations. The framework they have provided gives a starting point for examining more current literature, but there are other aspects of BEA (e.g., extended analysis, design of BEA, other academic content areas) that may provide information on the utility of BEA, as well as give insight into any trends among studies that have become popular.

Inconsistencies across BEA Studies

The meta-analysis conducted by Burns and Wagner (2008) was limited to the content area of reading, more specifically ORF. While published studies using BEA have been most evident for selecting ORF interventions, researchers have also looked at other academic domains. Other areas investigated include letter formation (Burns et al., 2009) and mathematics (Codding et al., 2009). Other than the variability in the academic areas, there is also variability in the design and implementation of BEAs addressing the same content area, as noted in Burns and Wagner's 2008 meta-analysis.

One source of variability noted by Burns and Wagner (2008) was the type of reading passage used to assess ORF. A vast majority of the articles reviewed in Burns and Wagner's (2008) meta-analysis utilized the Silver, Burdett, and Ginn reading series (Pearson et al., 1989) for the instructional level passages for the curriculum-based measurement. Since 2008, many studies continue to use the Silver, Burdett, and Ginn reading series, with a few studies that chose to utilize other CBM (e.g., DIBELS), or CBM that is custom created to have high content overlap for their particular study (Martens & Gertz, 2009). As previously noted, the type of reading passage (HCO or instructional level) does influence the magnitude of change reported. In terms of the passages used, Burns and Wagner discussed how many articles within their analysis were HCO or instructional. Some studies utilized both, whereas others only used one for the study. Since the outcomes in terms of strength of results vary according to the type of passage employed, comparison of results is difficult. It is not known if there is a recommended type of reading passage.

While the purpose of BEA is to compare multiple interventions to see which one is the most effective (and implement the intervention that has the highest chance of success), the brief nature of a BEA does not allow for an extensive investigation of each individual intervention. A common addition to BEAs is the use of an extended analysis to validate the effects of select interventions; however, the method of implementation of the extended analysis varies across studies. Extended analysis is similar to a mini-reversal in that there is a return to baseline, but the extended analysis consists of a longer period of time of implementing the most successful intervention and baseline (or least effective intervention) conditions. The extended analysis is administered after all interventions have been administered.

To identify the intervention that is the most effective, the CBM results for each intervention are usually compared to baseline (Burns et al., 2009). However, an extended analysis by VanAuken et al. (2002) compared the intervention with the highest effect to the intervention with the lowest effect after the BEA. Another use of an extended analysis is to take the interventions identified from the BEA and put them together in different combinations to assert which components create the greatest gain, as well as which ones could potentially be impeding a child's progress. This analysis essentially consists of testing individual intervention components via BEA with the intent of re-testing them as packages (Daly et al., 2006). In Burns and Wagner's (2008) meta-analysis, extended analysis was noted to be used to further investigate the utility of the selected interventions, but they did not set a criterion of what the extended results should look like and noted a lack of consistency in the application of an extended analysis.

While there are many studies that indicate that BEA helps identify interventions that lead to above-baseline performance, there is limited research that focuses on the investigation of BEA's sensitivity in detecting differential effectiveness between the interventions (Baranek et al., 2011). Interventions used within a BEA are selected based upon their ability to address the hypothesized problem and their known ability to improve performance. Therefore, interventions selected have been shown to produce increased performance at the group level and have a high probability of addressing the problem at the individual level. As a result, it is a typical finding in the BEA literature that more than one intervention produces an increase in performance (Daly, Murdoch, Lillenstein, Webber, & Lentz, 2002; Eckert, Ardoin, Daly, & Martens, 2002). Therefore, understanding BEA's ability to detect differential sensitivity is important. This would allow for a more comprehensive idea of the effectiveness or ineffectiveness of specific interventions for a particular student. A better understanding of the utility of BEA may lead to more implementation among practitioners in the future.

Burns and Wagner (2008) indicated that the implementation of BEA varies, especially in the areas of extended analysis, type of dependent measure (instructional level or HCO), and the content areas addressed (reading, math, writing). However, they did not conduct a direct analysis of these features. The problem this creates for practitioners is that there is no common thread of methods that has been explicitly outlined when looking at the BEA literature as a whole. Practical implementation may be more highly evident if there are guidelines to follow when conducting a BEA. Also, common techniques/methodology would allow for a more effective comparison of studies.

Purpose of the Present Study

While BEA allows for an efficient and effective method for assessment to inform intervention, Burns and Wagner (2008) discovered that there was no consistent method of examining the outcome data from a BEA. Burns and Wagner outlined some of the inconsistencies among the implementation of BEA that made it difficult, or in some cases impossible, to compare features across studies. However, their discussion provided some suggested methods for the analysis of BEA data and criteria for evaluating success based on the no-assumption effect size and percentage of non-overlapping data (PND). In addition, Burns and Wagner (2008) and others have pointed out inconsistences in the methods used to implement BEAs.

It has been eight years since Burns and Wagner (2008) published their study. The purpose of this study is to perform an analysis of the literature to determine what is currently known about BEA, and the current trends in BEA. This study will build on Burns and Wagner's findings and recommendations by examining more closely the BEA design, methods, methods used for analysis of results, and the use of extended analysis. Articles utilized within Burns and Wagner's (2008) meta-analysis will be revisited in order to review information that Burns and Wagner did not directly analyze and compare it to information from studies posted after the study was published in 2008. In addition, the articles that implement BEA in other academic areas will be analyzed and compared to the BEA studies conducted in the area of ORF to determine the consistency in implementation of BEA across academic areas. A preliminary search for studies using BEA published since 2008 indicates the continued publication of articles on BEA and the need to further assess, analyze and integrate the current findings and the nature of the

BEA studies. The addition of analyzing the content areas of math and writing also lend more information in regards to the flexibly and generalizability of BEA procedures. Obtaining data on the criteria for success, validity of intervention effectiveness through extended analyses, and similarities in studies across content areas should culminate into a contemporary understanding of the types and uses of BEA, which may potentially yield recommendations for the practitioner on the use of BEA. The following research questions form the foci of the present study.

1. How are BEA data being analyzed? The variety of methods for evaluating the outcome of various ORF instructional contingencies discussed in Burns and Wagner's (2008) meta-analysis suggest that the strategies at the time were a far cry from being consistent. However, Burns and Wagner's article did offer a model of three different methods to analyze the results of BEA: PND, percentage gain, and effect size. This research question will examine the types of analyses used within BEA studies after the publication of Burns and Wagner's (2008) meta-analysis.

2. How comparable are the BEA methods and designs across studies? Variability in the design of the BEA, the sequencing of the interventions, the conceptual models employed to develop the intervention, types of dependent measures used have all been noted and will be compared. Additional areas to be compared also include number of interventions compared, baseline determination, criteria for identifying interventions as successful, and strategies for intervention delivery.

3. How are extended analyses being conducted? Burns and Wagner (2008) indicated that extended analysis is a step used after a BEA to further analyze and validate the interventions selected by the BEA. The implementation of the extended analyses was noted by Burns and Wagner to lack consistency in terms of length and use. However, Burns and Wagner did not examine the utility of extended analyses in regards to how well they supported the findings of the BEA, as well as if there was a general standard for what to expect from extended analysis.

4. What is the status of BEA in academic areas other than reading? The utilization of BEA in areas outside of reading was noted in the preliminary identification of studies since 2008. The outcomes evident in other academic areas and the conceptual models used, methods of implementation, analysis of results and design of the BEA can be useful in determining the generalizability and comparability of BEA to other academic content areas.

Method

Data Collection

The Academic Complete, Education Full Text (H.W. Wilson), ERIC, MasterFILE Premier, Primary Search, PsycARTICLES, Psychology and Behavioral Sciences Collection, PsycINFO, and Teacher Reference Center was searched through WKU Libraries using the search terms “brief experimental analysis,” with “reading,” “mathematics,” and “writing” in order to capture all articles that include these key terms. The key terms used by these articles were then used to conduct additional searches. The search was limited to articles published between 1994 and 2016 for comparison of studies before and after Burns and Wagner’s meta-analysis (2008). The criteria for inclusion of a study for the current analysis is as follows:

1. The study utilizes BEA in any academic area.
2. The study involved children in grades Kindergarten through six.
3. The study was published in a peer-reviewed journal.
4. The study presented quantitative data (i.e., CBM scores, non-overlapping data (PND), percentage gain, and/or effect sizes) or graphs that can be converted to quantitative data.
6. The study was written in English.

A total of 60 articles were initially identified with 45 meeting criteria for inclusion in this study. The studies that were excluded were duplications or articles about BEA that did not conduct a BEA. The references of the articles meeting criteria were reviewed in Web of Sciences (a citation indexing service) for the purpose of searching for any other relevant articles that the original search may have missed.

Results

Description of BEA Articles

The 45 published articles meeting the study criteria generated the data to address the research questions. The journals in which the 45 articles were found are presented in Table 1 according to their publication before or after Burns and Wagner's 2008 meta-analysis. The 45 articles come from 16 different journals crossing the areas of school psychology, applied behavior analysis, education, and learning disabilities. Prior to 2008 there were 15 behavioral journals (57%), eight school psychology journals (31%), two education journals (8%), and one learning disability journal (4%). The articles after 2008 were published in eight school psychology journals (42%), nine in behavior journals (47%), and two in education journals (11%). These percentages indicate that there has been a proportional decrease of BEA articles in behavioral journals with a slight increase in school psychology and educational journals in the 2009-2016 group of studies.

Three of the articles identified were two part studies where each part was a separate BEA. For analysis purposes each separate BEA was treated as a separate study, bringing the number of BEA studies up to 48. In this document, published articles will be used to refer to the 45 journal articles in which the BEA studies were found and BEA studies will be used to refer to the 48 separate BEAs within those articles. Of the 48 BEA studies, 37 were in reading, seven were in math, and four were in writing. There were 28 BEA studies published prior to 2008 and 20 BEA studies published from 2009 up to May of 2016. The average number of published articles was 1.73 per year from 1994-2008 and 2.38 published articles per year from 2009-2016.

Table 1

Number of Research Articles on BEA by Publication

Journals	Before 2008	2009-2016
Assessment for Effective Intervention	1	0
Behavioral Interventions	1	0
Behavior Modification	0	2
Contemporary School Psychology	0	1
Education & Treatment of Children	1	0
Electronic Journal of Research In Educational Psychology	0	1
Journal of Applied Behavior Analysis	8	1
Journal of Behavioral Education	6	6
Learning Disability Quarterly	1	0
Journal of Applied Psychology	0	1
Journal of Applied School Psychology	1	2
Journal Of Direct Instruction	0	1
Journal of Evidence-Based Practices for Schools	1	1
Journal of School Psychology	1	1
Psychology in the Schools	2	2
School Psychology Review	3	0
Total	26	19

Research Question One: How are BEA Data Analyzed?

Data on how the most successful intervention(s) were identified and the criteria for the selection were gathered from each study. For each of these analyses the BEA studies were grouped so that they can be compared before and after the publication data of Burns and Wagner's (2008) meta-analysis.

Method of determining "best" intervention(s). The methods of analyzing the results of the BEA to determine the outcome were placed into six categories including visual analysis of graphed data to determine the highest scoring intervention, the percentage of non-overlapping data between the highest baseline and intervention (PND), and meeting a specific percentage increase over baseline. A measure of effect size, a method of determining intervention success suggested by Burns and Wagner (2008), was not utilized by any of the studies in the selection of interventions post-BEA. Other methods identified include selecting the intervention with the highest mean score and meeting a pre-selected goal created by the researchers (e.g., reaching a specific score in words read correctly per minute (WCPM) or meeting grade level CBM scores). Figure 1 displays the frequency of use of each of these strategies.

The use of visual analysis to analyze BEA data decreased slightly between the 1994-2008 group (71%) and the 2009-2016 group (60%). Of the other methods there was a slight decrease in use of percentage gain over baseline and a slight increase in the use of PND in the BEA studies published 2009-2016. Based on the data, there has been a slight shift in proportionality of methods for selecting the best intervention(s) from the BEA post Burns and Wagner's 2008 meta-analysis.

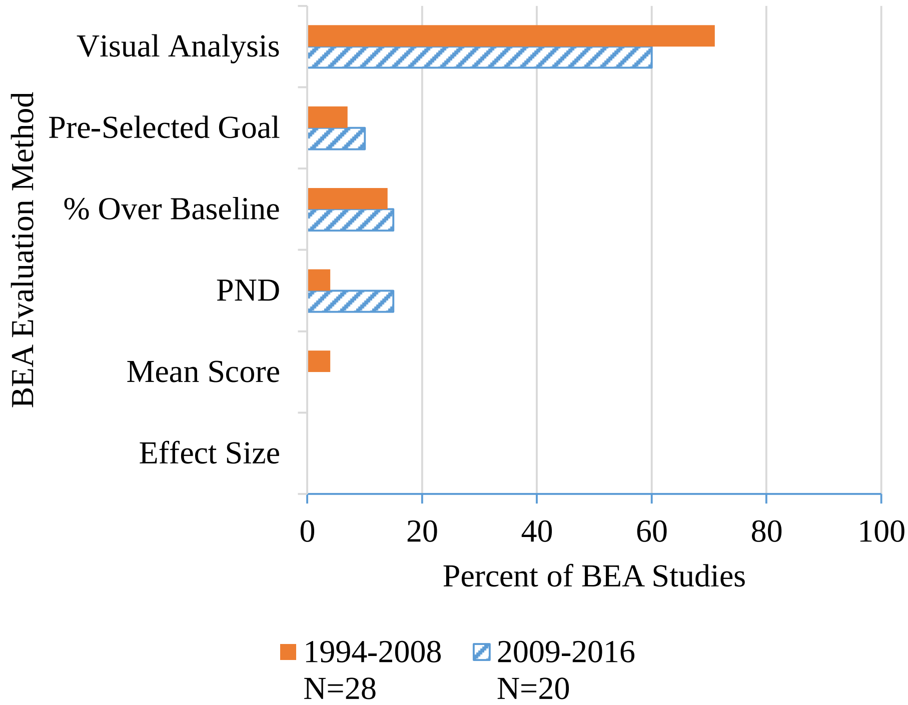

Figure 1. Analysis methods used to evaluate BEA data. Visual analysis = visual analysis of graphed or charted data points; pre-selected goal = a goal was set prior to the BEA; PND = percentage of non-overlapping data.

Criteria for "best" intervention. The criterion or criteria used to assess if an intervention was successful varied. Figure 2 shows the criteria that were used when visual analysis was not the method for selection of the intervention. Only 16 studies did not rely on visual analysis for selecting interventions, and are therefore the only studies displayed in Figure 2. From 1994-2008 there were six studies that did not utilize visual analysis, whereas from 2009-2016 there were 10 studies. The criteria were categorized into seven groups based on the authors' description and include a PND greater than or equal to 70%, the highest scoring intervention, greatest mean score, a predetermined target WCPM, expected instructional level of students' grade, a 20-30% improvement over baseline, or 50% improvement over baseline. None of the studies that utilized PND used Burns and Wagner's (2008) recommendation of 80%, with all using a slightly lower 70%. Of the studies that used percentage over baseline, none of the studies in either group used Burns and Wagner's recommended criteria of 73% over the median baseline score. The 50% over baseline was only noted in studies that utilized the Skill versus Performance conceptual model.

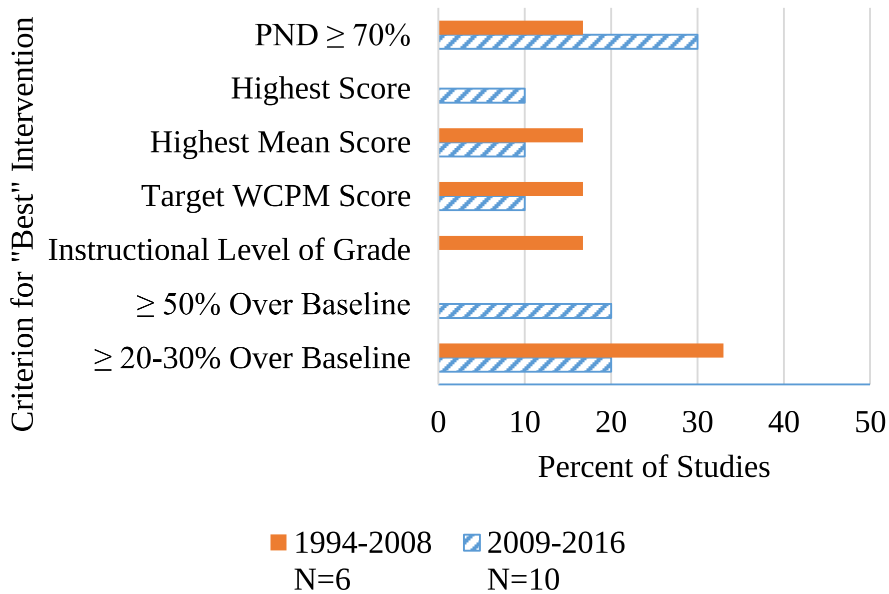

Figure 2. Percent of each evaluation method criterion when visual analysis was not used. PND ≥ 70% = percentage of non-overlapping data of 70 or higher; Highest Score = Intervention with highest CBM score. Highest Mean Score = Intervention with the highest mean score; Target WCPM Score = pre-selected CBM score. Instructional Level of Grade = score within instructional range for grade level; ≥ 50% Over Baseline = over 50% improvement with contingent reinforcement used; ≥ 20-30% Over Baseline = intervention with 20-30% over baseline deemed effective.

The methods used to determine intervention effectiveness can be further grouped into methods that use a comparison to baseline performance (intra-individual comparison) and those that set a criterion based on normative performance (intra-individual comparison). The PND and the 20%, 30% and 50% gain over baseline methods make up the comparison to baseline methods. The instructional level, highest mean score, highest mean and target WRCPM score make up the set criterion methods.

The criteria of the 1994-2008 studies are split (50/50) between baseline comparison methods and set criterion methods. The 2009-2016 articles weigh heavier on the baseline comparison (70%) than the set criterion methods (30%). Thus, a slight change to greater use of baseline comparison methods is evident in the BEA studies from 2009-2016.

Research Question Two: How Comparable are the BEA Methods and Designs?

Information was obtained from each article regarding the grade level of students number of interventions and intervention packages compared, number of baseline and intervention data points obtained, sequence of interventions, dependent variables, and CBM utilized. Data from each of these analyses were grouped so that the articles can be compared before and after the publication of Burns and Wagner's (2008) meta-analysis.

Grade level of students. The grade levels of students were categorized by the range of grade levels that each study contained. The results of the analysis can be seen in Figure 3. Half of the studies (50%) from 1994-2008, and slightly over half (55%) from 2009-2016 consisted only of one grade level within the BEA. Grade levels of students spanned from kindergarten to 6th grade. The most frequent grade levels from 1994-2008 were 3rd grade (36%), 4th grade (24%), and 2nd grade (21%). In the 2009-2016 studies the most frequent grade levels were 3rd grade (32%), 2nd grade (29%), and 4th grade (15%).

Conceptual model. The conceptual models fell into six categories and included the three models discussed in the literature review: instructional hierarchy (IH), five hypotheses for poor performance, and skill versus performance. Two additional conceptual categories were noted. One was identified as a problem-solving (PS) model as the studies provided information about the student's performance deficits based upon

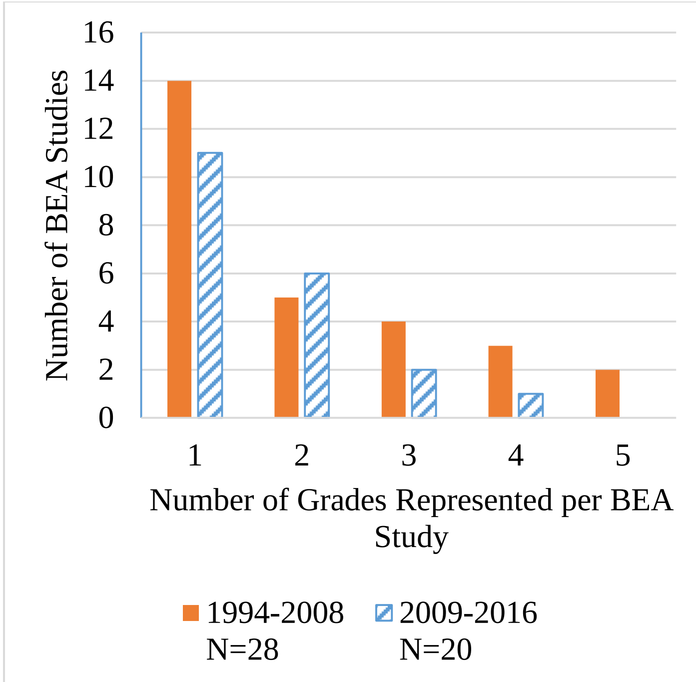

Figure 3. Range of grade levels of students within each BEA. The number of grades corresponds to the range of student instructional level(s) included in the BEA study.

data (e.g., classroom performance, screening specific to the study, school wide progress monitoring, task analysis) along with a logical connection to an evidence-based intervention. The last category was labeled no model (NM) and was used for studies that did not indicate any basis for selection of the interventions used. The conceptual models were identified as explicit if the article specifically stated the model and labeled as implicit if a description was provided in the article that was clear and consistent with a model, however, the model was never named. The diversity of the conceptual models across the studies is shown in Figure 4. Half (50%) of the studies post-2008 utilized the problem-solving model while only 14% of the models pre-2008 were problem solving. Also, conceptual models were evident for 100% of the studies post-2008 as opposed to 89% pre-2008. A vast majority (93%) of the conceptual models were explicit post-2008, as opposed to pre-2008 (58%). Also 43% of the 1994-2008 studies were implicit while

10% of the 2009-2016 studies were implicit. Therefore, the 2009-2016 had explicitly identified a higher proportion of conceptual models used for selection of interventions.

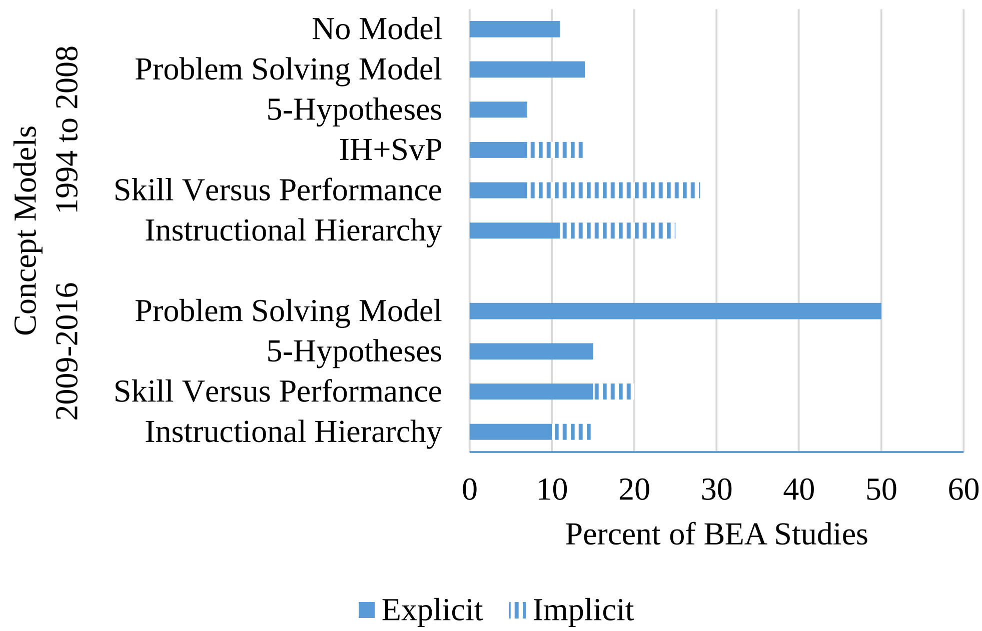

Figure 4. Conceptual models used to select interventions to be compared in each BEA. Studies were grouped by date of publication with 1994-2008 (*N*=28) and 2009-2016 (*N*=20). Problem-Solving Model = data provided that linked individual's performance deficits to an evidence based intervention; 5 Hypotheses = the five hypothesis for academic failure; IH+SP = Instructional Hierarchy combined with Skill versus Performance; Explicit = named that model; Implicit = description provided consistent with that model, but no model was named.

Single-Subject Design of the BEA. The design of the BEA studies consisted of five categories including alternating treatments, mini-reversal, multiple-baseline, and multiple probe. Alternating treatments is alternating interventions in quick succession. Mini-reversal consists of administrating baseline condition after administering all interventions and then proceeding with the most successful intervention to confirm the effects. Multiple-baseline refers to staggering the conditions across time for each participant. Multiple Probe involves testing multiple interventions per session. A fifth category, AB Design, was added due to studies only testing effects of motivation against control in skill versus performance. Figure 5 depicts the frequency of each design employed. The majority of BEA studies from both 1994-2008 (71%) and 2009-2016 (60%) utilized an alternating treatments design. The second most prevalent method for both 1994-2008 (14%) and 2009-2016 studies (35%) was mini-reversal. Proportionately there was a slight decrease in alternating treatments and a slight increase in mini reversals for the 2009-2016 studies.

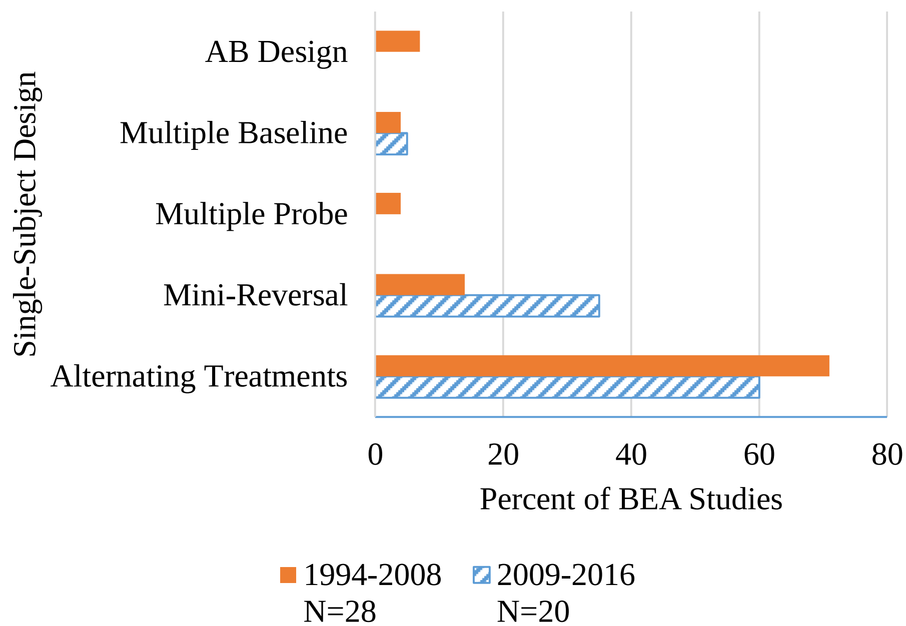

Figure 5. Single-subject design of BEAs. Alternating treatments = alternating interventions in quick succession; mini-reversal = administration of baseline condition after administering all interventions followed by the most successful intervention; multiple-baseline = conditions staggered across time; Multiple Probe = testing multiple interventions per session; AB Design = Only testing effects of motivation against control in skill versus performance procedure.

Number of interventions compared. Examining the number of interventions compared within each BEA was done by gathering frequency data on the interventions each BEA study used. Figure 6 represents how many interventions were in each BEA study. However, half (50%) of the BEA studies included single evidence based interventions that were packaged and delivered together as one instructional contingency. If a study combined multiple instructional contingencies together to form an instructional package, each instructional package was counted as one intervention in the analysis. When looking at the data proportionate to the number of studies in the pre- and post-2008 groups, they were roughly equivalent when looking at one to four interventions (79% and 80%) and five to eight interventions (21% and 20%).

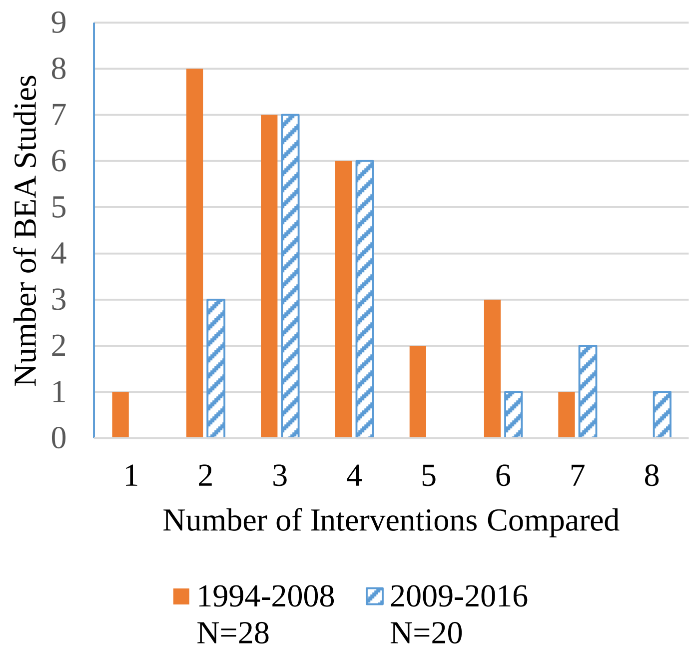

Figure 6. Number of interventions compared in a BEA study. Intervention packages were treated as one intervention.

There were 24 studies that included instructional packages (17 from 1994-2008 and seven from 2009-2016; see Figure 7). Instructional packages can be distinguished from multiple instructional contingencies in that dependent measure was obtained after each instructional contingency for multiple instructional contingencies and at the end of all of the instructional contingences for a package. Instructional contingency packages were compared to other instructional contingency packages as well as with a singular instructional contingency or contingent reinforcement.

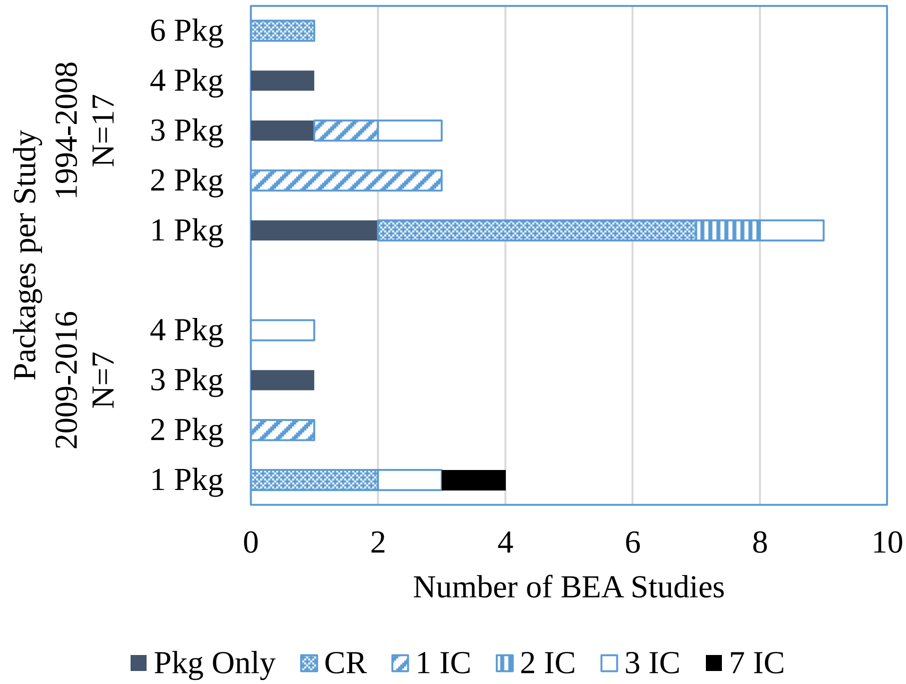

Figure 7. Frequency of instructional packages used in BEA studies and the additional interventions included. Pkg Only = Only packages were compared; CR = Packages were compared to contingent reinforcement; 1 IC = Package(s) were compared to 1 instructional contingency. 2 IC = Package(s) were compared to 2 instructional contingencies; 3 IC = Package(s) were compared to 3 instructional contingencies; 7 IC = Package was compared to 7 instructional contingencies.

There were predominantly two to four interventions utilized in BEA studies for both 1994-2008 (71%) and 2009-2016 (80%) groups. Proportionally, there were more instructional packages utilized in studies in the 1994-2008 group (61%) over the 2009-2016 group (35%). The 1994-2008 group of studies had a majority of BEA studies that compared one package with one to three single instructional contingencies.

Number of baseline data points. As represented in Figure 8, the frequency of baseline data fell into eight categories including 1-point, 2-point, 3-point, 4-point, 7-point, prescreen data, variable, and no information. For the prescreen data, no baseline was established in the process of the BEA, but instructional level was assessed prior to the BEA. Variable data meant that the baseline data varied from subject to subject due to the researchers administering baseline conditions until they established a steady response. Some of the studies were categorized as N/A because baseline was not applied due to using a set criterion or conducting the BEA with the purpose of observing the effects on a special population. The frequency of baseline data can be seen in Figure 8. One point of baseline data was used by most studies in both the 1994-2008 (57%) and 2009-2016 (55%) groups.

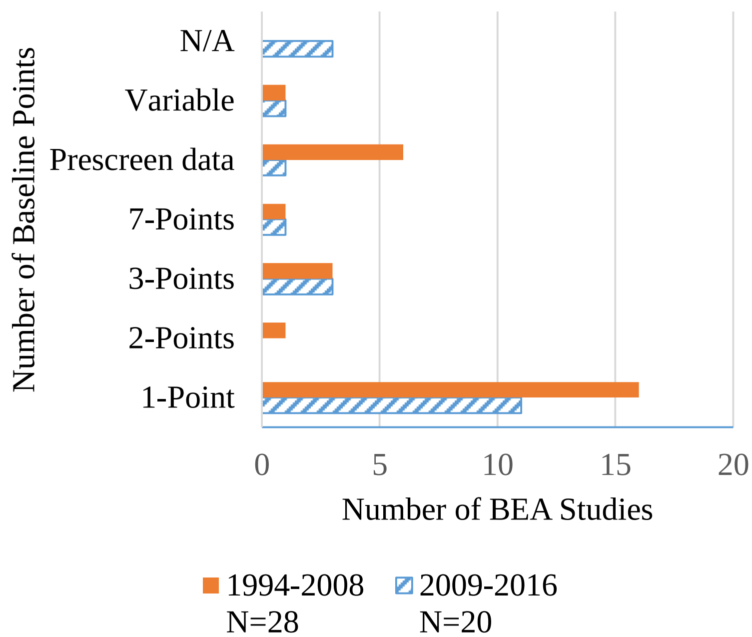

Figure 8. Frequency of baseline data points obtained within the BEA studies. Prescreen data = data was collected prior to the BEA and not included as a baseline point; Variable = baseline varied from individual to individual in order to identify a steady response.

Intervention delivery. The order in which the interventions were implemented was categorized into five groups including easy to hard (implementing interventions from least to most components as well as least to most adult effort) and randomized sequence. There was a suggested technique noted in the literature review that was not found in these BEA studies that consisted of deconstructing packages in order to test interventions from most complex to simple components. However, there were three additional categories noted including standard order, semi-random, and not applicable (N/A). Standard order (SO) sequencing meant that there was an order standardized across participants. Semi-random meant that a sequence started random, but the sequence was altered for each intervention session so that the same treatment would not be administered back to back in two consecutive sessions. Finally, one group of BEA studies either did not test multiple interventions or only manipulated one independent variable. These studies were included in a category labeled not applicable. The frequency of each intervention order category can be seen in Figure 9. The most frequent order of interventions used in BEA studies from 1994-2008 was easy to hard (39%), whereas the most frequent order from 2009-2016 was random (45%).

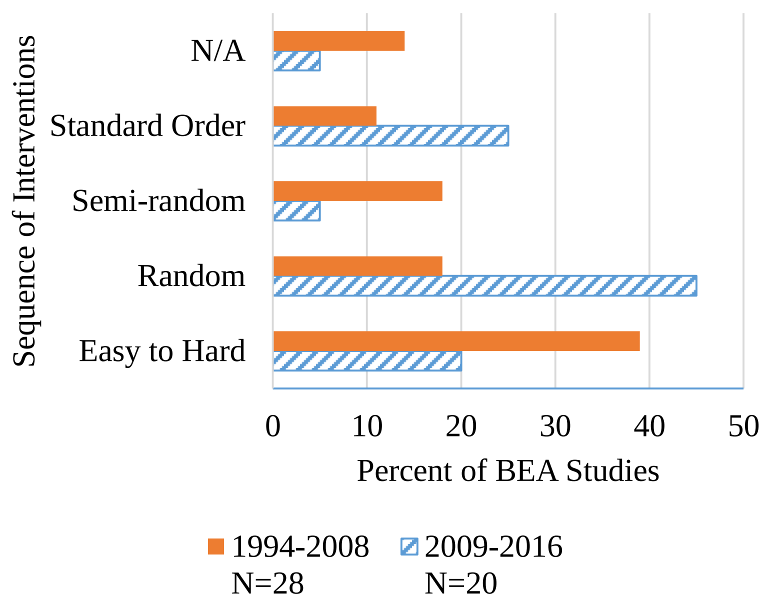

Figure 9. Percent of each intervention sequence used in the BEA. Easy to Hard = implementing interventions from least to most least to most adult effort and/or components; Random = randomized sequence; Standard Order = order standardized across participants; Semi-random = sequence started random with counter-balanced administration. N/A = not applicable to the study.

Number of intervention data points. The number of intervention data points are categorized by how many times each intervention was employed in the BEA. Frequency of the interventions could also be variable across participants if the procedure for determining the most effective intervention varied due to interventions being implemented until one intervention evidenced differentiation from the others. Proportionately, post-2008 there was an increase in use of one data point (55% versus 32%) and differentiation (49% versus 29%). Figure 10 represents the number of interventions points acquired for each intervention within a study.

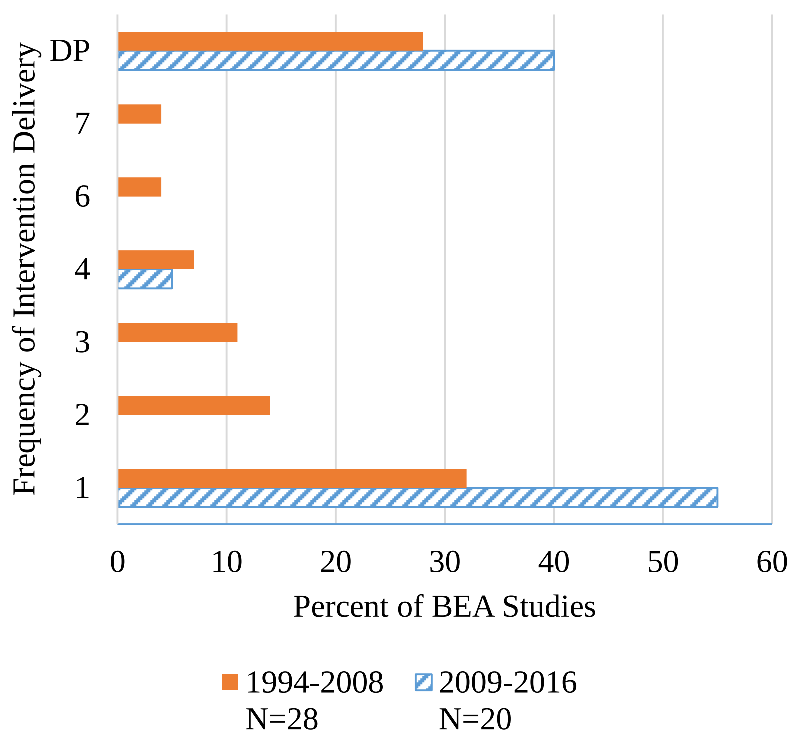

Figure 10. Percent of studies using specific frequencies of each intervention implemented in the BEA. DP = differentiation procedure implemented where intervention frequency varied from individual to individual with the purpose of identifying a steady response.

CBM utilized. Across all the studies, curriculum based measurement methods were used at the dependent measure. Measurement probes were either developed from basal reading series materials or selected from a packaged CBM system that supplies the probes to use in instruction and assessment. The sources of CBM were categorized by either the commercial CBM package or method in which they were created. The three categories of CBM were basal reader series, CBM package, and custom probes created by the researchers. Silver, Burdett, and Ginn (SBG), Ginn, and Harcourt Brace Jovanovich Treasury of Literature (HBJTL) are examples of basal reading series from which the measurement probes were drawn. AIMSweb, Edformation and DIBELS are commercially available CBM packages that are independent of the basal series and used for monitoring instruction, however, they were used for both instruction and monitoring in the BEA studies. In addition to these sources a custom category was added for any material that was created for the study such as story starters or math probes generated for a specific skill level. There was also an unknown category in which the material was only described as CBM. The frequency of the sources can be seen in Figure 11. Prior to 2008, 75% of the BEA studies used basal reading series to develop instructional and measurement probes whereas after 2008 most measures were CBM package (35%) and custom (50%).

Dependent variables. There were a total of 6 different categories of dependent variables across the studies spanning over all reading, math, and written expression. The categories include fluency, fluency and accuracy, fluency and mean score, accuracy, total score, and total score and accuracy. All studies that measured fluency variables included words correct per minute (WCPM) or digits correct per minute (DCPM), depending on the academic area. A majority of fluency-based dependent variables looked at general outcome measures. However, it was interesting to note that a subset of studies looked at

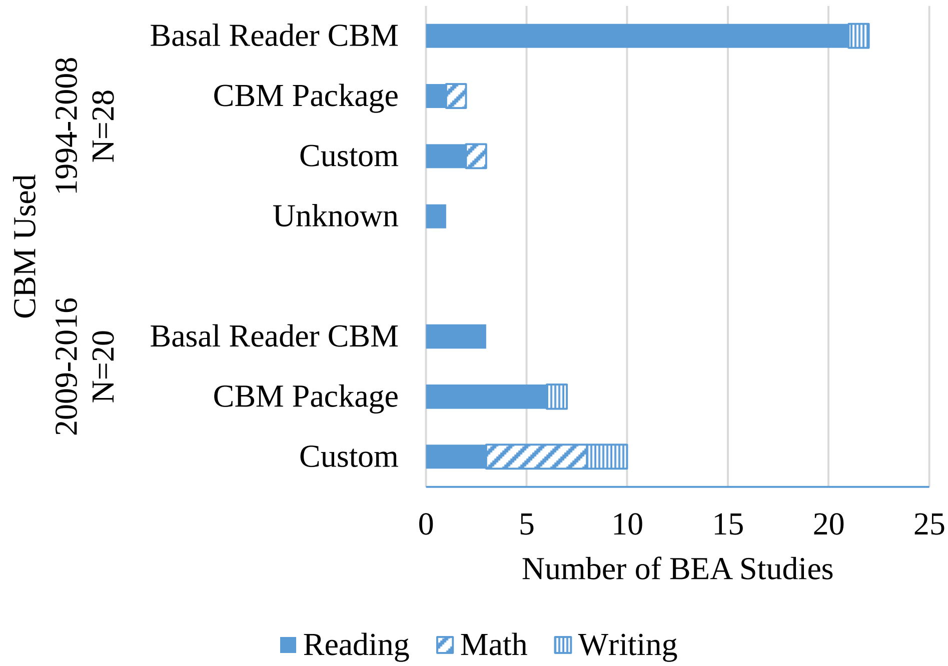

Figure 11. Sources of CBM across BEA studies. Basal Reader = instructional and measurement probes developed from published reading series; CBM Package = published CBM monitoring probes used for both instruction and monitoring; Custom = CBM utilized was created by researchers; Unknown = information was not provided.

fluency using a general outcome measure including a subskill fluency measure (i.e., letter sounds correct per minute, correct letter sequence). Studies looking at accuracy analyzed percentage of words correct, percentage of correct responses to comprehension questions, percentage of correctly formed letters, or percent of correct letter sequence. One study looked at fluency and mean reading comprehension levels. Total score refers to a total count for a specific measure that were not measured in terms of scores per minute (i.e., digits correct or words spelled correctly). Figure 12 depicts the frequency of each type of dependent variable group. The majority of 1994-2008 studies (64%) and 2009-2016 studies (60%) measured fluency alone.

Research Question 3: How are Extended Analyses Being Conducted?

After the completion of the BEA, most of the studies employed an extended

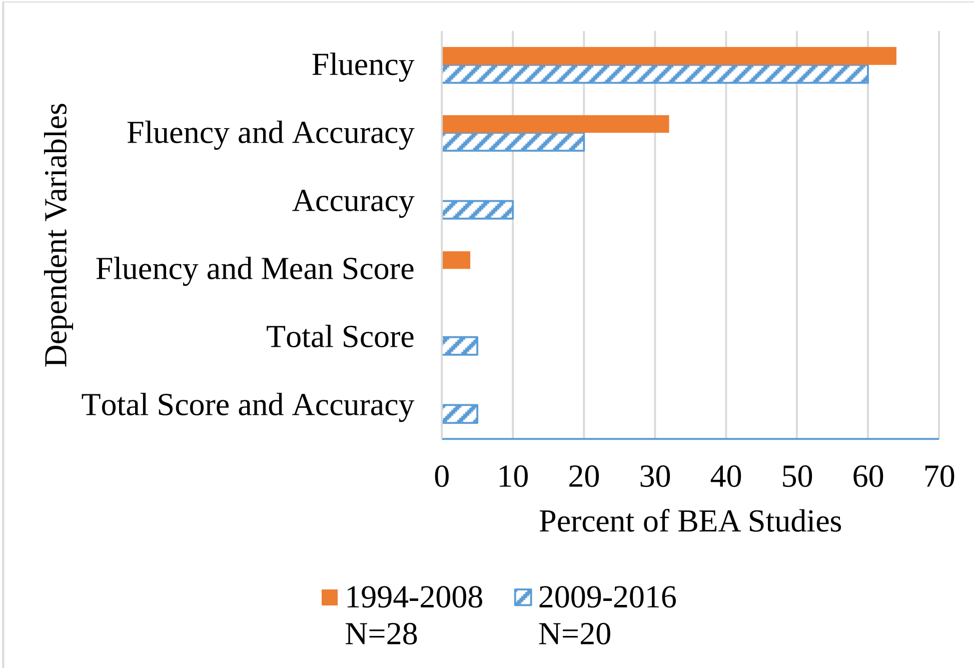

Figure 12. Percent of BEAs using each type of dependent variables. Fluency = variables included words correct per minute (WCPM) or digits correct per minute (DCPM); Accuracy = percentage correct of words responses to comprehension questions, correctly formed letters, or correct letter sequence; Fluency and Mean Score = WCPM and mean reading comprehension levels. Total score = total untimed score for a specific measure (i.e., digits correct or words spelled correctly).

analysis. The purpose of the extended analysis and the extent of the data collected in the extended analysis varied across the BEA studies. In general, an extended analysis is a procedure implemented subsequent to the BEA where one or a subset of the interventions is again implemented but over a longer period of time. In practice, extended analysis is employed to verify that the intervention(s) selected as most effective by the BEA continues to show the desired change. Amidst the BEA studies prior to Burns and Wagner's meta-analysis (2008), 10 of the 28 studies (36%) did not conduct an extended analysis. In fact, the first 5 studies chronologically from 1994 through 2000 did not include an extended analysis. Of the studies published after Burns and Wagner's meta-analysis (2008), only 2 of the 20 studies (10%) did not include an extended analysis.

Therefore, there were 18 BEA studies from 1994-2008 and 18 BEA studies from 2009-2016, for a total of 36, that serve as the data for the information on the use of extended analyses.

Design of extended analysis. In terms of design of the extended analyses, there were five categories of design types. The designs were either categorized as baseline followed by intervention sessions, intervention only, multiple baseline, or alternating between the interventions and baseline. The most frequently employed design for pre- and post-2008 was the alternating design (55%). In Figure 13 the frequency of designs is expressed before and after Burns and Wagner's (2008) meta-analysis.

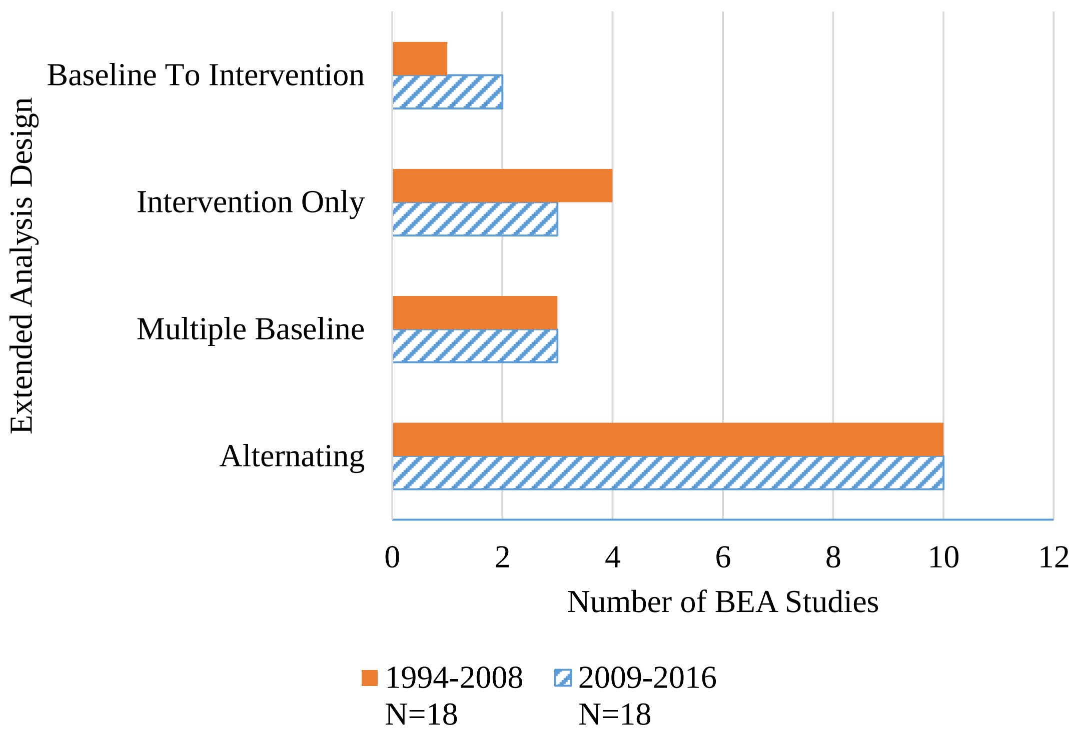

Figure 13. Design of extended analyses. Designs were categorized into Baseline to Intervention = baseline is established and followed by implementing the BEA selected intervention; Intervention Only = only the intervention was implemented; Multiple Baseline = intervention(s) implemented with a return to baseline; Alternating = conditions were alternating between intervention and baseline.

Number of interventions implemented. The number of interventions per extended analysis fell into five categories including one intervention up to four interventions and variable number. The extended analysis that included a variable number of interventions (i.e., two to three) adjusted according to the students' response with or without contingent reward. Figure 14 shows how the distribution of interventions among extended analyses from 1994-2008 and 2009-2016. The extended analyses from 1994-2008 were split between one and two interventions (50/50), with those from 2009-2016 not equivalent with one intervention used by 38% and two used by 53%. Only the post 2009 group evidenced interventions of three or more (17%).

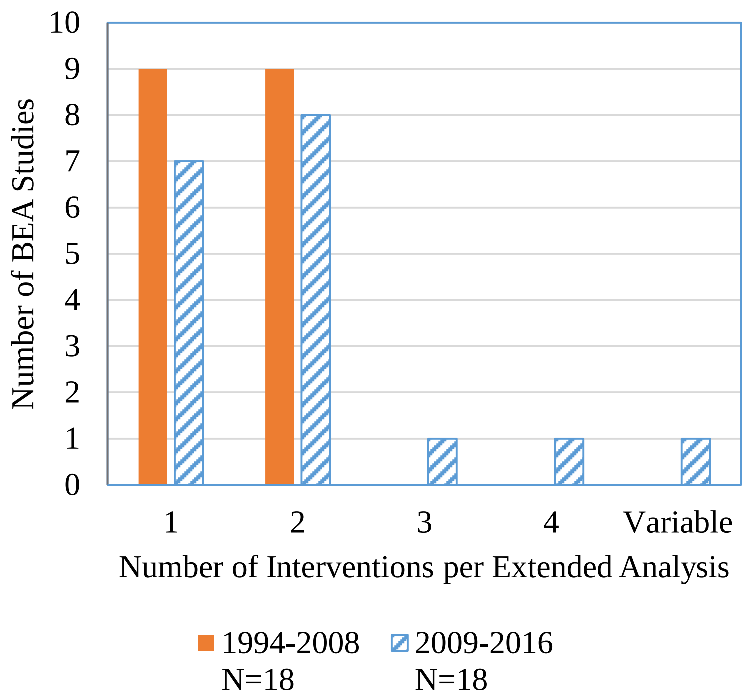

Figure 14. Number of interventions utilized in each extended analysis. Variable = number of interventions dependent on the students' performance.

Procedure for selecting interventions for extended analysis. The data were grouped into six categories that covered the methods for selecting the intervention(s) to be tested in the extended analysis. When the BEA identified only one intervention for use in the extended analysis it was categorized as selected intervention. Another category was created that included all extended analyses that compared the top two most effective interventions from the BEA which is labeled top two compared. Another category identified was when the intervention that produced the best results is compared to the intervention producing the least or poorest result (Most versus Least Effective). Another category labeled all interventions included studies where all interventions were again implemented for the extended analysis. The last category was student versus BEA selected interventions where a student selected a preferred intervention from the interventions to be implemented prior to the BEA and the extended analysis compared the student selected intervention to the highest performing one from the BEA. Selected intervention was used most frequent both pre- (61%) and post-2008 (56%). Figure 15 presents how the methods for selecting interventions chosen before and after Burns and Wagner's (2008) meta-analysis.

Length of extended analyses. The number of sessions from study to study varied dramatically. In the 1994-2008 studies there were 10 studies (56%) that conducted the same number of sessions across all students, and in the 2009-2016 group there were 9 studies (50%) that had consistent sessions. There was typically not an explicit rationale for varied lengths of extended analysis, but some of studies did mention that the number of sessions varied from student to student for miscellaneous reasons (e.g., scheduling problems, school breaks, no school because of snow). Extended analyses consisted of various session ranges within studies (e.g., 4-11, 9-15, 11-12, 10-19), with the purpose of finding patterns of stability and/or continued improvement with the implementation of the BEA selected intervention(s).

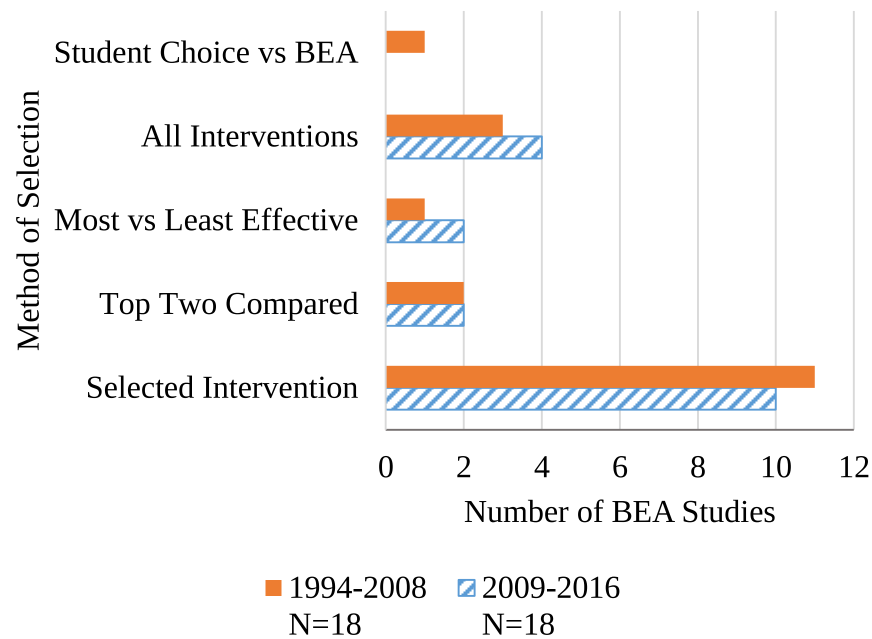

Figure 15. Frequency of methods to select the intervention(s) used in extended analyses. Student Choice vs BEA = student-selected interventions versus BEA identified interventions; All Interventions = all interventions from BEA utilized in extended analysis; Most vs Least Effective = most effective intervention selected; Top Two Compared = top two effective interventions were selected; Selected Intervention= The most effective intervention was selected.

Outcome of the extended analysis. Of the BEA studies that implemented extended analysis, a vast majority (94%) conducted it with the purpose of confirming the treatment validity of the BEA selected intervention(s). Of the two studies that did not use extended analysis in this manner, one was looking for the fewest number of components required to meet "mastery" level and the other study tested to find the best level of intervention intensity. Of those that did utilize extended analysis to evaluate BEA selected interventions, 17 (94%) of the studies from 1994-2008 and 11 (69%) of the studies from 2009-2016 confirmed that the BEA selected intervention was effective. Within the 1994-2008 group, one study observed the effect of the top two interventions in the extended analysis, finding mixed results with half of the BEAs not selecting the most effective intervention. The 2009-2016 group yielded four (25%) studies with mixed results due to inconsistent scores of the BEA selected intervention(s) when multiple interventions were compared in the extended analysis. There was also one extended analysis in a written expression study that was revealed to yield limited conclusions due to the variability in scores for the BEA selected intervention.

Research Question 4: What is the status of BEA in academic areas other than reading?

Of the 45 articles found, nine (20%) included BEA studies in written expression and math. Of these nine articles, five were in mathematics, three were in written expression, and one article contained both academic areas. Two articles contained two BEA studies each, one in math and one that included both math and writing. Therefore, there was a total of seven mathematics and four written expression BEA studies. The academic areas of mathematics and written expression were analyzed in terms of conceptual model utilized, methods of implementation, analysis of results, and design of the studies. One of the mathematics studies was published prior to Burns and Wagner's (2008) meta-analysis, but the rest of the mathematic and written expression studies were published after 2008.

Mathematics. The seven mathematical studies were found across four different journals (i.e., Journal of Behavioral Education, Journal of Applied Psychology, Psychology in the Schools, and School Psychology Review). Of the seven mathematics studies, there were three conceptual models utilized (i.e., instructional hierarchy, skill versus performance, and the problem-solving model), two sequences of intervention implementation (i.e., easy to hard and random), two methods of analyzing the results (i.e., visual analysis and percentage over baseline), as well as four designs (i.e., alternating treatments, mini-reversal, multiple-baseline, and AB design). Comparison between studies yields little consistencies due to the low number of studies as well as the variety within the components (with the exception of the two studies within the same article). There was commonality in the fact that all of the studies used randomly generated probes due to measuring a specific skill within mathematics (e.g., 2 digit-by-2 digit multiplication without regrouping or 3 digit-by-3 digit multiplication with regrouping). Baseline data across these studies varied from one to four points of data, with one study using at least three points or until a stable pattern was established. There were two to four interventions implemented in each study, with all of the interventions being single instructional contingencies. Each math intervention was implemented one to three times, with five of the studies repeating the interventions two to three times dependent on student performance. The dependent measures consisted of general outcome measures and subskills using fluency, accuracy, and total scores on probes. All of the studies in mathematics did employ both the BEA and an extended analysis with the purpose of identifying effective interventions and then confirming validation over an extended period of time. All of the extended analyses within these studies validated the BEA selected intervention.

Written expression. The written expression studies were found across three journals (i.e., Journal of Behavioral Education, Journal of Applied School Psychology, and School Psychology Review). The studies covered the areas of written letter formation, spelling, early writing, and general written expression. Of the four written expression studies, there were three conceptual models utilized (i.e., the five hypotheses for academic failure, skill versus performance deficit, and the problem-solving model), three sequences of intervention implementation (i.e., easy to hard, standard order, and randomized), two methods of analyzing the results (i.e., visual analysis and percentage over baseline), as well as three designs (i.e., alternating treatments, mini-reversal, and AB design). There was also no intervention that spanned across the four studies. Like mathematics, comparison between studies did not result in any notable consistencies across the variables of study due to the low number of studies as well as the variety of BEA components. Baseline data across these studies consisted primarily of implementing one baseline point, with the exception of one study that used prescreening information to determine student level of performance. There were two to four interventions implemented in each BEA study, with three of the studies only measuring individual instructional consistencies and one measuring three instructional contingencies and one instructional package. Each intervention was implemented one to two times, with two of the studies repeating the interventions dependent on student performance. The dependent measures consisted of general outcome measures and subskills using fluency, accuracy, and total scores. However, all of the studies did employ both the BEA and an extended analysis with the purpose of identifying effective interventions and then confirming validation over an extended period of time. Three of the studies confirmed that the BEA selected intervention was effective or the most successful (when more than one BEA intervention was used). One study made limited conclusions from the extended analysis due to high variability in the intervention performance over time.

Discussion

Brief experimental analysis (BEA) has been present in the literature for the past 23 years. Operationally defined in 1997 (Daly et al.), BEA allows for conceptually-based interventions to be rapidly compared in order to identify effective interventions for individual students. Burns and Wagner (2008) conducted a meta-analysis that consisted of BEA studies analyzing ORF. From this analysis they made recommendations for future practice of BEA. The current study examined the BEA literature in order to analyze the methods and strategies of conducting a BEA before and after Burns and Wagner's (2008) meta-analysis.

The search for BEA articles via WKU's Online Library databases yielded a total of 45 relevant articles. The 45 articles come from 16 different journals crossing the areas of school psychology, applied behavior analysis, education, and learning disabilities. Post-2008 there was a slight shift noted in more publications in school psychology journals than in behavioral journals. Also, the rate of publications on the topic of BEA has increased post-2008. This change indicates a slight shift in the disciplines conducting and the consumers reading the research on BEA. The rate increase indicates BEA is a slightly more frequent area of study among researchers.

The first research question addressed involved analyzing the methods of determining which interventions the BEA identified as "successful." The present study yielded some shifts from pre-2008 to post-2008 studies. The majority of the studies both pre- and post-2008 used visual analysis, however, the proportion of BEA studies using visual analysis as a means of determining the most effective intervention was slightly lower in the post-2008 studies as opposed to the pre-2008 group. PND was proportionately the highest used method for the post-2008 studies when visual analysis was not utilized. Use of effect size as recommended by Burns and Wagner (2008) was not evident pre- or post-2008. This could be due to the fact that effect sizes are not being employed by practitioners as a common practice. PND is a much more realistic tool of measurement, due to its ease of computation and intuitive nature. Only one study from each group utilized a target WCPM score. Of the seven studies that used percentage over baseline criteria, none set the criterion for 73% over the median baseline as Burns and Wagner recommended. Also, none of the studies that used PND mentioned that Burns and Wagner's recommended 80% was required for an intervention, but many used 70% as a minimum requirement. Only one study used Burns and Wagner's (2008) criteria in the area of ORF for an increase of 30 WCPM (Fienup, Reyes-Giordano, Wolosik, Aghjayan, & Chacko, 2015). All other studies used criteria that were less than that of Burns and Wagner's (2008) for percentage gain and PND. One explanation for this is that extended analysis was used in the majority of these studies both pre- and post-2008. A less stringent criterion can be supported if extended analysis serves to verify the effects of selected interventions.

The second research question sought to obtain information about any changes in the specific details of the BEAs. The span of grade levels of the participants indicated that a majority of the studies pre- and post-2008 only implemented BEAs in grades one through three. While CBM allows for testing across grades and adjusting to individual levels of performance, using the same criteria to select interventions for a student in 1st grade may yield different results for students in 6th grade. However, both groups of studies primarily focused on 3rd grade students. The focus on the lower grade levels is understandable in that development of literacy skills by third grade predicts level of academic attainment in future grades (Stanovich, 1986).

The use of conceptual models to shape the selection of interventions for the BEA was more evident post-2008, as every study included some rationale for the selection of interventions. The problem solving model post-2008 was the most frequently used conceptual model (50%), which could partially be due to the establishment of evidence-based interventions and the availability of progress monitoring data emerging with the implementation of response to intervention in the schools. In other words, the evidence-based interventions were chosen because the interventions were previously shown to address the problem identified through progress monitoring data. Also, the post-2008 studies were proportionally providing more explicit reference to models used than the pre-2008 studies.

In the selection of single-subject designs, the vast majority of both groups used alternating treatments in the BEA studies. This indicates some consistency in selection of single subject design. Also, mini reversals were employed more frequently post-2008. Even though mini-reversals were not predominantly used in the designs, many studies implemented an extended analysis that included some sort of control measure that allowed for the highest performing interventions to be compared directly to baseline. Whether it is done during the BEA or extended analysis, a return to baseline is significant in that treatment effects from different interventions back to back can impact a student's score (which is especially true if more than one session is conducted in the same day). Implementing the best performing intervention again directly after baseline gives a better representation of its direct effect on the student's performance.

Within each BEA, two to four interventions were compared in the majority of studies both pre- and post-2008 studies. Of the studies that compared instructional packages, the pre-2008 group included more packages per study. The shift to using less packages post-2008 could be due to a greater focus post-2008 on the impact of single instructional contingencies. Treatment effects of combining multiple components into a package make it difficult to know which specific instructional contingencies are responsible for the results, especially if they are administered in the same session. Also, if one instructional contingency out of a comprehensive package is responsible for a vast majority of the improvement, then it would be more practical to implement the single instructional contingency over the whole package.

More studies utilized the easy-to-hard sequencing of interventions in the pre-2008 studies, which could be due to a higher proportion of researchers implementing instructional packages of increasing complexity in a specific sequence. While instructional packages are more comprehensive interventions, knowing the effect of singular instructional contingencies allows for precision in determining what has an effect on the child's performance. A majority of the post-2008 studies likely focused on implementing one round of single instructional contingencies (65%) considering a majority of interventions were administered one time each within the BEA (55%). Standardizing the number of sessions in BEA studies is more manageable for practitioners, but there is a balance between frequency of implementation and differential results that are of use in drawing conclusions for research and practice.

Baseline data provided by the studies revealed that a majority of studies pre- and post-2008 utilized one data point for baseline. Some studies alternated baseline within the

BEA design, but ultimately the data collected before interventions were implemented were used to establish the student's level of performance. The use of one data point is practical and consistent with the intent of the purpose of BEA.

A vast majority of the CBM probes used in studies were from basal series pre-2008, whereas the post-2008 studies consisted of more custom probes created by the researchers. This is largely due to the much larger proportion of math and written expression studies that were conducted post-2008. The analysis of dependent variables reveals that, while both pre- and post-2008 studies predominantly measured fluency, post-2008 studies consisted of more variety in what was being measured. Again, this is partly due to the academic areas of mathematics and written expression being investigated among the post-2008 studies. There were also notably more reading-focused BEAs that looked at other factors outside of reading fluency (e.g., comprehension). Also, there were more studies that looked at subskills (e.g., correct letter sounds) in conjunction with general outcome measures (e.g., WCPM), which is also represented more in the post-2008 studies. However, some form of CBM is the preferred measurement tool.

In summary, post-2008 there have been noticeable proportional shifts towards utilizing methods of identifying the most effective interventions other than visual analysis, reversal single-subject designs utilized, comparing individual instructional contingencies over instructional packages, use of randomized intervention sequence, implementing each intervention one time in the BEA, and including more CBM packages and custom probes to measure progress. It would seem that some of these trends are consistent with current trends in the schools including progress monitoring and RTI.

The third research question sought to obtain information about the means of conducting extended analysis following the administration of BEA. The designs of the extended analyses were primarily alternating between intervention and baseline for both pre- and post-2008 studies. The pre- and post- 2008 groups had proportionately used similar designs, however, there were proportionally more interventions tested in the analyses post-2008. Some of the extended analyses post-2008 analyzed all of the interventions in the BEA in order to collect information on treatment validity in the BEA selected intervention(s). Therefore, more extended analyses were geared towards looking at the utility of extended analysis in relation to BEA selected interventions post-2008. Of the procedures used for extended analyses, the predominant method of analyzing the BEA-selected intervention(s) was to solely observe the effect of the intervention with or without baseline.

As noted in Burns and Wagner's (2008) meta-analysis, the length of the extended analysis can vary dramatically from study to study. Most researchers focus more on collecting data on stability and/or trend in improvement than creating a standardized procedure for practitioners. However, the vast majority of studies did implement extended analysis with the purpose of confirming the treatment validity of the BEA selected intervention(s). Any set criteria used to assess extended analysis was dependent on either reaching mastery level or reaching a specific score with no confinements to a specific increase over time. Also, it was noted that only two studies (Baranek et al., 2011; Welsch, 2007) analyzed differentiation between interventions during the extended analysis. The findings from both studies revealed that there was great variability between the two most successful BEA selected interventions in the extended analysis. The recommendation from the researchers was for future analysis of intervention differentiation to also include a measurement of variability and trend. While the limited analyses of intervention differentiation do not lead to any generalizable conclusions, they do give potential methods for future design.

The fourth research question examined the BEAs conducted in mathematics and written expression, which consisted of 23% of the total studies analyzed. While the overall analysis of BEA yielded many findings on shifting patterns in methods from the pre-2008 studies to the post-2008 studies, the specific analysis of BEA in mathematics and written expression was mainly for the post-2008 group. This analysis revealed that the studies contained many variations within the academic areas. However, there were some notable outcomes for each academic area. Across both areas, for example, an extended analysis was done with a majority indicating that the BEA selected interventions were effective.

In math, all of the interventions tested were individual instructional contingencies, with no two methods being packaged together. The math BEA studies shared commonality in the fact that all of the studies used randomly generated probes due to measuring a specific skill within mathematics. The dependent measures consisted of general outcome and sub-skill fluency, accuracy measures, and total scores on probes.

The BEA studies that looked at written expression covered the areas of written letter formation, spelling, early writing, and general written expression. The difference in specific subject areas resulted in no intervention being used across all studies. Like mathematics, comparison between studies did not result in any notable consistencies in most of the areas assessed in prior research questions. However, like math, the dependent measures consisted of general outcome and sub-skills fluency, accuracy measures, and total scores. All of the studies did employ both the BEA and an extended analysis with the purpose of identifying effective interventions and then confirming validation over an extended period of time. Three of the studies confirmed that the BEA selected intervention was the most successful, and one was inconclusive due to high variability in the intervention performance over time.

Strengths and Limitations

A strength of the current study is that it provides a unique approach at evaluating the various aspects of BEA over time and analyzes the various components evident in the published literature. The information provided can both give insight into the trends of each component over time as well as serve as a resource to inform future research of the areas which to conduct more individualized analysis. The various components of the study also give a broad perspective of how BEA is being conducted across reading, math, and written expression.

Another strength of the study is that many consistencies were identified across the pre- and post-2008 studies. Conceptual models have shown to be consistently utilized methods to rationalize the choice of interventions. Alternating treatments and mini-reversals were the most frequent single-subject designs used. Most BEAs were followed up with an extended analysis to verify the results. Most BEAs focus on one to two instructional levels. Also, many of the studies utilize some form of CBM. These findings show proportionally consistent patterns that resonate through the majority of BEAs conducted over the years. Overall, the present study gives insight into how the application of applied behavioral analysis techniques can effectively be used to identify relevant and practical instructional contingencies for a child's specific needs.

A limitation of this study is that while the general purpose of BEA is to test multiple interventions on a particular student, the BEA studies were drawn from research that was sometimes focused on one particular aspect of BEA. For example, early studies tended to focus on the usefulness of BEA in identifying interventions. Later studies tended looked at variables that influence outcomes, such as packages of interventions, the sequence of delivery of the interventions, and the efficacy of BEA with special populations and in content areas outside of reading. Underlying trends in the studies may be responsible for some of the shifts from pre- to post-2008 studies. Understanding the researcher's goals in conjunction with the assessment of interventions could provide rationale for specific shifts in the data.

Another possible limitation for this study was in how the articles were chosen to be included in this study based on keyword searches and a review of citations. Even though the articles were carefully reviewed for inclusion in the study, it is possible that some articles were not found due to errors in the database search, or due to authors using alternative terminology outside of the search terms. Also, there could have been something miscoded or misinterpreted in the article reviewed.

Future Research

One potential direction for future research would be to take the information from this study and construct a prototypical model or set of procedures for conducting a BEA. An understanding of the common trends and consistencies across studies gives insight into procedures needed to develop BEAs that can then be disseminated for widespread use in the schools.

Future research could also compare the various aspects of BEA analyzed in the current study in a different manner in order to inform best practice. For example, the current study analyzed BEA and extended analysis components individually, however a future study that looks into how specific strategies and methods employed could directly affect other components (e.g., how intervention frequency might impact the extent of data needed to validate the BEA-identified intervention). Developing a standard procedure for extended analysis would be of great utility for practitioners to efficiently indicate treatment validity. Through more specific comparisons of the interactions between BEA components, researcher might be able to find more commonalities and patterns that lead to successful outcomes for students.

Summary

Brief experimental analysis (BEA) in academics has shown to be an effective method of rapidly testing the relative effects of two or more interventions in order to provide evidence-based intervention that informs instruction. The findings of this study indicate that there have been shifts in the procedures of BEA since the publication of Burns and Wagner's (2008) meta-analysis with recommendations for practice. There are also consistencies in the methodology that span over the 23 years that this study analyzed. BEA has shown to still be a viable option as a means to inform instruction and intervention, with trends in research seeming to emphasize the practical application of BEA over analyzing its utility. Also, the BEAs covering the academic areas of math and written expression contained varied methodologies across different subskills that could provide the groundwork for researchers to further develop BEA for those areas. The content areas of math and written expression used subskill measures along with general outcome measures, whereas reading typically utilized general outcome measures only. Because of this trend, BEA procedures may need to be different for mathematics and written expression.

References

Banda, D. R., & Therrien, W. J. (2008). A teacher's guide to meta-analysis. *Teaching Exceptional Children, 41*(2), 66-71.

Baranek, A., Fienup, D. M., & Pace, G. (2011). Brief experimental analysis of sight word interventions: A comparison of acquisition and maintenance of detected interventions. *Behavior Modification*, *35*(1), 78–94. http://doi.org/10.1177/0145445510391242

Burns, M. K., & Wagner, D. (2008). Determining an effective intervention within a brief experimental analysis for reading: A meta-analytic review. *School Psychology Review*, *37*, 126–136.

Burns, M., Ganuza, Z., & London, R. (2009). Brief experimental analysis of written letter formation: Single-case demonstration. *Journal of Behavioral Education*, *18*(1), 20-34.

Busk, P. L., & Serlin, R. C. (1992). Meta-analysis for single-case research. In T. R. Kratochwill & J. R. Levin (Eds.) *Single-case research design and analysis: New directions for psychology and education* (pp. 187–212). Hillsdale, NJ: Lawrence Erlbaum.

Carson, P. M., & Eckert, T. L. (2003). An experimental analysis of mathematics instructional components: Examining the effects of student-selected versus empirically-selected interventions. *Journal of Behavioral Education*, *12*(1), 35-54.

Codding, R. S., Baglici, S., Gottesman, D., Johnson, M., Kert, A. S., & Lebeouf, P. (2009). Selecting intervention strategies: Using brief experimental analysis for mathematics problems. *Journal of Applied School Psychology, 25*(2), 146-168.

Daly, E. J., Andersen, M., Gortmaker, V., & Turner, A. (2006). Using experimental analysis to identify reading fluency interventions: Connecting the dots. *Behavior Analyst Today, 7*, 133-150.

Daly, E. J., Bonfiglio, C. M., Mattson, T., Persampieri, M., & Foremann-Yates, K. (2005). Refining the experimental analysis of academic skills deficits: Part I. An investigation of variables that affect generalized oral reading performance. *Journal of Applied Behavior Analysis, 38*, 485–497.

Daly, E. J., Martens, B. K., Dool, E. J., & Hintze, J. M. (1998). Using brief functional analysis to select interventions for oral reading. *Journal of Behavioral Education, 8*, 203–218.

Daly, E. J., Martens, B. K., Hamler, K. R., Dool, E. J., & Eckert, T. L. (1999). A brief experimental analysis for identifying instructional components needed to improve oral reading fluency. *Journal of Applied Behavior Analysis, 32*, 83–94. doi:10.1901/jaba.1999.32-83

Daly, E. J., Murdoch, A., Lillenstein, L., Webber, L., & Lentz, F. E. (2002). An examination of methods for testing treatments: Conducting brief experimental analyses of the effects of instructional components on oral reading fluency. *Education & Treatment of Children, 25*(3), 288-316.

Daly, E. J., Witt, J. C., Martens, B. K., & Dool, E. J. (1997). A model for conducting a functional analysis of academic performance problems. *School Psychology Review*, *26*, 554–574.

Deno, S. (2003). Developments in curriculum-based measurement. *The Journal of Special Education*, *37*, 184-192.

Duhon, G. J., Noell, G. H., Witt, J. C., Freeland, J. T., Dufrene, B. A., & Gilbertson, D. N. (2004). Identifying academic skill and performance deficits: The experimental analysis of brief assessments of academic skills. *School Psychology Review*, *33*(3), 429-443.

Eckert, T. L., Ardoin, S. P., Daisey, D. M., & Scarola, M. D. (2000). Empirically evaluating the effectiveness of reading interventions: The use of brief experimental analysis and single case designs. *Psychology in the Schools*, *37*(5), 463-473.

Eckert, T. L., Ardoin, S. P., Daly, E. I., & Martens, B. K. (2002). Improving oral reading fluency: A brief experimental analysis of combining an antecedent intervention with consequences. *Journal of Applied Behavior Analysis*, *35*(3), 271-281. doi:10.1901/jaba.2002.35-271

Fienup, D. M., Reyes-Giordano, K., Wolosik, K., Aghjayan, A., & Chacko, A. (2015). Brief experimental analysis of reading deficits for children with attention-deficit/hyperactivity disorder. *Behavior Modification*, *39*(1), 191–214. http://doi.org/10.1177/0145445514550393

Harding, J., Wacker, D. P., Cooper, L. J., Millard, T., & Jensen-Kovalan, P. (1994). Brief hierarchical assessment of potential treatment components with children in an outpatient clinic. *Journal of Applied Behavior Analysis, 27*, 291–300.

Haring, N. G., Lovitt, T. C., Eaton, M. D., & Hansen, C. L. (1978). *The fourth r: Research in the classroom*. Columbus, OH: Charles E. Merrill Publishing Co.

Hosp, J. L., & Ardoin, S. P. (2008). Assessment for instructional planning. *Assessment for Effective Intervention, 33*, 69-77.

Hosp, M. K., Hosp, J. L., & Howell, K. W. (2007). *The ABCs of CBM: A practical guide to curriculum-based measurement*. New York, NY: Guilford Press.

Individuals with Disabilities Education Improvement Act (IDEA). (2004). *Public Law No. 104-446.*

Jones, K. M., & Wickstrom, K. F. (2002). Done in sixty seconds: Further analysis of the brief assessment model for academic problems. *School Psychology Review*, *31*(4), 554-568.

Lentz, F. E., Jr. (1988). Effective reading interventions in the regular classroom. In J. L. Graden, J. E. Zins, & M. L. Curtis (Eds.), *Alternative educational delivery systems: Enhancing instructional options for all students* (pp. 351–370). Washington, DC: National Association of School Psychologists.

Loverich, T. P., O'Donohue, W. T., & O'Donohue, W. T. (2010). Functional analysis. In W. Craighead & C. Nemeroff (Eds.), *The Corsini encyclopedia of psychology and behavioral science*. Hoboken, NJ: Wiley. Retrieved from http://search.credoreference.com/content/entry/wileycorsini/functional_analysis/0

Martens, B., Eckert, T., Bradley, T., & Ardoin, S. (1999). Identifying effective treatments from a brief experimental analysis: Using single-case design elements to aid decision making. *School Psychology Quarterly, 14*(2), 163-181. doi:10.1037/h0089003

Martens, B., & Gertz, L. (2009). Brief experimental analysis: A decision tool for bridging the gap between research and practice. *Journal Of Behavioral Education*, *18*(1), 92-99.

McComas, J. J., Wacker, D. P., Cooper, L. J., Asmus, J. M., Richman, D., & Stoner, B. (1996). Brief experimental analysis of stimulus prompts for accurate responding on academic tasks in an outpatient clinic. *Journal of Applied Behavior Analysis, 29*, 397–401.

Noell, G. H., Gansle, K. A., Witt, J. C., Whitmarsh, E. L., Freeland, J. T., LaFleur, L. H., … Northup, J. (1998). Effects of contingent reward and instruction on oral reading performance at differing levels of passage difficulty. *Journal of Applied Behavior Analysis*, *31*(4), 659–663. http://doi.org/10.1901/jaba.1998.31-659

No Child Left Behind Act [NCLB]. (2001). *Public Law 107-15.*

Pearson, P. D., Johnson, D. D., Clymer, T., Indirsano, R., Venezky, R. L., Baumann, J. F., & … Toth, M. (1989). *Silver, Burdett, and Ginn*. Needham, MA: Silver, Burdett, and Ginn, Inc.

Petursdottir, A., McMaster, K., McComas, J. J., Bradfield, T., Braganza, V., Koch-McDonald, J., ... Scharf, H. (2009). Brief experimental analysis of early reading interventions. *Journal of School Psychology*, *47*(4), 215-243.

Plavnick, J., & Ferreri, S. (2013). Single-case experimental designs in educational research: A methodology for causal analyses in teaching and learning. *Educational Psychology Review*, *25*(4), 549–569.

Stanovich, K. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Research Quarterly, 22*, 360-407.

VanAuken, T. L., Chafouleas, S. M., Bradley, T. A., & Martens, B. K. (2002). Using brief experimental analysis to select oral reading interventions: An investigation of treatment utility. *Journal Of Behavioral Education*, *11*(3), 163-179. doi:10.1023/A:1020126003221

Welsch, R. G. (2007). Using experimental analysis to determine interventions for reading fluency and recalls of students with learning disabilities. *Learning Disability Quarterly*, *30*(2), 115–129.

Wilber, A., & Cushman, T. P. (2006). Selecting effective academic interventions: An example using brief experimental analysis for oral reading. *Psychology in the Schools*, *43*(1), 79-84.

Ysseldyke, J., & Reschly, D. (2014). The evolution of school psychology: Origins, contemporary status, and future directions. In P. Harrison & A. Thomas (Eds.), *Best Practices in School Psychology – Data Based and Collaborative Decision Making* (pp. 71- 86). Bethesda, MD: National Association of School Psychologists.

APPENDIX: ARTICLES INCLUDED IN THE CURRENT DATA ANALYSES

Andersen, M. N., Daly, E. J., III, & Young, N. D. (2013). Examination of a one-trial brief experimental analysis to identify reading fluency interventions. *Psychology in the Schools*, *50*(4), 403–414.

Baranek A, Fienup D. M., & Pace G. (2011). Brief experimental analysis of sight word interventions: A comparison of acquisition and maintenance of detected interventions. *Behavior Modification*, *35*(1), 78–94. http://doi.org/10.1177/0145445510391242

Bonfiglio, C. M., Daly, E. J., Martens, B. K., Lin, L.-H. R., & Corsaut, S. (2004). An experimental analysis of reading interventions: generalization across instructional strategies, time, and passages. *Journal of Applied Behavior Analysis*, *37*(1), 111–114. http://doi.org/10.1901/jaba.2004.37-111

Burns, M., Ganuza, Z., & London, R. (2009). Brief experimental analysis of written letter formation: Single-case demonstration. *Journal of Behavioral Education*, *18*(1), 20-34. Retrieved from http://www.jstor.org/stable/41824449

Carson, P. M., & Eckert, T. L. (2003). An experimental analysis of mathematics instructional components: Examining the effects of student-selected versus empirically-selected interventions. *Journal of Behavioral Education*, *12*(1), 35–54. http://doi.org/10.1023/A:1022370305486

Cates, G. L., Thomason, K., & Havey, M. (2007). A preliminary investigation of the effects of reading fluency interventions on comprehension: Using brief experimental analysis to select reading interventions. *Journal of Applied School Psychology*, *23*(1), 133-154.

Codding, R. S., Baglici, S., Gottesman, D., Johnson, M., Kert, A. S., & Lebeouf, P. (2009). Selecting intervention strategies: Using brief experimental analysis for mathematics problems. *Journal of Applied School Psychology*, *25*(2), 146–168. http://doi.org/10.1080/15377900802484661

Daly, E. J., Bonfiglio, C. M., Mattson, T., Persampieri, M., & Foreman-Yates, K. (2005). Refining the experimental analysis of academic skills deficits: Part I. An investigation of variables that affect generalized oral reading performance. *Journal of Applied Behavior Analysis*, *38*(4), 485–497. http://doi.org/10.1901/jaba.2005.113-04

Daly, E. J., Bonfiglio, C. M., Mattson, T., Persampieri, M., & Foreman-Yates, K. (2006). Refining the experimental analysis of academic skills deficits: Part II. Use of brief experimental analysis to evaluate reading fluency treatments. *Journal of Applied Behavior Analysis, 39*(3), 323–331. http://doi.org/10.1901/jaba.2006.13-05

Daly, E. J., & Martens, B. K. (1994). A comparison of three interventions for increasing oral reading performance: Application of the instructional hierarchy. *Journal of Applied Behavior Analysis*, *27*(3), 459–469. http://doi.org/10.1901/jaba.1994.27-459.

Daly, E., Martens, B., Dool, E., & Hintze, J. (1998). Using brief functional analysis to select interventions for oral reading. *Journal of Behavioral Education, 8*(2), 203-218.

Daly, E., Martens, B., Hamler, K., & Dool, E. (1999). A brief experimental analysis for identifying instructional components needed to improve oral reading fluency. *Journal of Applied Behavior Analysis*, *32*(1), 83–94. http://doi.org/10.1901/jaba.1999.32-83

Daly, E. J., Murdoch, A., Lillenstein, L., Webber, L., & Lentz, F. E. (2002). An examination of methods for testing treatments: Conducting brief experimental analyses of the effects of instructional components on oral reading fluency. *Education & Treatment of Children*, *25*(3), 288.

Daly E. J., Shroder J., & Robinson, A. (2001). Testing treatments for oral reading fluency problems: Two case studies. *Journal of Evidence-Based Practices for Schools, 7*, 4–26.

Dufrene, B. A., & Warzak, W. J. (2007). Brief experimental analysis of Spanish reading fluency: An exploratory evaluation. *Journal of Behavioral Education*, *16*(2), 143–154.

Duhon, G. J., Noell, G. H., Witt, J. C., Freeland, J. T., Dufrene, B. A., & Gilbertson, D. N. (2004). Identifying academic skill and performance deficits: The experimental analysis of brief assessments of academic skills. *School Psychology Review*, *33*(3), 429-443.

Eckert, T. L., Ardoin, S. P., Daisey, D. M., & Scarola, M. D. (2000). Empirically evaluating the effectiveness of reading interventions: The use of brief experimental analysis and single case designs. *Psychology in the Schools*, *37*(5), 463-473.

Eckert, T. L., Ardoin, S. P., Daly, E. J., & Martens, B. K. (2002). Improving oral reading fluency: A brief experimental analysis of combining an antecedent intervention with consequences. *Journal of Applied Behavior Analysis*, *35*(3), 271–281. http://doi.org/10.1901/jaba.2002.35-271

Eckert, T., Dunn, E., & Ardoin, S. (2006). The effects of alternate forms of performance feedback on elementary-aged students' oral reading fluency. *Journal of Behavioral Education*, *15*(3), 148–161.

Fienup, D. M., Reyes-Giordano, K., Wolosik, K., Aghjayan, A., & Chacko, A. (2015). Brief experimental analysis of reading deficits for children with attention-deficit/hyperactivity disorder. *Behavior Modification*, *39*(1), 191–214. http://doi.org/10.1177/0145445514550393

Gortmaker, V. J., Daly, E. J., McCurdy, M., Persampieri, M. J., & Hergenrader, M. (2007). Improving reading outcomes for children with learning disabilities: Using brief experimental analysis to develop parent-tutoring interventions. *Journal of Applied Behavior Analysis*, *40*(2), 203–221. http://doi.org/10.1901/jaba.2007.105-05

Guzel-Ozmen, R. (2011). Evaluating the effectiveness of combined reading interventions on improving oral reading fluency of students with reading disabilities. *Electronic Journal of Research in Educational Psychology*, *9*(3), 1063–1086. Retrieved from http://www.investigacion-psicopedagogica.org/revista/articulos/25/english/Art_25_606.pdf

Hofstadter-Duke, K. L., & Daly, E. J. (2011). Improving oral reading fluency with a peer-mediated intervention. *Journal of Applied Behavior Analysis*, *44*(3), 641–646. http://doi.org/10.1901/jaba.2011.44-641

Jones, K. M., & Wickstrom, K. F. (2002). Done in sixty seconds: Further analysis of the brief assessment model for academic problems. *School Psychology Review*, *31*(4), 554-568.

Jones, K. M., Wickstrom, K. F., Noltemeyer, A. L., Brown, S. M., Schuka, J. R., & Therrien, W. J. (2009). An experimental analysis of reading fluency. *Journal of Behavioral Education*, *18*(1), 35–55. http://doi.org/10.1007/s10864-009-9082-9

Kuhn, L. R., Watson, L., Ota, M., Cole, M., & Johnson-Gros, K. N. (2009). Effects of a brief experimental analysis and a reading intervention on multiple dependent variables. *Journal of Evidence-Based Practices for Schools*, *10*(1), 23–42.

Malloy, K. J., Gilbertson, D., & Maxfield, J. (2007). Use of brief experimental analysis for selecting reading interventions for English language learners. *School Psychology Review*, *36*(2), 291–310.

McComas, J., Wagner, D., Chaffin, M., Holton, E., McDonnell, M., & Monn, E. (2009). Prescriptive analysis: Further individualization of hypothesis testing in brief experimental analysis of reading fluency. *Journal of Behavioral Education, 18*(1), 56–70.

McCurdy, M., Clure, L. F., Bleck, A. A., & Schmitz, S. L. (2016). Identifying effective spelling interventions using a brief experimental analysis and extended analysis. *Journal of Applied School Psychology*, *32*(1), 46–65. http://doi.org/10.1080/15377903.2015.1121193

McCurdy, M., Daly, E., Gortmaker, V., Bonfiglio, C., & Persampieri, M. (2007). Use of brief instructional trials to identify small group reading strategies: A two experiment study. *Journal of Behavioral Education*, *16*(1), 7–26.

Mong, K. W., Mong, M. D., Henington, C., & Doggett, R. A. (2012). A Comparison of Four Reading Interventions from Struggling Elementary Students Using Brief Experimental Analysis and Extended Intervention Analysis. *Journal of Direct Instruction*, *12*, 13–24.

Mong, M. D., & Mong, K. W. (2012). The utility of brief experimental analysis and extended intervention analysis in selecting effective mathematics interventions. *Journal of Behavioral Education*, *21*(2), 99–118. http://doi.org/10.1007/s10864-011-9143-8

Noell, G. H., Freeland, J. T., Witt, J. C., & Gansle, K. A. (2001). Using brief assessments to identify effective interventions for individual students. *Journal of School Psychology*, *39*(4), 335–355. http://doi.org/10.1016/S0022-4405(01)00072-3

Noell, G. H., Gansle, K. A., Witt, J. C., Whitmarsh, E. L., Freeland, J. T., LaFleur, L. H., … Northup, J. (1998). Effects of contingent reward and instruction on oral reading performance at differing levels of passage difficulty. *Journal of Applied Behavior Analysis*, *31*(4), 659–663. http://doi.org/10.1901/jaba.1998.31-659

Parker, D., Dickey, B., Burns, M., & McMaster, K. (2012). An application of brief experimental analysis with early writing. *Journal of Behavioral Education*, *21*(4), 329–349.

Persampieri, M., Gortmaker, V., Daly III, E. J., Sheridan, S. M., & McCurdy, M. (2006). Promoting parent use of empirically supported reading interventions: Two experimental investigations of child outcomes. *Behavioral Interventions*, *21*(1), 31–57. http://doi.org/10.1002/bin.210

Petursdottir, A.-L., McMaster, K., McComas, J. J., Bradfield, T., Braganza, V., Koch-McDonald, J., … Scharf, H. (2009). Brief experimental analysis of early reading interventions. *Journal of School Psychology*, *47*(4), 215–243.

Powers, K., & Mandal, A. (2011). Tier III assessments, data-based decision making, and interventions. *Contemporary School Psychology, 15*, 21–33.

Reisener, C. D., Dufrene, B. A., Clark, C. R., Olmi, D. J., & Tingstrom, D. H. (2016). Selecting effective interventions to increase math computation fluency via brief experimental analyses. *Psychology in the Schools*, *53*(1), 39–57. http://doi.org/10.1002/pits.21887

Rich, H., E., S., & Duhon, G. J. (2014). Using brief academic assessments to determine generalization strategies. *Journal of Behavioral Education*, *23*(4), 401–420. http://doi.org/10.1007/s10864-014-9212-x

Schreder, S. J., Hupp, S. D. A., Everett, G. E., & Krohn, E. (2012). Targeting reading fluency through brief experimental analysis and parental intervention over the summer. *Journal of Applied School Psychology*, *28*(2), 200–220.

VanAuken, T. L., Chafouleas, S. M., Bradley, T. A., & Martens, B. K. (2002). Using brief experimental analysis to select oral reading interventions: An investigation of treatment utility. *Journal of Behavioral Education*, *11*(3), 163-179.

Wagner, D., McComas, J. J., Bollman, K., & Holton, E. (2006). The use of functional reading analysis to identify effective reading interventions. *Assessment for Effective Intervention*, *32*(1), 40–49. http://doi.org/10.1177/15345084060320010601

Welsch, R. G. (2007). Using experimental analysis to determine interventions for reading fluency and recalls of students with learning disabilities. *Learning Disability Quarterly*, *30*(2), 115–129.

Wilber, A., & Cushman, T. P. (2006). Selecting effective academic interventions: An example using brief experimental analysis for oral reading. *Psychology in the Schools*, *43*, 79–84.