Best Practices and Interventions in Special Education: How do we Know What Works?

Lucinda S. Spaulding

A Feature Article Published in

TEACHING Exceptional Children Plus

Volume 5, Issue 3, January 2009

Best Practices and Interventions in Special Education: How do we Know What Works?

Lucinda S. Spaulding

Abstract

The critical issue in special education today is no longer the assurance of access, but rather, the assurance of effectiveness. Determining which practices and interventions are most effective and efficient for ensuring optimal student achievement is a fundamental concern of special education teachers in this era of accountability. In this discussion I examine three designs commonly used in special education research (experimental research designs, meta-analyses, and narrative research syntheses) and their utility and appropriateness for determining the efficacy of classroom practices and interventions.

Keywords

best practices, research designs, experimental research designs, meta-analysis, narrative research syntheses

SUGGESTED CITATION:

Spaulding, L. S. (2009). Best practices and interventions in special education: How do we know what works? *TEACHING Exceptional Children Plus*, 5(3) Article 2. Retrieved [date] from http://escholarship.bc.edu/education/tecplus/vol5/iss3/art2

Introduction

While the paramount issue in special education 40 years ago was *access*, the critical issue today is *effectiveness* (Katsiyannis, Yell, & Bradley, 2001; Kavale, 2007; Keogh, 2007). Public Law 94-142 (1975) (now the Individuals with Disabilities Education Act [IDEA]) ensured students with disabilities were educated, but it did little to influence, regulate, or assess the effectiveness of services provided. As a result, although students with disabilities finally began receiving a public education, a gap developed between the academic achievement of those with disabilities and those without. Addressing and reducing this achievement gap was a key focus of the No Child Left Behind Act (NCLB, 2001). NCLB recognized that "ineffective teaching practices and unproven education theories are among the chief reasons children fall behind" (p. 1). Consequently, NCLB requires the use of scientifically based instructional programs and provides guidelines for evaluating if an intervention is supported by rigorous evidence (see Coalition for Evidence-Based Policy, 2003).

Some researchers question both the utility of relying solely on a single experimental design for evaluating the efficacy of a given intervention or program and the validity of generalizing classroom research to other settings.

Moreover, United States Federal regulations define *special education* as "specially designed individualized or group instruction or special services or programs . . . to meet the unique needs of students with disabilities" (Department of Education, 2006, p. 223). Hence, the fundamental challenge in special education is determining which instructional interventions, services, and programs most effectively and efficiently achieve this federal mandate of meeting the unique needs of students with disabilities, with the natural corollary of reducing the achievement gap.

Although NCLB emphasizes evidence-based practices and special education professionals have traditionally endorsed the scientific method for making decisions about the efficacy of services and interventions (Kavale, 2007), several paradigm wars divide the field (Forness, 2001), with the least being qualitative versus quantitative research (Hirsch, 2002), to the greatest being modernism versus postmodernism (Mostert, Kauffman, & Kavale, 2003). With such discord among researchers alongside the myriad of poorly designed and advocacy-driven studies permeating the field (Coalition for Evidence-Based Policy, 2003), it begs the question: Is there is any hope of objectively knowing what works and what does not work in special education?

The purpose of this discussion is to examine which research designs are more or less effective for empirically establishing best practices in special education, and to determine when it is appropriate to implement or rely on the following methods: experimental research designs, meta-analyses, and narrative research syntheses (see Table 1).

Experimental Research Designs

Many argue true experimental research designs yield the answers to special education's fundamental question, what works? There are several key characteristics of experimental research designs including random assignment, manipulation of the treatment conditions, outcome measures, and group comparisons (Cresswell, 2005). Ran-

dom assignment refers to the process of assigning participants at random to either a control group (having no exposure to the intervention) or an experimental group (receiving the intervention) in order to distribute participants and their personal characteristics evenly across groups. Experiments with random assignment are considered "true experiments" and are more rigorous than "quasi-experiments" which lack random assignment. Manipulation of treatment conditions in educational experiments typically involves introducing a treatment condition or independent variable (e.g., intervention, treatment, program) and measuring the results or dependent variable (e.g., academic achievement, improved behavior). Outcomes for the control and experimental group are measured to determine the effect of the treatment and to make group comparisons.

Table 1: Characteristics of Research Designs

Characteristics of Research Designs
Experimental Research
❑ Compare two (or more) groups: Group 1: No intervention Group 2: Receives an intervention (Group 3: Receives an alternative intervention) ❑ Participants are randomly assigned so groups are equal ❑ Often include pretests and posttests
Meta-analyses
❑ Include many experimental research studies on a topic ❑ Combine statistical/numerical results to determine the overall magnitude of results ❑ Used to determine the strength of an intervention or amount of difference between groups ❑ Used to refute or support general findings
Narrative Research Syntheses
❑ Include multiple kinds of studies on a topic (i.e., experimental, quasi-experimental, survey research, etc.) ❑ Serve to find patterns, trends, or themes in research ❑ Used to analyze the strengths and weaknesses of primary studies ❑ The purpose is to summarize and draw conclusions from multiple studies

According to the Coalition for Evidence-Based Policy (2003), true experimental research designs should be considered the benchmark for measuring the effects of an intervention. On this premise, the Coalition outlined the criterion (i.e., a control and an experimental group, random assignment, etc.) for evaluating whether or not interventions are backed by strong evidence.

An emphasis on experimental research is also reflected in the suggestions of special education researchers assembled by the Office of Special Education Programs (OSEP) (see Gersten, Baker, & Lloyd, 2000). Summarizing the guidelines developed by this group, Gersten et al. contended that experimental group designs are the most powerful method available for evaluating the effectiveness of

interventions, and "maintaining a focus on conducting intervention research *in real school settings* [italics added] is imperative" (p. 3).

However, some researchers question both the utility of relying solely on a single experimental design for evaluating the efficacy of a given intervention or program and the validity of generalizing classroom research to other settings. In his article *Classroom Research and Cargo Cults*, Hirsch (2002) asserts that educational research is generally inconclusive: "The process of generalizing directly from classroom research is inherently unreliable" (p. 53). Hirsch argues that most classroom studies are a-theoretical, lacking usefulness for advancing research agendas or directing policy. Hirsch claims, "the limitations of classroom research eliminate not only certainty, but also the very possibility of scientific consensus" (p. 54). His explanation is that because schooling is "context-dependent," there are simply too many extraneous variables (e.g., teacher quality, school culture, etc.) that cannot be adequately controlled in a classroom setting, thereby eliminating the opportunity to conclude that any specific independent variable (e.g., intervention, treatment, program) is responsible for a specific dependent variable (e.g., academic achievement, improved behavior). While Hirsch's solution is to place less reliance on traditional educational research, he concedes that synthesizing research on a certain topic is "a more dependable guide to education policy than the data derived from classrooms" (p. 59). He explains that theories can gain consensus when data from many kinds of studies and sources are explained. Hirsch concludes by challenging educational policy makers to demand consensus from the research community.

Narrative research syntheses serve as valuable research methods for integrating and synthesizing findings.

Hirsch does not stand alone in his conclusion. Research demonstrates that experimental treatments often produce unpredictable results, and the variability of effects is often greater than the average effectiveness of that treatment (Mostert, 2001a). Furthermore, although empirical evidence is available to determine whether methods for special education instruction are effective, the evidence too frequently remains isolated and irrelevant when the results of individual studies conflict (Kavale, 2007). Consequently, "a single study, no matter how elegant, is unlikely to provide a definitive evaluation" (Mostert & Kavale, 2001, p. 57). Hence, when an area in the field possesses a number of unresolved issues, quantitative review methods should be employed to "impart an objective, explicit, and systematic attitude to the review process" (Kavale & Forness, 1996, p. 228).

Recognizing the imperative to "converge on a consensus view," leading special education researchers emphasize the importance of synthesizing research (i.e., Forness, 2001; Kavale, 2007; Mostert, 1996; Swanson, 1996). While other methods of reviewing literature have been emphasized in the past, meta-analysis has increasingly become the preferred method for conducting rigorous reviews of special education research: "What the research says is most clearly revealed in rigorous narrative reviews, quantitative approaches in general, and *meta-analysis in particular* [italics added]" (Mostert & Kavale, 2001, p. 65).

Meta-analysis

In 1976 Gene Glass reintroduced meta-analysis as a method of quantitative research for assisting the process of combining research findings. Meta-analysis relies on the basic statistic of effect size (ES) and involves averaging ESs across a domain in order to determine either the level of differentiation between a group (e.g., students with disabilities versus students without), or the magnitude or strength of a treatment effect (e.g., the effectiveness of a particular intervention). ESs can be interpreted as *z* scores or standard deviation (SD) units. ESs range from 0 (no effect) to 1.00 (large effect), with an ES of 1.00 indicating that the two groups being compared differ by 1 SD, or, if using a standardized achievement test, an ES of 1.00 can be translated into one year of academic growth. By relying on the quantitative and objective parameter of ES, meta-analysis represents a decision-oriented form of evaluation that "transcends other forms of opinion, assertion, and belief" (Mostert & Kavale, 2001, p. 61).

Furthermore, meta-analysis follows the methodology of other primary research studies. Kavale (2001) explained that meta-analysis parallels the scientific method by incorporating the following procedures: formulating problems, sampling, classifying and coding research studies, data analysis, and ES interpretation. Moreover, in addition to determining the magnitude of an intervention or amount of differentiation among groups, meta-analysis provides a methodology for investigating main effects, interactions, and covariation (Kavale, 2001; Mostert, 1996). For these reasons, meta-analysis is considered by many to be the "gold standard" of research in special education. Mostert (2004) asserts there is little doubt that meta-analysis is a "powerful technique that provides very useful answers for theory, policy, and practice. In terms of uncovering meta-answers to questions of intervention efficacy, it continues to be useful for theorists and practitioners alike" (p. 114).

A good example of a significant educational meta-analysis is the National Reading Panel's meta-analysis of phonics instruction (see Ehri, Nunes, Stahl, & Willows, 2001). Commissioned in 1997 by the U.S. Congress, this quantitative research synthesis evaluated the effects of systematic phonics instruction compared to non-phonics instruction or unsystematic phonics instruction. Thirty-eight primary experimental research studies yielding 66 comparisons between treatment and control groups met the inclusion criteria for the study and generated the following results: The overall effect of phonics instruction on reading was moderate (ES = 0.41); effects were larger when instruction began early, and effects persisted after instruction ended; phonics benefited word reading, decoding, comprehension, and spelling; phonics helped low and middle SES readers, younger students at risk for reading disability (RD), and older students with RD; and systematic instruction of phonics was more effective for teaching students to read than all forms of control group instruction, including whole language.

However, although meta-analysis is an incredibly useful summative tool for answering major research questions in special education, it must be used wisely (Kavale, 2001; Mostert, 2004; Swanson, 1996). Several researchers demonstrated the need to strengthen the face validity of meta-analyses (Mostert, 1996; Swanson, 1996). Although the techniques of meta-analysis have "witnessed a number of technical advances that have served to enhance the objectivity, verifiability, and replicability of the meta-analytic review

process" (Kavale & Forness, 1996, p. 226-237), meta-analytic findings are not absolutely definitive or unimpeachable for several reasons (Mostert, 2004).

First, it must be acknowledged that a meta-analysis "can only be as valid as the expertise of the meta-analyst" (Mostert, 1996, p. 8). By its very nature, conducting a meta-analysis requires many critical decisions on the part of the researcher. Meta-analysts must: specify research questions and establish inclusion and exclusion criteria to discriminate among primary studies based on the research purpose(s); make decisions about coding study features in order to identify and separate independent variables in the study; decide how to calculate outcomes, for example, deciding among the Glassonian Meta-Analysis (entering multiple ESs from each primary study into the analysis without averaging) or using the Study Effect Meta-Analysis (averaging multiple effect sizes from a primary study to determine one average ES for the study); decide which ES statistic to use (e.g., dividing by the standard deviation [SD] of the control group or pretest SD, or the pooled SD); and finally, meta-analysts must determine the appropriate amount of detail to include in their discussion and analysis of findings.

Clearly, as NCLB posits, there are more objective ways of knowing what works, and therefore, there is hope of reducing the academic achievement gap between students with disabilities and those without.

Second, even the most competent and experienced meta-analyst is bound by the amount of information reported in the primary study: "Meta-analysis relies heavily on the information reported in the primary studies, which themselves may not be complete" (Mostert, 1996, p. 2). Moreover, ESs are often derived from studies of interventions with different purposes, research samples, and outcome measures (Forness, 2001). This is referred to as the "apples and oranges problem," the argument that diversity in primary studies makes comparisons inappropriate (Wolf, 1986). Jackson (1980) highlighted that although meta-analysis can be used for evaluating results within a set of studies on a given topic, "it cannot weave together the evidence across sets of studies on related topics" (p. 452). Other criticisms assert that meta-analytic results are uninterpretable because results from poorly designed studies are included with results from rigorous studies, and published research is biased because significant findings are more often published than insignificant findings, tending toward biased results (Wolf, 1986). Consequently, despite the best efforts of the researcher, the face validity of the meta-analysis may be limited.

Finally, meta-analytic results can be misleading; they tend to give the impression that their results are definitive (Forness, 2001; Mostert, 2001). However, Mostert (2001) explains that this impression may be challenged for three reasons:

> (a) Meta-analytic results rely heavily on how the independent variables from the primary studies are defined, related and coded, (b) the meta-analytic information provided is often too sparse for readers to make reasonable judgments regarding the face validity of the meta-analysis, and (c) some evidence suggests that meta-analyses conducted on the same body

of primary studies can yield different results. (p. 200)

For example, Hammill and Swanson (2006) provided an alternative interpretation of the National Reading Panel's meta-analysis of phonics instruction. Using a different form of analysis, Hammill and Swanson argued that the effects of phonics instruction are not moderate, but rather small: "In general, although effect sizes may favor phonics instruction, the magnitude of these differences on a practical level is in most cases small" (p. 25). In another example, a reanalysis by Inglis and Lawson (1987) of a Kavale and Forness (1984) study revealed opposite conclusions as a result of different statistical manipulations to the same set of data.

Further exemplifying the way results can be misleading or misinterpreted, Forness (2001) demonstrated the necessity of looking closely at data and interactions among variables. For example, a mega-analysis (a meta-analysis of meta-analyses) of special education and related services revealed an overall average special education intervention ES of 0.55. However, when dividing the interventions into three categories, (a) *special* education interventions (i.e., unique and different), (b) special *education* interventions (i.e., adapting and modifying instruction), and (b) related services (i.e., dependent on other professionals), analysis revealed an ES of 0.20 for *special* education, an ES of 0.84 for special *education*, and an ES of 0.53 for related services. It is clear that data must be carefully reported, analyzed, and interpreted to ensure findings are not errantly misleading.

However, to address criticisms and improve face validity, much attention has been directed toward developing criteria for evaluating the quality of published meta-analyses. Drawing from the growing literature addressing issues in meta-analyses, Mostert (1996, 2001a, 2004) methodologically outlined and illustrated (in learning disabilities, mental retardation, and emotional and behavioral disorders) a set of prototypical criteria for judging the quality of meta-analyses. Mostert's criteria spanned six domains: locating studies/context, specifying inclusion criteria, coding study features, calculating individual study outcomes, data analysis, and limits of the meta-analysis; and included (but were not limited) to the following criteria: greater accuracy and specificity of populations under study, descriptions of coded studies rather than lists, providing examples of included and excluded studies, and report the range of ESs.

Swanson (1996) also noted a deficiency in the literature related to available criteria for judging the quality of meta-analyses. Observing few replications, Swanson developed a checklist of suggested criteria for evaluating synthesis reports using meta-analysis. The major criteria categories included: qualification of effect sizes; criteria for the source (e.g., article) selection; basis for article inclusion; coding of variables; methodological rigor of studies; descriptive or statistical analysis; and interpretation and discussion related to the synthesis.

Since Mostert (1996) and Swanson (1996) proposed guidelines for better evaluation and replication of meta-analyses, recent reviews suggest that later meta-analyses in special education research "appear to be reporting more of the domain criteria than earlier studies, a significant improvement given the importance of reporting domain criteria for judging the face validity of published meta-analyses" (Mostert, 2001a, p. 218). Mostert (2004) observed a "fairly strong trend" (p. 114) in meta-analyses to increasingly report necessary information for judging the face validity and permitting replication.

Hence, when seeking answers to major research questions in special education, meta-analysis is clearly an appropriate evaluative method. However, meta-analyses are useful only if studies pertaining to a research question are predominantly *quantitative* (i.e., numerical/statistical) and use *experimental* designs. When the existing body of research on a topic contains a wide range of study designs (i.e., experimental, quasi-experimental, qualitative, and case study designs), meta-analysis is impossible and "analytical narrative synthesis may well be the only way of evaluating research to generate usable knowledge" (Mostert & Kavale, 2001, p. 57).

Narrative Research Syntheses

Conducting a meta-analysis is clearly not always possible because primary study results cannot always be transformed into ESs and many qualitative studies fail to use traditional research designs (Mostert & Kavale, 2001). In this case, narrative research syntheses serve as valuable research methods for integrating and synthesizing findings. Narrative reviews integrate various research studies on a topic by analyzing individual studies to draw an overall conclusion (Kavale, 2001). Narrative reviews can be arranged in one of four ways: (1) through identifying or discussing new developments in a field, (b) by illustrating, assessing, or proposing theory, (c) by organizing knowledge from divergent lines of research, or . . . (d) through integrative review methods (Mostert, 2001b).

When it comes to making decisions about implementing new practices and programs, the needs are too great and the time and resources too few to invest in interventions that have not been validated by multiple research studies.

Narrative research syntheses have many purposes and benefits. They serve to analyze the strengths and weaknesses of primary studies in detail, rather than refute or support general findings; attempt to make sense of divergent research findings around a similar research hypothesis; provide a summary of what is already known; allow researchers to uncover patterns and consistencies across studies; allow researchers to place more weight on studies using valid designs and reporting more complete data; and finally, narrative syntheses allow researchers to draw meta-conclusions (Mostert, 2001b).

Several highly informative and conclusive narrative reviews of research findings in special education have been conducted. One example illustrating the evaluative function of an analytical narrative synthesis is Mostert's (2001b) assessment of facilitated communication (FC) as a technique with autistic people and others who are noncommunicative. In his review, Mostert described, analyzed, and summarized primary study characteristics, followed by a summative discussion of findings supporting and opposing the efficacy of FC.

However, when data and research designs permit, *quantitative* research syntheses (i.e., meta-analyses) have largely replaced the reliance on narrative research because it is often difficult to objectively determine whether an intervention is better for certain types of children, more effective for certain types of problems, or has greater efficacy than other interventions (Forness, 2001). For these reasons, when tenable, the synthesis of cumulative research findings

in special education has generally begun to rely on meta-analysis.

It is imperative to note, however, that the importance and continued necessity of implementing true experimental research designs should not be mitigated. Both meta-analyses and narrative research syntheses rely on the availability of rigorous and sound primary research studies for synthesizing and drawing conclusions. When an intervention has not yet been evaluated or a body of literature on a topic is still developing, it is impossible to conduct a valid meta-analysis or comprehensive narrative review. In this case, implementing experimental research is often the only choice, and the efforts of NCLB, the Coalition for Evidence-Based Policy, and OSEP to establish criteria for implementing and evaluating experimental research are necessary and important.

However, when it comes to making decisions about implementing new practices and programs, the needs are too great and the time and resources too few to invest in interventions that have not been validated by multiple research studies in various sites and with diverse student populations. Educational decisions are best made based on conclusions from research designs such as meta-analysis that synthesize the results of *multiple* rigorous experimental research studies on a single topic or intervention, as they are more likely to produce accurate and less biased findings that can be translated into effective practice.

Conclusions

Is it possible to objectively know what works and what does not work in special education? The answer is "yes." Rigorous narrative research syntheses and meta-analytic syntheses "offer a methodology of enormous potential for judging the worth of special education practices because their relative objectivity brings greater logic and reason to judgments about what works" (Mostert & Kavale, 2001, p. 65). Clearly, as NCLB posits, there are more objective ways of knowing what works, and therefore, there is hope of reducing the academic achievement gap between students with disabilities and those without.

However, the centerpiece of special education is individualization (Yell, Rogers, & Rogers, 1998) and despite federal and state efforts to standardize teaching, the special education teacher is ultimately responsible for employing best practices and implementing interventions to build on students' present levels of performance in order to meet their social, emotional, behavioral, and intellectual needs. Mostert (1999-2000) demonstrates that teachers need discriminative ability—"the ability to know and understand what works effectively, what does not work effectively, and the ability to tell the difference" (p. 119). I would expand Mostert's argument and add that in addition to discriminating what works and what does not work, special educators must know their students' individual needs, and this comes directly from time with the student, effective forms of assessment, and accrued teacher experience. However, while the development of skills to discriminate and prioritize the specific needs of students may be gradual and take time, there is clearly a growing wealth of empirical evidence available *right now* for teachers to judge the efficacy of special education interventions.

As Hirsch (2002) aptly observed, "common sense will remain a valuable classroom commodity" (p. 67). While a wide-range of interventions and practices are being promoted by teacher education programs, local education agencies, and commercial organizations, there is also a large and growing special education research base available for evaluating the efficacy of many of these in-

terventions. It would seem highly sensible to rely on practices and interventions that have been demonstrated to be more effective than others. Answers are certainly available, most optimally from meta-analytic findings and narrative research syntheses; however, if they have not yet been conducted on a topic, there are clear guidelines for identifying rigorous experimental research studies to guide and inform teacher practices.

References

Coalition for Evidence-Based Policy (2003, December). Identifying and implementing educational practices supported by rigorous evidence: A user friendly guide. Washington, DC: Retrieved June 13, 2008, from http://www.ed/rschstat/pubs/rigorousevid/rigorous/evid.pdf

Cresswell, J. W. (2005). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research*. Upper Saddle River, NJ: Pearson Education.

Department of Education. (2006, August 14). Assistance to States for the education of children with disabilities and preschool grants for children with disabilities; Final rule. *Federal Register, 71*(156), 46540-46845.

Ehri, L.C., Nunes, S. R., Stahl, S. A., & Willows, D. M. (2001). Systematic phonics instruction helps students learn to read: Evidence from the National Reading Panel's meta-analysis. *Review of Educational Research, 71*(3), 393-447.

Forness, S. R. (2001). Special education and related services: What have we learned from meta-analysis? *Exceptionality, 9*(4), 185-197.

Gersten, R., Baker, S., & Lloyd, J. W. (2000). Designing high-quality research in special education: Group experimental design. *The Journal of Special Education, 34*(1), 2-18.

Glass, G. E. (1976). Primary, secondary, and meta-analysis of research. *Educational Researcher, 5*, 3-8.

Hammill, D. D., & Swanson, H. L. (2006). The National Reading Panel's meta-analysis of phonics instruction: Another point of view. *The Elementary School Journal, 107*(1), 17-26.

Hirsch, E. D., Jr. (2002). Classroom research and cargo cults. *Policy Review*, 51-69.

Inglis, J., & Lawson, J. S. (1987). Reanalysis of a meta-analysis of the validity of the Wechsler Scales in the diagnosis of learning disability. *Learning Disability Quarterly, 10*(3), 198-202.

Jackson, G. B. (1980). Methods for integrative reviews. *Review of Educational Research, 50*, 3, 438-460.

Katsiyannis, A., Yell, M. L., & Bradley, R. (2001). Reflections on the 25th anniversary of the Individuals with Disabilities Education Act. *Remedial and Special Education, 22*(6), 324-334.

Kavale, K. A. (2001). Decision making in special education: The function of meta-analysis. *Exceptionality, 9*(4), 245-268.

Kavale, K. A. (2007). Quantitative research synthesis: Meta-analysis of research on meeting special educational needs. In L. Florian (Ed.), *The Sage Handbook of Special Education* (pp. 207–221). Thousand Oaks, CA: Sage Publications.

Kavale, K. A., & Forness, S. R. (1996). Social skill deficits and learning disabilities: A meta-analysis. *Journal of Learning Disabilities, 29*(3), 226-237.

Keogh, B. K. (2007). Celebrating PL 94-142: The Education of All Handicapped Children Act of 1975. *Issues in Teacher Education, 16*(2), 65-69.

Mostert, M. P. (1996). Reporting meta-analysis in learning disabilities. *Learning Disabilities Research & Practice, 11*(1), 2-14.

Mostert, M. P. (1999-2000). A partial etiology and sequelae of discriminative disability: Bandwagons and beliefs. *Exceptionality*, 8(2), 117-132.

Mostert, M. P. (2001a). Characteristics of meta-analysis reported in mental retardation, learning disabilities, and emotional and behavioral disorders. *Exceptionality, 9*(4), 199-225.

Mostert, M. P. (2001b). Facilitated communication since 1995: A review of published studies. *Journal of Autism and Developmental Disorders, 31*(3), 287-313.

Mostert, M. P. (2004). Face validity of meta-analysis in emotional and behavioral disorders. *Behavioral Disorders, 29*(2), 89-118.

Mostert, M. P., Kauffman, J. M., & Kavale, K. A. (2003). Truth and consequences. *Behavioral Disorders, 28*, 333-347.

Mostert, M. P., & Kavale, K. A., (2001). Evaluation of research for usable knowledge in behavioral disorders: Ignoring the irrelevant, considering the germane. *Behavioral Disorders, 27*(1), 53-68.

No Child Left Behind. (2001). The facts about. . . investigating in what works. Retrieved June 15, 2008, from http://www.ed.gov/nclb/methods/whatworks/what_works.pdf

Swanson, H. L. (1996). Meta-analysis, replication, social skills, and learning disabilities. *The Journal of Special Education, 30*(2), 213-221.

Wolf, F. M. (1986). *Meta-analysis: Quantitative methods for research synthesis*. Series: Quantitative Applications in the Social Sciences. Beverly Hills, CA: Sage Publications.

Yell, M. L., Rogers, D., & Rogers, E. L. (1998). The history of special education: What a long, strange trip it's been! *Remedial and Special Education, 19*(4), 219-228.

About the Author:

Lucinda S. Spaulding is a Doctoral Fellow and Special Education/ Educational Psychology Ph.D. candidate at Regent University in Virginia Beach, VA.