

Evidentiary Sleight of Hand: The High Stakes of Silencing Teachers

Deborah Duncan-Owens

June 4, 2008

Evidentiary Sleight of Hand: The High Stakes of Silencing Teachers

Few would dispute the far reaching effects of the No Child Left Behind (NCLB) legislation on education. By far, however, its effects have been experienced most

profoundly by everyone involved in literacy instruction – from students, teachers, and administrators, to researchers and teacher trainers. The National Reading Panel’s (NRP, National Institute of Health, 2001) publication of their meta-analysis of reading methodology brought sweeping changes in how reading is taught in public schools, forming the basis for Reading First, the component of NCLB that governs reading instruction. The NRP established a standard for evaluating the effectiveness of reading instruction, mandating that instructional methodologies be grounded in scientifically based reading research. Likewise, the NRP’s decision to exclude qualitative studies from their meta-analysis had the effect of canonizing empirical quantitative evidence as the only acceptable means of evaluating instructional programs and methodologies. The term scientifically based reading research, used over 110 times in the Reading First portion of NCLB, has become the mantra of educational policy makers (Pearson, 2005). The implication that qualitative research is a less valid means for evaluating instructional practice has had a profound impact on reading educators and researchers by silencing the voices of the teacher practitioners on the front lines of literacy instruction (Garan, 2004). The disequilibrium that has marked the state of reading education since the publication of the NRP report and the signing of the NCLB legislation has left a void between theory and practice and a disconnect between higher education and public schools and districts. While schools clamber to align their curriculums with state and federal accountability standards, higher education has often been slow to respond with assistance. Commercial program developers have exploited the disconnect and rushed in to fill the void with expensive products that guarantee alignment with NCLB and state frameworks, promising impressive improvements in student achievement. This paper examines

commercial program developers' incursion into the educational arena and the evidentiary sleight of hand techniques practiced by program developers as teachers are marginalized and silenced and as program developers assume the dual roles of program promoter and evaluator or their own products.

Evidentiary Sleight of Hand Technique #1: Misdirection

Magicians often define misdirection as “getting the audience to look in the wrong place, at the right time” (Robinson, 2004, p. 1). While the audience's attention is focused elsewhere, the magician can pull a rabbit out of a hat or a card can magically appear out of nowhere. The key to the success of misdirection is the conditioning of the audience to focus their attention on the waving wand or card flourish. By the time the NRP issued its report the public was accustomed to the reading wars and bantering between phonics and whole language advocates. The NRP's initial charge was to settle the reading wars; however, far from putting an end to the reading wars, the NRP ratcheted up and rearmed the debate, extending the battle front to include arguments over NCLB, the NRP, and research methodologies. The stage was set for commercial reading program developers to promote their programs as the perfect tool for aligning school curricula with NRP standards and NCLB mandates.

At times the ratcheting up of the reading wars created a hostile environment marked by inflammatory statements and claims. Phonics proponent, Louise Moats (2007), for example, asserted:

The failures of whole language are many – from failure to teach phonics and other language skills explicitly and systematically, to an overly personalized, nondirective approach to reading comprehension. For millions of children who struggle to learn to read, the results are disastrous. (p. 4)

Whole language proponents often vilify phonics advocates and assign them ulterior motives such as attempting to gain absolute control over teachers and threatening the literacy of American youths (Hempenstall, 2008). Evidence of the degree of the rancor that has developed since the publication of the NRP report is the publication by the International Reading Association (IRA) of a Statement Calling for Civil Dialogue in which the IRA Board stated,

A civil, courteous, and professional public dialogue in support of strengthening education and improving student achievement is essential. The use of abusive language and violent metaphors by U.S. government officials and by education professionals alike is unacceptable, unprofessional, and unproductive. A civil dialogue is needed, and it must be the basis for a better understanding of how to provide excellent reading instruction that leads to high achievement of all students and communities we serve. (2003, p. 1)

The rearming of the reading wars has extended beyond the realm of reading instruction, too, as members of each reading camp lay claims to research methodologies. Whole language's alignment with qualitative research is not new. Frank Smith (1989) claimed, "Only one kind of research has anything useful to say about literacy, and that is ethnographic or naturalistic research" (p. 356). Proponents of phonics instruction stake their claim of quantitative research as proof positive of the superiority of their approach to teaching reading and the NRP report solidified their resolve. The exclusion of qualitative research from the NRP's analysis bolstered phonics proponents argument about, not only what constitutes effective reading instruction, but also which methodologies are acceptable for evaluating instructional efficacy. In all wars there are profiteers who have a vested interest in keeping the wars armed. In the case of the reading wars, commercial program developers are content to have the attention of reading

professionals focused on arguments about methodology. This misdirection of public and professional attention on the reading wars allows them to continue promoting the use of their products in schools seeking alignment with state and federal accountability mandates.

At the heart of the matter, for researchers, is what constitutes evidence of effective reading instruction and, more importantly, what constitutes evidence of authentic literacy learning. Reliance solely on quantitative evidence provides a limited understanding of how children learn to read. Most reading researchers seek a deeper understanding of literacy learning. Qualitative research has the ability to complete the portrait of a young literacy learner and develops an understanding of factors that impact literacy development, such as socioeconomic status, race, and ethnicity (Purcell-Gates, 2000). Wren (2001) contends that,

As long as educators are in any way expected to base their educational decisions on the issues, debates, politics and polemics of the Great Debate, and as long as we limit our horizons to approaches and philosophies that have been advocated by one faction or another, there is no reason to believe that real progress in reading education will ever be made. (p. 1)

The result of the fulminating debates over instructional and research methodology has been the perfect arena for commercial program developers to practice the art of misdirection. While public attention remains fixed on the public and often contentious debates by reading professionals, policy makers, and researchers, commercial program developers and profiteers have been met with open arms by school districts as partners in navigating their way through the accountability quagmire created by NCLB legislation.

Evidentiary Sleight of Hand Technique #2: Palming the Evidence

Palming is a sleight of hand technique in which a card or coin is concealed in the palm of the hand, used to trick an opponent in a card game or create an illusion. Commercial program developers employ this sleight of hand technique in program evaluations published to promote their products by palming evidence that does not support their program's effectiveness and revealing only supportive evidence and positive findings. School districts, working within the confines of NCLB, are compelled to ensure that all reading methodologies used within their schools are grounded in scientifically based reading research. Commercial program developers exploit this directive by publishing empirical evidence to promote their products. The evidence of efficacy program developers embed within promotional materials reveals several hallmark features:

- The evidence is generated and compiled by the program developers.
- Descriptive statistics are provided based on pre-test/post-test data.
- Graphic representations are artfully used as evidence.
- Terminology is used to develop confidence in alignment with federal accountability standards.
- There is very little or no peer-reviewed research.
- There is a layering of evaluation reports: reports readily available for prospective purchasers and more complete technical reports. Different publications lead to different conclusions of the program's effectiveness.

The evidence commercial programs publish of their program's effectiveness is primarily used to sell their products and, therefore, takes the form of an infomercial. Websites are often heavily laden with glowing testimonials from administrators, teachers, and parents.

Brightly colored graphs depict remarkable increases in student achievement on standardized measures. Photographs of happy children and teachers are displayed as proof of a program that provides sure solutions for academic woes. However, just as infomercials should not be relied upon for accurate representations of a product's worthiness, commercial programs' evidence of effectiveness should not be relied upon as proof positive that use of their products will ensure success for children learning to read.

Prospective purchasers of programs often find it difficult to cull from the wealth of impressive promotional materials the factual reality of the effects of the program on student learning. Some program developers make technical reports of program evaluations available. However, the evidence that is reported in promotional materials only reflects the most positive findings from the technical reports and it is only the most astute administrator who will read beyond the promotional materials and critically discern the validity of the actual evidence. Commercial programs employ a sleight of hand technique in order to palm the true evidence about their program's effectiveness while using marginal evidence to promote their products. An examination of one scripted, commercial reading program, Read Well, illustrates how commercial programs employ evidentiary sleight of hand to palm evidence that discredits their program's effectiveness and reveal only the evidence that is useful for promotional purposes.

Read Well is a commercial, scripted program, developed by Sopris West (a subsidiary of Cambium Learning) that is marketed in several states as a scientifically based core reading program aligned with state and federal standards. Among the research available on the Read Well website is a study of 144 kindergarten and 1st grade students in three schools in two Mississippi school districts. The primary document highlighting

data from the study is colorful and features graphic depictions of pre-test/post-test data depicting apparently incredible gains made by students who used Read Well compared with the dismal lack of gains made by students in a comparison group using basal readers or a literature based reading curriculum (Cambium Learning, 2007). One graph, for example, indicates (based on performance on the DIBELS Letter Naming Fluency that at the beginning of the study 67% of the Read Well kindergarteners were in the low-risk category and at the end of the study 73% were in the low-risk category. The graph, likewise, indicates that 81% of the comparison group were in the low risk category at the beginning of the study and 64% at the end of the study. In other words, the number of children in the Read Well group who were on the road to reading success increased by 6%, while the comparison group lost ground with 17% more students at risk for reading failure at the end of the year than at the beginning of the year. Five more graphs follow that seem to depict equally impressive results for students in the Read Well group.

However, there are caveats that detract from the impressiveness of the evidence of Read Well's effectiveness. One factor that taints the validity of the evaluation is the population size. While there were a reported 144 participants in the study, data from only 112 are depicted in graphic evidence. Data from 32 students – more than 20% of the total population – are missing and, therefore, not analyzed or reported. Additionally, it's notable that the kindergarten Read Well group was made up of only 15 students and, therefore, the Letter Name Fluency data is not nearly as impressive as the graphic depiction makes it appear. Given 15 students, reporting that 67% were in the low risk category translates into 10.05 students considered low risk at the beginning of the study. At the end of the study, 73% of the 15 students were in the low risk category – 10.95

students. Read Well's own evidence of their effectiveness doesn't make much sense because, according to their own data, 9/10ths of one student (or less than 1 student) made enough improvement to be no longer considered at risk for reading failure based on a measure of Letter Naming Fluency.

The flawed nature of the evidence touted by Read Well as proof of effectiveness is even more evident in their technical report. An examination of the full report highlights the manner in which the program developers cherry-pick a few items – those which would translate into a dramatic (albeit misleading) graphic rendering – for exploitation in promotional materials. Read Well's technical report reveals several key factors that dramatically alter the nature of their evidence. According to the authors of the technical report:

The study was originally intended to span one academic year, however, due to the impact of Hurricane Katrina the school year was shortened and the pretest did not occur until mid-February. The conditions under which instruction occurred differed somewhat across the groups. The Read Well program was not taught under typical classroom instruction conditions because the Read Well classrooms had an additional full-time teacher assisting in the classroom along with a paraprofessional whereas comparison classrooms had only a paraprofessional assisting in the classroom. Hence the results for the evaluation reflect the impact of students receiving 90 minutes of instruction per day over 3.5 months implemented by two teachers in each of the Read Well classrooms and by one teacher in each of the comparison classrooms. (Cambium Learning, 2007, pp 13-14)

Of course the research published by Read Well in their promotional material does not explain that the participants in the study were living and working within the aftermath of Hurricane Katrina. Arguably, the most compelling evidence – the part relegated to the last pages of the technical report – is that associated with kindergarten and 1st grade students returning to school after one of the deadliest natural disasters in the history of

the United States. Regardless, it is apparent that the conditions surrounding the study constituted a threat to the validity of the findings. The fact that the students in the Read Well classrooms were staffed with two full-time teachers and one paraprofessional, while the comparison classrooms had one teacher and one paraprofessional, would seem to give the Read Well groups an advantage and, therefore, calls into question the validity of the findings as well. The author of the technical report also discusses limitations of the study.

One set of limitations concerns the small number of participants which reduced statistical power. ... A related issue concerns differential attrition among the study groups. The reasons for the missing data were not recorded, therefore it was not possible to evaluate whether the loss of data was systematic or differentially impacted the study groups. (p. 14)

It's not unreasonable to speculate that this research would not fare well within a peer review process. However, it points out the problem with evidence that is propagated for the purpose of selling a product and how commercial program developers palm evidence that might negatively impact sales. Garan (2002) demonstrates similar examples of commercial programs (Open Court, Direct Instruction, Jolly Phonics, and Orton-Gillingham) palming evidence of negative or marginal evaluations and using only positive data to promote greater sales of their products. The fact that the program's reporting of its evaluations masquerades as scientific evidence, and is widely accepted by schools and school districts as evidence of an effective reading program, is particularly problematic.

Evidentiary Sleight of Hand Technique #3: The Shell Game

Sleight of hand practitioners know that in order to profit from the shell game, it's necessary for most of their audience to participate in the illusion. In other words, the people watching the shell game actually know that sleight of hand is being practiced and the only person who falls for the trick is the mark – the person who actually bets money on his ability to choose the right shell. The other participants in the shell game play a role in the illusion, giving the mark false confidence that he can win the game. Of course the perpetrators of the shell game know that the mark will never, can never, win. The shell game, for commercial program developers is employed to engage perspective purchasers. It works because teachers are unwittingly complicit in the deception – feigning fidelity to the program.

A hallmark of most, if not all, commercial reading programs is a demand for high fidelity implementation. According to Duncan-Owens and Hare (2007) the locus of control in reading instruction when commercial programs are used is within the program itself. Commercial program developers, either explicitly or through implication, promote the idea that the role of the teacher is limited to that of the deliver of the program and, therefore, the program will work equally well with any teacher (regardless of their level of training, experience, or competence) and for any children. Diamand (2004), contends fidelity to tightly designed and well engineered designs is what yields results. Mc-Gill-Franzen (2005), however, questions mandates for fidelity of implementation because it silences teachers' voices in decision making.

For more than 40 years researchers have investigated the effectiveness of commercial reading programs and their findings support McGill-Franzen's (2005) concerns about demands for fidelity of implementation. Bond and Dykstra (1967), after

analyzing 27 individual studies of reading instructional methods and programs, concluded that teacher quality was the single biggest contributing factor to student success. This conclusion about the role of teachers in reading instruction has been supported by contemporary studies (Allington, 2002; Pressley, Wharton-McDonald, Block, Morrow, Baker, Brooks, et al, 2001; Ryder, Sekulski, & Silbert, 2002).

Other researchers have examined how teachers use programs and instructional material and have concluded that fidelity of implementation is, for the most part, an illusion. Datnow and Castellano (2000) concluded that, in spite of demands by administrators and program developers for fidelity of implementation, teachers typically made adaptations in the reading programs they were assigned to use based on their own teaching styles, pedagogical beliefs, and the needs of their students. Sosniak and Stodolsky (1993) examined teachers' use of instructional textbooks and found that teachers maintained a great deal of autonomy about how and when they used the textbooks and other instructional materials. Duncan-Owens (2007) examined teachers' implementation of the Read Well program in demonstration classrooms and found that, while the teachers claimed to like the program in general, they found flaws in the program's design and found it necessary to veer from the program's script and supplement the program in order to meet the needs of their students. In an on-line forum for teachers using the Open Court Reading program one teacher recommended modification of the program, stating, "modify, modify, modify!!! ... What good teacher follows the curriculum to the letter? Not a single one" (Open Court, 2005). Researchers have noted benefits, such as higher student achievement in reading comprehension, associated with reading programs that facilitate teacher involvement in making decisions

about how or what to teach their students (Fang, Fu & Lamme, 2004; Tivnan & Hemphill, 2005; Wilson, Martens, & Poonam, 2005).

The concept of implementation fidelity is grounded in empirical research and is defined as the extent to which delivery of an intervention, or treatment, adheres to the program model (Mowbray, Holter, Teague, & Bybee, 2003). However, researchers have noted, in the case of commercial reading programs, teachers tend to abandon strict fidelity to the program in favor of meeting the needs of their students. Program developers, while making attempts to obtain fidelity of implementation, acknowledge variability in implementation. Borman, Slavin, Cheung, Chamberlain, Madden, and Chambers (2006) evaluated the effects of the Success for All program and noted that, while “many efforts were made to ensure fidelity to the experimental treatment,” there were factors that “inhibited quality implementation” (pp. 23-24). It’s notable, however, that program developers seem relatively untroubled by lack of strict fidelity when the results of the study yield conclusions favorable to the program. Borman, et al, concluded that Success for All yielded positive benefits for students within 35 schools in spite of the fact that few schools implemented all components of the program adequately and few had full time program facilitators to ensure fidelity of implementation. On the other hand, when there is evidence of adequate implementation fidelity, fidelity is cited as an essential component of the program’s effectiveness. Borman, Dowling, & Schneck (2006) in an evaluation of the efficacy of the Open Court Reading program attributed the positive effects of the program to the fact that “the treatment fidelity and OCR implementation quality seem reasonably good” (p. 27). The America’s Choice program, a for-profit subsidiary of a non-profit organization, National Center on Education and the

Economy, note on their website, “Districts and schools that adopt the America’s Choice School Design make a commitment to implement the design with fidelity” (America’s Choice, no date, p. 1). Any claims, therefore, of the program’s effectiveness are, therefore, predicated on the assumption that it has been implemented with fidelity. The underlying demand for and assumption of fidelity perpetuates the flawed logic employed by program developers that when programs are successful it is because of implementation fidelity. However, when programs fail, it is attributed to a lack of implementation fidelity.

Teachers often find themselves in the precarious position of having to decide between meeting the needs of their students or complying with demands by administrators and program developers for fidelity of implementation. Shanton and Valenzuela (2005) describe a teacher who questioned the ability of Success for All to meet the needs of his students and found himself labeled as insubordinate and being sanctioned. Teachers implementing Read Well consistently abandoned fidelity to the program in spite of possible recriminations by administrators (Duncan-Owens & Hare, 2007). One teacher commented, “Sometimes I feel like I’m having to choose between being a good employee and being the best teacher for my students” (p. 13). Another Read Well teacher simply adopted a “don’t ask, don’t tell” attitude, stating that as long as she wasn’t explicitly asked about fidelity to the program, she didn’t feel that she needed to explain how she veered from the program’s script.

The demand for program fidelity is a necessary component of the shell game as perpetrated by program developers because it makes teachers, albeit unwittingly, accomplices in the evidentiary sleight of hand. In the same way that scam artists rely on

skills to validate that the shell game is winnable, program developers solicit statements of fidelity from schools in order to maintain the illusion of their program's effectiveness. Teachers, under threat of recrimination by administrators, may be unwilling to admit lack of fidelity of implementation. Their implied fidelity, however, even in the absence of true fidelity, is sufficient for program developers whose primary interest is in the pre-test/post-test data for evidence of their program's success. The teachers, who are ancillary to reading instruction as deliverers of the program, are unwittingly used as participants in the evidentiary sleight of hand.

The High Stakes of Silencing Teachers

The evidence of program effectiveness would be greatly altered if the voices of the teachers who used them were not silenced by the presumption of fidelity of implementation. There are high stakes associated with program evaluations that harvest pre-test/post-test data, while overlooking what happens between measures. From wasted time, wasted evidence, and wasted resources, the cost of evidentiary sleight of hand are too high to ignore.

The first year a school implements a commercial program, it's assumed that there will be a certain amount of time lost on training teachers to use the material. It is often in the second year of implementation when teachers are comfortable using the program and can truly begin to form an opinion about its effectiveness. Therefore, it may not be until the third year that teachers are able to make a case for retaining or abandoning the program. Rather than engaging in meaningful explorations of how students learn to read, the teachers have spent several years implementing a program that may or may not be the best curriculum for their students.

Commercial reading programs are big business. The decision to implement a commercial reading program is a costly one. When a program fails to yield promised results those funds are wasted, never to be reclaimed. Funds that could have been used for high quality professional development designed to equip teachers as decision makers about their students' academic needs have been lost to program materials that are often relegated to dusty shelves or supply closets once they have been found to be inadequate in meeting the needs of students.

Most importantly, there is a loss of valuable evidence of how children learn to read. Between the pre-test and post-test measures designed to measure the effectiveness of a program, teachers, as they make the difficult decisions about how to alter their use of the program in order to meet their students' needs, often develop strategies that are worthy of further investigation. Program developers, who insist on strict implementation of fidelity and fail to engage in discussions with teachers about how they alter their use of programs, miss opportunities to make their programs better. Because the alterations represent lack of fidelity, teachers are often not free to share their strategies with colleagues and miss the opportunity to participate in constructive discourse about how children develop as literacy learners. The high stakes of silencing teachers extend beyond threats to the internal validity of program evaluations, they represent a wealth of wisdom and practice lost in an environment that is only focused on pre- and post-test results and producing evaluations useful for promotional materials.

REFERENCES

Allington, R. L. (2002) *Big brother and the national reading curriculum: how ideology trumped evidence*. Portsmouth, NH: Hienemann.

- America's Choice (no date). Retrieved February 6, 2008, from <http://www.americaschoice.org/index.cfm>
- Bond, G.L., & Dykstra, R. (1967). The cooperative research program on first-grade reading instruction. *Reading Research Quarterly* (11)4, 5-142.
- Borman, G. D.; Dowling, N. M. & Schneck, C. (2006). The national randomized field trial of Open Court Reading. The University of Wisconsin at Madison. Retrieved February 7, 2008, from http://www.wcer.wisc.edu/itp/Fall%2007%20seminar/BormanFINAL%20OCR%20RFT9%20to%20EEPA_2_.pdf
- Borman, G.D., Slavin, R.E., Cheung, A. & Chamberlain, A., Madden, N. & Chambers, B. (2006). Final reading outcomes of the national randomized field trial of Success for All. Retrieved February 6, 2008, from http://www.successforall.com/_images/pdfs/Third_Year_Results_06.doc
- Cambium Learning (2007). An evaluation of the Read Well curriculum with at-risk student populations. Retrieved February 6, 2008, from http://store.cambiumlearning.com/Resources/Research/pdf/sw_Research_ReadWell_MI01.pdf
- Datnow, A. & Castellano, M. (2000). Teachers' responses to Success for All. How beliefs, experiences and adaptation shape implementation. *American Educational Research Journal*, 37, 775-799.
- Diamond, L. (2004). High fidelity: It's not about music or marriage. It's all about instructional materials. Bay Area School Area Collaborative.
- Duncan-Owens, D. (2007). *Reforming reading instruction in Mississippi through demonstration classes: Barksdale literacy teachers' first year experiences*. Unpublished doctoral dissertation, Mississippi State University, Starkville.
- Duncan-Owens, D. & Hare, D. (in press) Expert reading teachers and commercial programs: Lessons learned from demonstration classrooms. *Language Arts*.
- Fang, Z., Fu, D., & Lamme, L.L. (2004). From scripted instruction to teacher empowerment: Supporting literacy teachers to make pedagogical transitions. *Literacy*, 38(1) 58-64.
- Garan, E. M. (2002). *Resisting reading mandates: How to triumph with the truth*. Portsmouth, NH: Heinemann.
- Garan, E. M. (2004). *In defense of our children: When politics, profit, and education collide*. Portsmouth, NH: Heinemann.

- Hempenstall, K. (2008). The three-cueing model: Down for the count? EdNews.Org, February 11, 2008. <http://www.ednews.org/articles/4084/1/The-three-cueing-model-Down-for-the-count/Page1.html>
- International Reading Institute. Board Position: Statement Calling for Civil Dialogue. Newark, Delaware. http://www.ira.org/downloads/positions/ps0301_civil.pdf
- McGill-Granzen, A. (2005). In the press to scale up, what is at risk? *Reading Research Quarterly*, 40(3), 366-370.
- Moats, L. (2007). Whole-Language high jinks. Thomas B Fordham Institute. Available: <http://vcww.edexcellence.net/institute/publication/publication.cfm?id=367&pubsid=1455#1455>
- Mowbray, C. T.; Holter, M. C., Teague, G. B., & Bybee, D. (2003). Fidelity criteria: Development, measurement, and validation. *American Journal of Evaluation*, 24, 315-340.
- National Institute of Health (2000). *Report of the National Reading Panel: Teaching children to read: An evidence-based assessment of the scientific literature on reading and its implications for reading instruction: Reports of the sub-groups* (NIH Publication No. 00-4754). Washington DC: U.S. Government Printing Office. Available on-line: <http://www.nichd.nih.gov/publications/nrp/report.htm>
- Open Court Reading (2005). Teach-nology: The online teacher resource. Available: <http://www.teach-nology.com/forum/showthread.php?t=402>
- Pearson, P. D. (2005). *Using research to build a responsible pedagogy: Making a difference for students and teachers*. Retrieved August 2, 2006, from www.scienceandliteracy.org
- Pressley, M., Wharton-McDonald, R., Allington, R., Block, C.C., Morrow, L., Tracey, D., et al (2001). A study of effective grade-1 literacy instruction. *Scientific Studies in Reading*, 5, 35-58.
- Purcell-Gates, V. (2000). The role of qualitative and ethnographic research in educational policy. *Reading Online*, 4(1). Retrieved March 12, 2006, from <http://www.readingonline.org/articles/purcell-gates/>
- Robinson, B. (2004). Misdirection. <http://www.illusiongenius.com/3-04.html>
- Ryder, R.J., Sekulski, J., & Silbert, A. (2003). *Results of Direct Instruction reading program evaluation first through second grade, 2000-2002*. Wisconsin Department of Public Instruction, Madison, WI.

- Shanton, K. D. & Valenzuela, T. C. (2005). *Not in the script: The missing discourses of parents, students, and teachers about Success for All*. In L. Poyner & P. M. Wolfe (Eds.), *Marketing Fear in America's Public Schools*. Mahwah, N.J.: Lawrence Erlbaum Associates, Publishers.
- Smith, K. (1989). Overselling literacy. *Phi Delta Kappan*, 70, 353-359.
- Sosniak, L.A., & Stodolsky, S.S. (1993). Teachers and textbooks: Materials use in four fourth grade classrooms. *The Elementary School Journal*, 93(3), 249-275.
- Tivnan, T. & Hemphill, L. (2005). Comparing four literacy models in high poverty schools: Patterns of first-grade achievement. *The Elementary School Journal*, 105, 419 – 441.
- Wilson, P., Martens, P., & Poonam, A. (2005). Accountability for reading and readers: What the numbers don't tell. *The Reading Teacher*, 58(7), 622-631.
- Wren, S. (2001). Topics in early reading coherence: What does a balanced literacy approach mean? Southwest Educational Development Laboratory (SEDL). <http://www.sedl.org/pubs/catalog/items/read13.html>