Reporting Data with "Over-the-Counter" Data Analysis Supports Increases Educators' Analysis Accuracy

Jenny Grant Rankin, Ph.D.

Northcentral University

Department of Education

Paper presented at the 2013

California Educational Research Association

Conference

Anaheim, California

December 5, 2013

To contact the author of this paper, visit her personal website: www.overthecounterdata.com.

Abstract

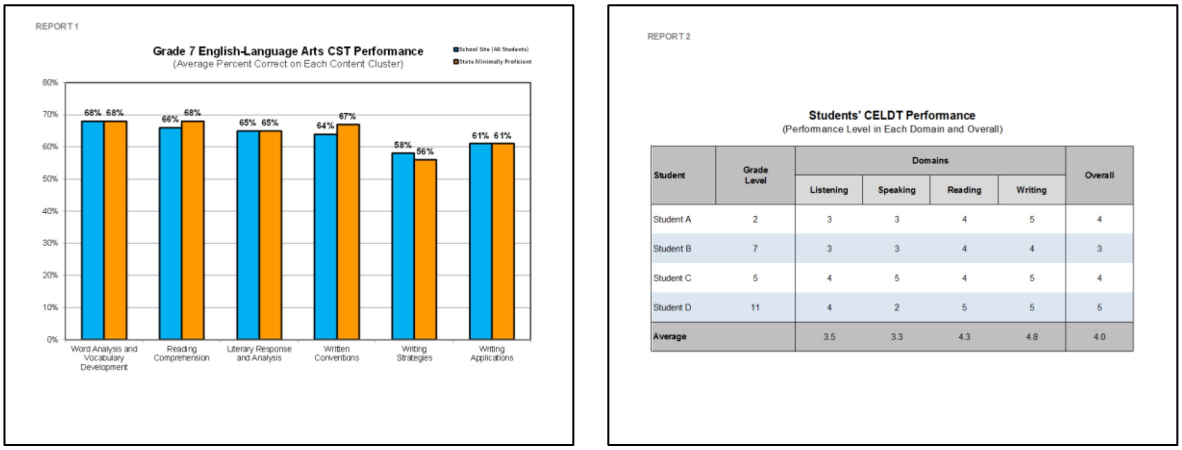
There is extensive research on the benefits of making data-informed decisions to improve learning, but these benefits rely on the data being effectively interpreted. Despite educators' above-average intellect and education levels, there is evidence many educators routinely misinterpret student data. Data analysis problems persist even at districts where there is proactive support for data use, as another variable plays a significant role in rendering successful or unsuccessful data use: the tool educators use for data analyses, which is typically a data system. These data systems and their reports usually display figures *without* supporting guidance concerning the data's proper analysis. A solution to analysis errors lies in the data-equivalent to over-the-counter medicine, termed *over-the-counter data*: essentially, enlisting medical labeling conventions to pair education data reports with straightforward verbiage on the proper interpretation of report contents. The researcher in this experimental, quantitative study explored the inclusion of such supports in education data systems and their reports, while also investigating varied formats for each support. The cross-sectional sampling procedure incorporated responses from 211 educators of varied backgrounds and roles at nine elementary and secondary schools throughout California. Participants answered survey questions regarding student data reports with varied forms of analysis guidance. Respondents' data analyses were found to be 307% more accurate when a report footer was present, 205% more accurate when an abstract was present, and 273% more accurate when an interpretation guide was present. Findings were significant and fill a void in field literature with evidence that can be used to identify how data systems can increase educators' data analysis accuracy by offering analysis support through labeling and supplemental documentation.

Purpose of the Study

The purpose of the experimental, quantitative study was to facilitate causal inferences concerning the degree to which including different forms of data usage guidance within a data system reporting environment can improve educators' understanding of the data contents, much like including different forms of usage guidance with over-the-counter medication is needed to improve use of contents. The researcher presented student achievement data report sets to 211 elementary and secondary educators in California. Each of these report sets fit into one of the following treatment categories:

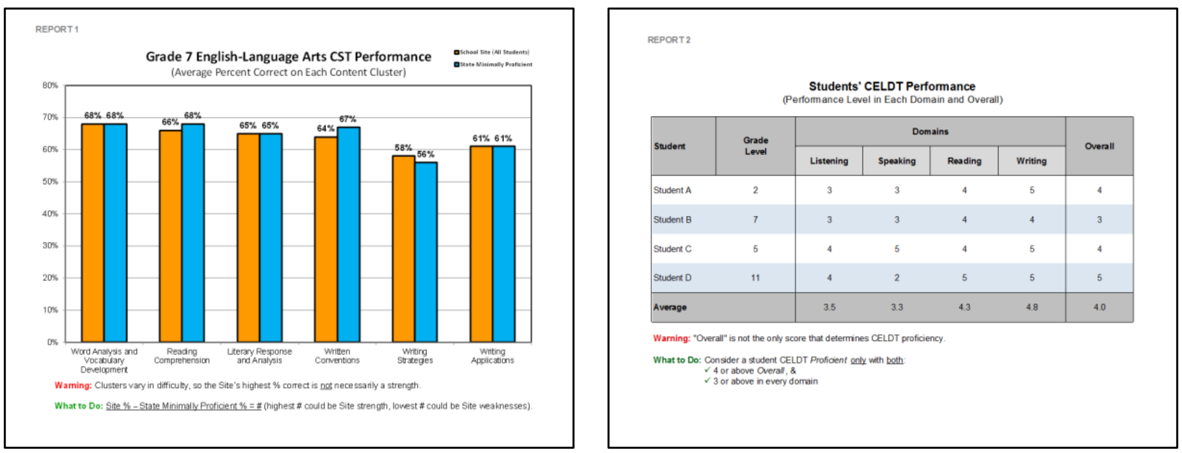
- (a) control group with no added analysis support (Figure 1);
- (b) analysis support by way of footers directly on the reports, which were offered in two different framing styles (Figures 2 and 3);
- (c) analysis support by way of report abstracts, which accompanied the reports and were offered in two different framing styles (Figures 4 and 5); and
- (d) analysis support by way of interpretation guides, which accompanied the reports and were offered in two different framing styles (Figures 6 and 7).

The study's primary independent variables included the three data analysis supports (b-d, above), which can be generated within a data system, in varied formats. The dependent variable was accuracy of data analysis-based responses. See Table 1 for primary research questions and hypotheses. Findings from this research are suited to identify how data systems used by educators can help prevent common analysis mistakes by providing analysis support within the interface and the reports they are used to generate. Secondary research questions concerning the impact of site demographics and educator



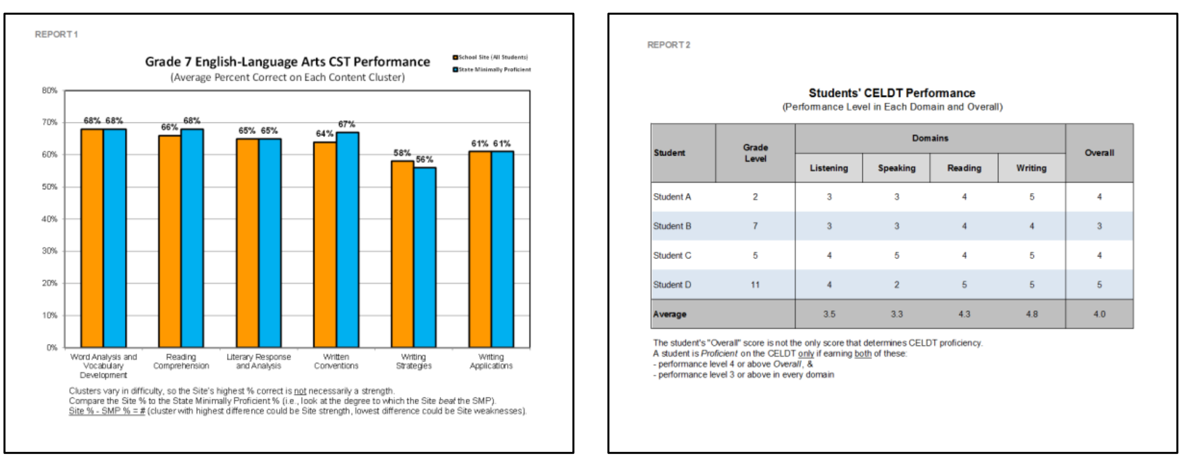
Student	Grade Level	Domains				Overall
		Listening	Speaking	Reading	Writing	
Student A	2	3	3	4	5	4
Student B	7	3	3	4	4	3
Student C	5	4	5	4	5	4
Student D	11	4	2	5	5	5
Average		3.5	3.3	4.3	4.8	4.0

Figure 1: *Scenario 1 Participant (Control Group) Handouts*



Student	Grade Level	Domains				Overall
		Listening	Speaking	Reading	Writing	
Student A	2	3	3	4	5	4
Student B	7	3	3	4	4	3
Student C	5	4	5	4	5	4
Student D	11	4	2	5	5	5
Average		3.5	3.3	4.3	4.8	4.0

Figure 2: *Scenario 2 (Footer A) Participant Handouts*



Student	Grade Level	Domains				Overall
		Listening	Speaking	Reading	Writing	
Student A	2	3	3	4	5	4
Student B	7	3	3	4	4	3
Student C	5	4	5	4	5	4
Student D	11	4	2	5	5	5
Average		3.5	3.3	4.3	4.8	4.0

Figure 3: *Scenario 3 (Footer B) Participant Handouts*

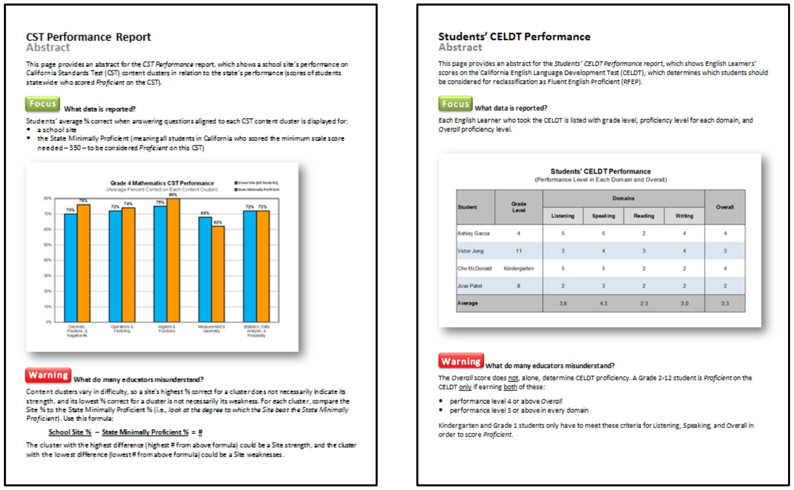


Figure 4: *Scenario 4 Participant (Abstract A) Handouts; These Participants Also Received Figure 1 Handouts*

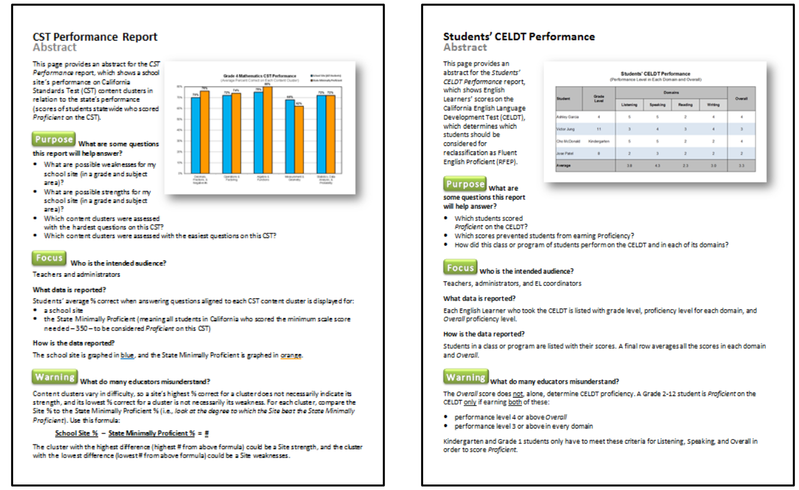


Figure 5: *Scenario 5 Participant (Abstract B) Handouts; These Participants Also Received Figure 1 Handouts*

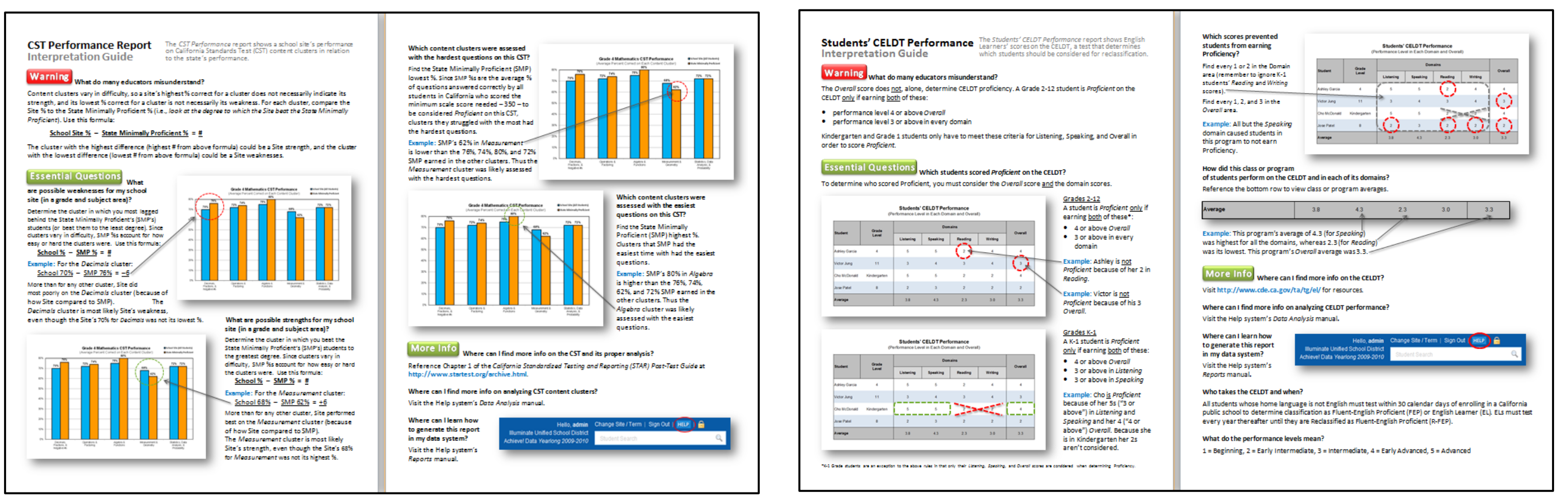


Figure 6: *Scenario 6 Participant (Interpretation Guide A) Handouts; These Participants Also Received Figure 1 Handouts*

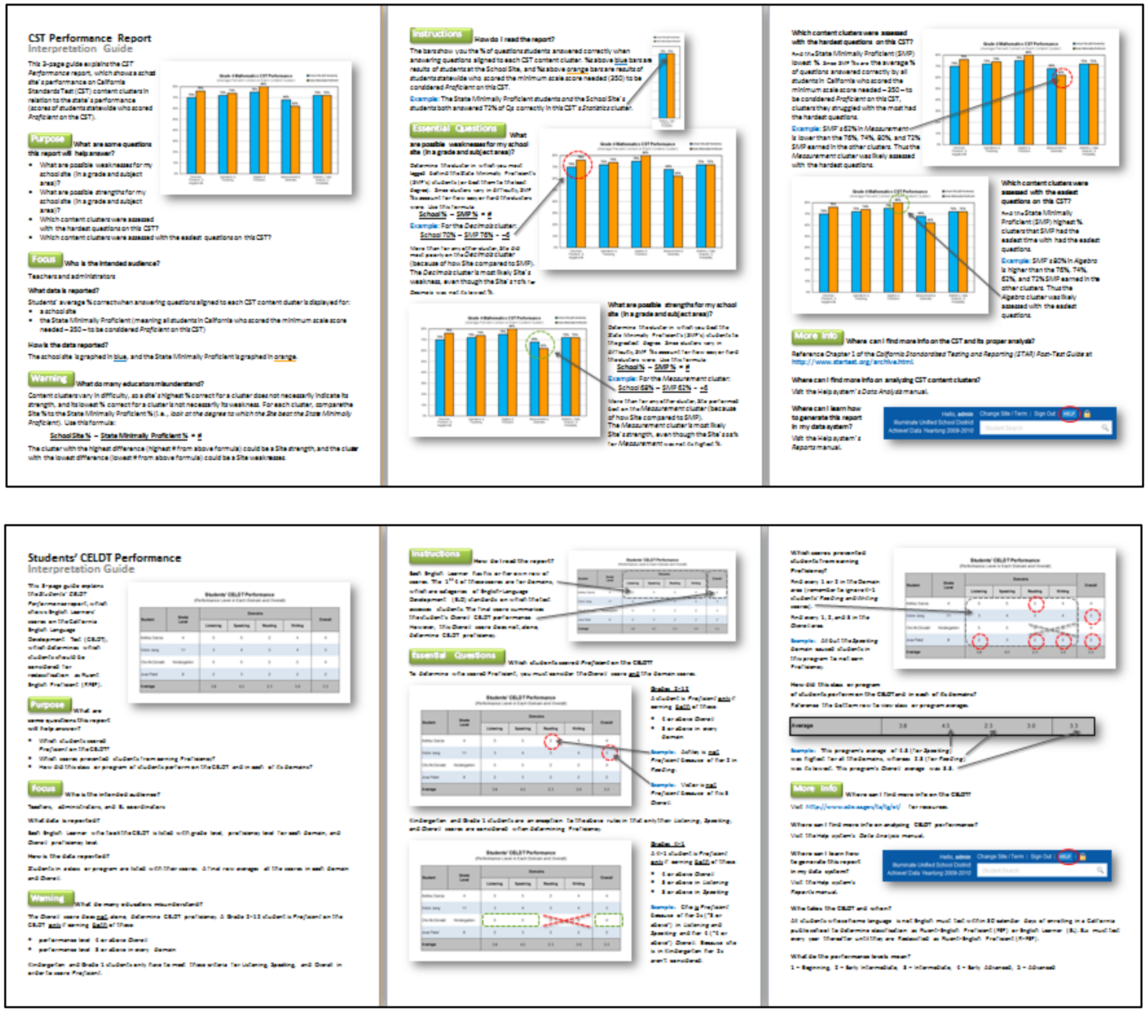


Figure 7: *Scenario 7 Participant (Interpretation Guide B) Handouts; These Participants Also Received* Figure 1 *Handout*

Table 1: *Primary Research Questions with Alternative Hypotheses and Linear Regression Analyses Applied to Research Question Variables*

Abbreviated Research Question	Alternative Hypothesis	Linear Regression Relationships
Q1. Support's (Meaning Footer, Abstract, or Interpretation Guide) impact on analysis accuracy	**H1$_a$.** Accompanying a report with a support containing analysis guidance in the form of footer, abstract, or interpretation guide would have a positive impact on the frequency of accurate conclusions educators drew concerning student achievement data.	A = f(S) *A (Analysis Accuracy) is a function of S (Support)* $A = \alpha+\beta S$
Q2a. Footer's impact on analysis accuracy	**H2a$_a$.** Accompanying a report with a supportive footer would have a positive impact on the frequency of accurate conclusions educators drew concerning student achievement data.	A = f(F) *A (Analysis Accuracy) is a function of F (Footer)* $A = \alpha+\beta F$
Q2b. Footer framing's impact (moderate variations) on analysis accuracy	**H2b$_a$.** The manner in which a footer was framed, in terms of moderate differences in length and text color, would have an impact on the frequency of accurate conclusions educators drew concerning student achievement data.	A = f(FF) *A (Analysis Accuracy) is a function of F (Footer's Framing)* $A = \alpha+\beta FF$
Q3a. Abstract's impact on analysis accuracy	**H3a$_a$.** Including a report abstract with a report would have a positive impact on the frequency of accurate conclusions educators drew concerning student achievement data.	A = f(B) *A (Analysis Accuracy) is a function of B (Abstract)* $A = \alpha+\beta B$
Q3b. Abstract framing's impact (moderate variations) on analysis accuracy	**H3b$_a$.** The manner in which an abstract was framed, in terms of moderate differences in density and header color, would have an impact on the frequency of accurate conclusions educators drew concerning student achievement data.	A = f(BF) *A (Analysis Accuracy) is a function of BF (Abstract's Framing)* $A = \alpha+\beta BF$
Q4a. Interpretation guide's impact on analysis accuracy	**H4a$_a$.** Including an interpretation guide with a report would have a positive impact on the frequency of accurate conclusions educators drew concerning student achievement data.	A = f(I) *A (Analysis Accuracy) is a function of I (Interpretation Guide)* $A = \alpha+\beta I$
Q4b. Interpretation guide framing's impact (moderate variations) on analysis accuracy	**H4b$_a$.** The manner in which an interpretation guide was framed, in terms of moderate differences in length and information quantity, would have an impact on the frequency of accurate conclusions educators drew concerning student achievement data.	A = f(IF) *A (Analysis Accuracy) is a function of IF (Interpretation Guide's Framing)* $A = \alpha+\beta IF$

demographics were also investigated in relation to their impact on the primary research questions.

Theoretical Framework

The Food and Drug Administration (FDA) requires over-the-counter medication to be accompanied by textual guidance proven to improve its use, deeming it negligent to do otherwise (DeWalt, 2010). With such guidance, patients may take over-the-counter medication with the goal of improving wellbeing while a doctor is not present to explain how to use the medication. No or poor medication labels have resulted in many errors and tragedy, as people are left with no way to know how to use the contents wisely (Brown-Brumfield & DeLeon, 2010).

Labeling conventions can translate to improved understanding on non-medication products, as well (Hampton, 2007; Qin et al., 2011). Thus, in the way over-the-counter medicine's proper use is communicated with a thorough label and added documentation, a data system used to analyze student performance can include components to help users better comprehend the data it contains. Yet data systems display data for educators without sufficient support to use their contents – data – wisely (Coburn, Honig, & Stein, 2009; Data Quality Campaign [DQC], 2009, 2011; Goodman & Hambleton, 2004; National Forum on Education Statistics [NFES], 2011). Labeling and tools within data systems to assist analyses are uncommon, even though most educators analyze data alone (U.S. Department of Education Office of Planning, Evaluation and Policy Development [USDEOPEPD], 2009). Essentially, data systems do not commonly present data in an "over-the-counter" format for educators, whose primary option for using data to treat students is thus akin to ingesting medicine from an unmarked or marginally marked container.

Unfortunately, the resultant data analyses are flawed. Educators often do not use data correctly, and there is clear evidence many users of data system reports have trouble understanding the data (Hattie, 2010; Wayman, Snodgrass Rangel, Jimerson, & Cho, 2010; Zwick et al., 2008). For example, in a national study of districts known for *strong* data use, teachers incorrectly interpreted 52% of data (USDEOPEPD, 2009). Teachers at 13 school districts considered exemplars of active data use, where teachers receive support in using data systems to make decisions, only achieved 48% correct when making data inferences involving basic statistical concepts, and it is unlikely teachers at other districts would perform better (USDEOPEPD, 2011). Stakeholders at all levels have trouble interpreting data, including principals and teacher coaches (Underwood, Zapata-Rivera, & VanWinkle, 2008). Data interpretation has become increasingly vital to school reform (Minnici & Hill, 2007), yet misunderstandings about how to use data and a data system can cripple data use in a school district (Wayman, Cho, & Shaw, 2009). If data system users do not understand how to properly analyze data, the data will be used incorrectly if it is used at all (NFES, 2011).

Professional development (PD) can improve educators' data analysis accuracy (Lukin, Bandalos, Eckhout, & Mickelson, 2004; Sanchez, Kline, & Laird, 2009; Zwick et al., 2008). Staff resources such as site leaders, data teams, data experts, and/or instructional coaches can improve educators' data analysis accuracy (Bennett & Gitomer, 2009; McLaughlin & Talbert, 2006). However, PD is not without limitations (Lock, 2006; Kidron, 2012; O'Hanlon, 2013; USDEOPEPD, 2011; Zapata-Rivera & VanWinkle, 2010), nor are staff supports without limitations (McDonald, Andal, Brown, & Schneider, 2007; Underwood et al., 2008; Wayman et al., 2010). Since even districts enlisting these approaches continue to struggle with data use, more needs to be done to support educators.

Data analysis difficulties should not be mistaken as criticisms of educators, and the problem should not be mistaken as failure on the part of educators. Rather, this study was based on recognition that a population surpassing the general public in schooling and intellect yet still struggling with data analyses, despite its own efforts to rectify the problem, might be using tools that are flawed in their ability to render accurate analyses.

The power of data systems will not be realized until researchers contribute to improving data system design to improve analysis (DQC, 2011). Literature that did examine data system and report format, including how effectively this format communicates data to users, focused on participants' preferences and perceived value of supports. However, user preference can be the opposite of the reporting format that actually renders more accurate interpretation (Hattie, 2010). In order to improve data use, practitioners and researchers need to gather empirical evidence to support different ways in which data is reported (Lyrén, 2009). This study was unique in determining the specific extent to which each form of analysis guidance improves analysis accuracy rather than relying on participants' *perceived* value of supports. The findings of this study filled a gap in education field literature by containing evidence that can be used to identify how data systems can help increase educators' data analysis accuracy by providing analysis support within data systems and their reports, and rendered examples and templates for real-world implementation. Improvements data system and report providers make in light of this study have potential to improve the accuracy with which educators analyze the data generated by their data systems. This improvement will likely benefit students.

Methods

This experimental, quantitative study measured how effective three data analysis supports, which can be featured in data systems but typically are not, are in improving educators' data analysis accuracy:

- labeling in the form of brief, cautionary verbiage in data system report footers (Figures 2 and 3);
- supplemental documentation in the form of report abstracts that can be reached via link in a data system and can also be printed to accompany printed reports (Figures 4 and 5); and
- supplemental documentation in the form of interpretation guides that can be reached via link in a data system and can also be printed to accompany printed reports (Figures 6 and 7).

Participants answered survey questions regarding student data reports they received, which featured varying levels and forms of embedded analysis guidance. In addition to establishing the data analysis accuracy rendered by educators using reports with no added supports (Figure 1), the survey was used to measure the specific impact the three above-listed variables have on educators' data analysis accuracy. The study was pilot-tested first, subscribed to all Institutional Review Board (IRB) and ethical guidelines, and reflected precautions to avoid or overcome threats to external and internal validity.

Sample

A priori two-tailed t-test (effect size d = 0.5, α error of probability = 0.05, power = 0.95), rendered a recommended sample size of at least 210 participants. A priori F-test linear multiple regression analysis (effect size $f^2 = 0.15$, α error of probability = 0.05, power = 0.95, predictors based on independent variables = 7) rendered a recommended sample size of at least 153 participants. The study employed a random, cross-sectional sampling procedure when incorporating responses from 211 educators of all school levels spanning transitional kindergarten (TK) through twelfth grade, at all veteran levels, working in varied roles, and at schools with a range of demographics. These educators were employed at nine schools in six school districts, six cities, and three counties in California. The sample accurately reflected the study's population, which is comprised of public educators of all primary and secondary school levels.

Behavioral Economics

This study related to improving the accuracy of educators' data analyses, as enacted in the thought portion – or "data-informed" portion – of data-informed decision-making. The process of thinking and deciding is influenced by behavioral economics facets such as priming, biases, heuristics, prototypes, judgments, anchoring, and framing (Kahneman, 2011). Thus data-informed thoughts are believed to influence decision-making. For example, even small and seemingly insignificant differences in how content is arranged can mean a significant difference in the decisions people make based on that content (Thaler & Sunstein, 2008). This study's design reflected consideration of all key facets of behavioral economics but related particularly to framing.

Framing applies to the presentation of information, and presenting the same information to someone in different ways will often result in different levels of difficulty in understanding or analyzing the information (Kahneman, 2003, 2011). The manner in which content is organized for people using it to make decisions significantly impacts those decisions (Thaler & Sunstein, 2008). Framing thus plays a large role in data analysis accuracy and data-informed decision-making.

The reports used in this study subscribed to leading research-based recommendations concerning the best ways in which to frame the data in report format, though they did so in a way that did not deviate from what is commonly seen in data systems currently on the market. In other words, the study's report handouts adhered to the better data presentations commonly seen in data systems, but they did not adhere to the best data presentations that – despite being more effective – are not yet commonly seen in student data systems. Suggested ways to present analysis guidance in footers, abstracts, and interpretation guides were utilized in this study, but the best manner in which to frame these resources had not yet been determined in regards to direct impact on analysis accuracy. Thus each of the three support resources used in this study were framed in two different formats for respondents.

Materials/Instruments

Survey. Participant responses were collected through an anonymous, web-based survey crafted and administered in Google Docs, employing the Google Form feature, with the researcher present. The survey included 10 multiple choice questions involving respondent background and the analysis of data contained in report handouts. The survey was crafted with attention to validity and reliability considerations, as well as opportunities for within-method methodological triangulation.

All *analysis* survey questions concerned data from state assessments with which the Californian study participants were most likely to be familiar with analyzing: the California Standards Test (CST), constituting the largest component of California's Standardized Testing and Reporting (STAR) Program, and the California English Language Development Test (CELDT), which California educators must use when determining reclassification recommendations for any English Learner (EL).

Handouts. All participants received two reports containing the same data. The control group received plain reports with no analysis supports, whereas all other participants also received either footers, abstracts, or interpretation guides (see Figures 1-7). Data analysis supports used in the study adhered to research-based best practices to the fullest extent possible, inspired by literature such as Odendahl (2011) and Sabbah (2011). However, given controversies concerning framing, each support was framed in two slightly different ways, with minor differences in length, density, and color usage. In order to mimic real-world conditions, the abstracts and interpretation guides addressed all major questions the reports were designed to answer, as opposed to being geared exclusively toward the questions asked in this study's survey.

Likewise, all handouts mimicked real world environments by being distributed in hard copy format. While some teachers (44%) use their data system directly, most (56%) have access but do not use their data system directly and instead only read printed versions of reports others used the data system to generate (Underwood et al., 2008). This design was also needed to better isolate the impact of study variables, as viewing a report on the computer can negatively impact how it is interpreted. For example, someone who correctly interprets a printed report can make mistakes when scrolling is involved, users are more likely to scan a report on a computer that they would read carefully when printed, and users' inability to mark on the screen can reduce the credibility users attribute to reports (Hattie, 2010; Leeson, 2006).

Coding and analysis. The Google Docs Form tool automatically assigned an anonymous ID to each respondent's data, which was used in complete absence of participant names or employee numbers. The data was automatically, securely stored and password-protected online as soon as it was entered, and was exported into Microsoft Excel® shortly afterwards in order to be coded in accordance with a code book (columns A-JH for each respondent) and analyzed with the Microsoft 2010 Data Analysis feature and Predictive Analytics Software (PASW) Version 18 with the Statistical Package for the Social Sciences (SPSS) Data Access Pack. Results were analyzed to (a) answer research questions with related hypothesis strands, and (b) identify themes, patterns, relationships, and implications.

Independent samples T-Tests and crosstabulations with Chi-square were used to investigate variables. The dependent variable was data analysis accuracy. The primary independent variables included brief, cautionary verbiage in (a) report footers, (b) report-specific abstracts, and (c) report-specific interpretation guides, as well as the framing of these supports. Secondary variables were investigated to add insight to the primary research questions: school site demographics (school level type, school level, academic performance, EL population, Socioeconomically Disadvantaged population, and Students with Disabilities population) and educator demographics (veteran status, current professional role, perception of his or her own data analysis proficiency, data analysis PD time, and number of graduate-level educational measurement courses).

Results

This paper uses the following terms:

- *support*, meaning any or one of the following supports: footer, abstract, or interpretation guide
- *support use*, meaning instances in which respondents indicated they used the available support
- *data analysis accuracy*, meaning the mean value of participants' percent correct scores earned when answering survey questions measuring data analysis accuracy

All supports used in the study – footers, abstracts, and interpretation guides – had a significant, positive impact on the participating educators' data analysis accuracy. This resulted in acceptance of the alternative hypotheses for primary Research Questions Q1, Q2a, Q3a, and Q4a (described in Table 1). Specifically, educators' data analyses were:

- 264% more accurate (with an 18 percentage point difference) when any one of the three supports was present and 355% more accurate (with a 28 percentage point difference) when respondents specifically indicated having used the support,
- 307% more accurate (with a 23 percentage point difference) when a footer was present and 336% more accurate (with a 26 percentage point difference) when respondents specifically indicated having used the footer,
- 205% more accurate (with a 12 percentage point difference) when an abstract was present and 300% more accurate (with a 22 percentage point difference) when respondents specifically indicated having used the abstract, and

- 273% more accurate (with a 19 percentage point difference) when an interpretation guide was present and 436% more accurate (with a 37 percentage point difference) when respondents specifically indicated having used the guide.

Overall, the 211 study participants indicated they used supports 62% of the time. 87% of participants who receive no supports indicated they would have used footers, abstracts, or interpretation guides if the supports had been available.

When no supports were used, data analysis accuracy was 11%. All 211 participants, regardless of support use, averaged a data analysis accuracy of 26%. In cases where respondents indicated they used an available support, data analysis accuracy was 39%. See Figure 8 for visual representation of the breakdown of support impact shown in Table 2.

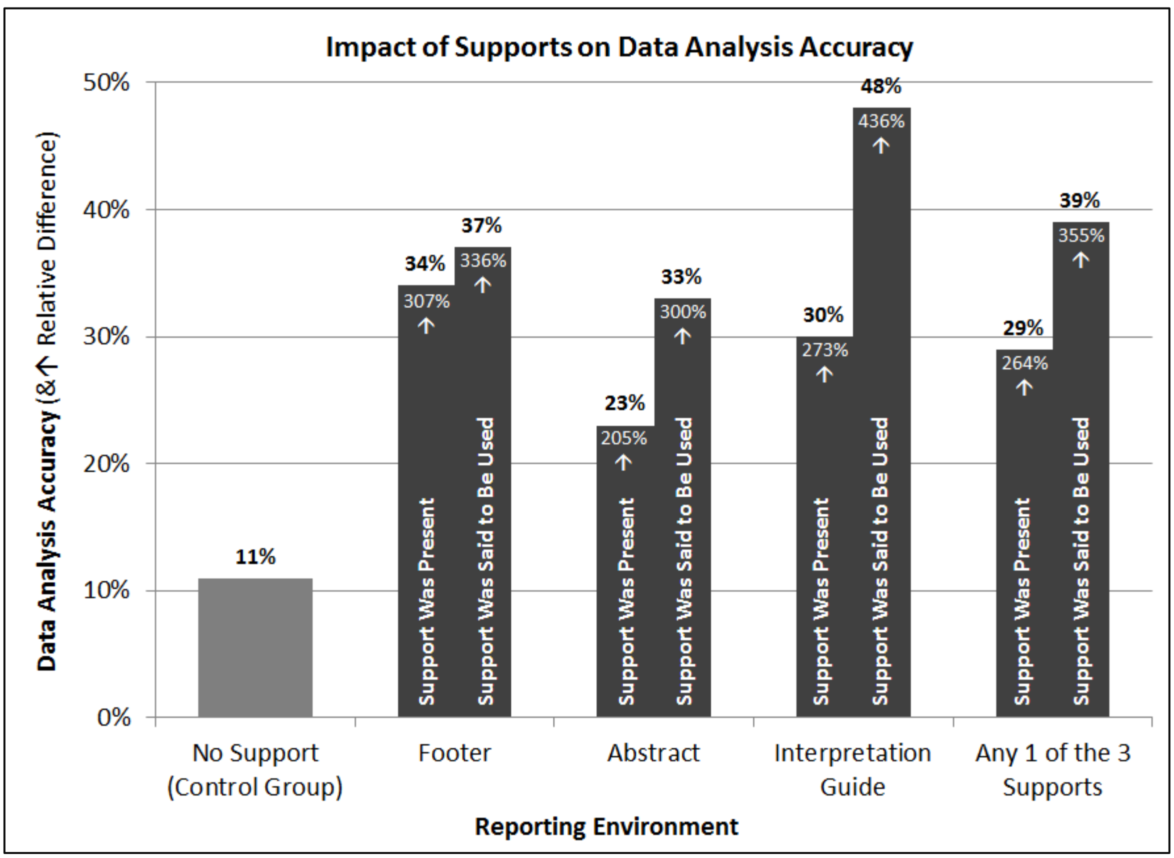


Figure 8: *Impact of Supports in Terms of Analysis Accuracy and Relative Difference*

Table 2: *Support Use and Data Analysis Accuracy in Each Report Environment*

	Participants		Use	Data Analysis Accuracy (% Correct)		
Report Environment	*n*	%	% Used/ Wanted Support	Did Not Use Support	Regardless of Support Use	Used Available Support
Plain Report (Control Group)	**31**	**15%**	**87%**	**11%**	**11%**	**n/a**
Report with Shorter Footer	30	14%	75%	27%	36%	33%
Report with Longer Footer	30	14%	70%	6%	32%	40%
Report with Any Footer	**60**	**28%**	**73%**	**15%**	**34%**	**37%**
Plain Report + Less Dense Abstract	30	14%	53%	11%	21%	31%
Plain Report + Denser Abstract	30	14%	47%	9%	24%	36%
Report with Any Abstract	**60**	**28%**	**50%**	**10%**	**23%**	**33%**
Plain Report + 2-Page Interpretation Guide	30	14%	52%	0%	32%	48%
Plain Report + 3-Page Interpretation Guide	30	14%	52%	3%	28%	48%
Report with Any Interpretation Guide	**60**	**28%**	**52%**	**2%**	**30%**	**48%**
Report with Any Support	**180**	**85%**	**58%**	**8%**	**29%**	**39%**

Results were expected to be positive *when supports were used* given previously-existing literature recommending the presence of footers, abstracts, and interpretation guides. However, some literature suggested the supports would not be utilized and would be rendered ineffective. Not only did the supports prove to have a significant, positive impact on data analysis accuracy, but the substantial rate at which they were utilized rendered their value significant for *all* educators as a whole, even when respondents' use of the supports was not considered. Nonetheless, respondents' data analyses were even higher when they indicated having used the available support.

The minor modifications in support format, mainly in terms of length and color usage, had no significant impact on participating educators' data analysis accuracy. This resulted in acceptance of the null hypotheses for primary Research Questions Q2b, Q3b, and Q4b (questions in Table 1). These results were somewhat unexpected given literature on behavioral economics, particularly in the area of framing, and literature on report and documentation design. However, it is important to note all support format variations used in the study subscribed to leading best practices in design. Thus the variations were minor and designed to garner more specificity in these best practices. It was thus concluded such minor variations are also minor in their impact on educators' data analyses.

Additional, secondary research questions were used to add insight to the primary research questions. Findings in relation to these questions determined that educators' school site demographics had no significant impact on their data analysis accuracy that might impact the primary research questions. In other words, an educator's school level type, school level, academic performance, EL population, Socioeconomically Disadvantaged population, or Students with Disabilities population had no significant impact on data analysis accuracy. Likewise, findings in relation to the secondary questions determined that educators' demographics had no significant impact on their data analysis accuracy that might impact the primary research questions. In other words, an educator's veteran status, current professional role, perception of his or her own data analysis proficiency, data analysis PD time, and number of graduate-level educational measurement courses had no significant impact on data analysis accuracy. This resulted in acceptance of the alternative hypotheses for secondary research questions. These results were expected given the lack of literature indicating the impact of such school site and educator demographic variables. The variables were examined, nonetheless, given common-yet-unsubstantiated theories they are of import to data analyses and thus support use and effectiveness.

Conclusions

Most educators have access to data systems to generate and analyze score reports (Aarons, 2009; Herbert, 2011). However, many educators do not use this data correctly, and there is clear evidence many users of data system reports have trouble understanding the data (Wayman et al., 2010; Zwick et al., 2008). Despite this, labeling and tools within data systems to assist analysis are uncommon (USDEOPEPD, 2009). The *Over-the-Counter Data's Impact on Educators' Data Analysis Accuracy* study rendered findings that data system-embedded data analysis support in the forms of footers, abstracts, and interpretation guides all have a significant, positive impact on the accuracy of educators' data analyses.

Findings rendered implications there are direct benefits to educators' data use when a data system and its reports embed at least one of the three data analysis supports investigated in this study. Findings also supported experts' assertions that educators desire more data analysis support from their data systems and its reports, and that the majority of educators use such supports when they are available. Likewise, findings negated literature suggesting the added supports would not be used. In addition, secondary research questions concerning educators' personal and school site demographics were answered with the finding that such demographics have no significant bearing on the supports' success, and thus the supports can be implemented with expected success at varied locations and for varied users.

Given the significant success of footers, abstracts, and interpretation guides, the study warranted related recommendations for three key roles:

- data system and report providers, such as data system vendors and also district staff who maintain in-house data systems, who can embed a footer, abstract, and interpretation guide for every report in the data system;
- educators who use data systems and reports, who can argue for these supports in their current and future data systems; and
- the education research community, who can further research in determining how best to provide added "over-the-counter" data analysis support to educators.

Likewise, the education research community is encouraged to explore best practices for other over-the-counter data aspects such as non-footer aspects of report labeling, the data system's help system, report packaging and data display, and report contents in order to inform better data systems and reports that provide optimal support for educators' data analyses.

Study findings fill a void in education field literature by containing evidence that can be used to identify:

- whether data systems can help increase data analysis accuracy by providing analysis support within data systems and their reports, with the finding being that they can.
- three specific data system/report-embedded supports that increase educators' data analysis accuracy.
- the specific degree to which these supports increase educators' data analysis accuracy (*Figure 8*).
- how likely educators are to use each support.

- examples showing what effective footers, abstracts, and interpretation guides look like (*Figures 1-7*).
- whether minor modifications in support format, mainly in terms of length and color usage, impacted educators' data analysis accuracy, with the findings being that differences in data analysis accuracy were insignificant.

Significance

To offer over-the-counter medication without evidence-based textual guidance would be negligent (DeWalt, 2010). Nonetheless, data systems display data for educators without sufficient support to use their contents – data – wisely (Coburn, Honig, & Stein, 2009; DQC, 2009, 2011; NFES, 2011). Labeling and tools within data systems to assist analysis remain uncommon (USDEOPEPD, 2009). Thus educators' primary option for data use is not typically presented in an "over-the-counter" format, and educators using unmarked or marginally marked data to treat students can be imagined as akin to ingesting medicine from an unmarked or marginally marked container. This analogy is appropriate considering many – and some studies indicate most – educators are making flawed data analyses when using data to impact students.

Despite the common use of data systems to generate reports, research on aspects of report format and system support that could enhance analysis accuracy had been scarce (Goodman & Hambleton, 2004). Literature that did examine data system and report format, including how effectively the format communicates data to users, focused on participants' preferences and *perceived* value of supports. However, user preference can be the opposite of the report format that actually renders more accurate interpretations (Hattie, 2010).

This study examined how effective varied analysis supports are in improving data analysis accuracy, and it did not rely on participants' preferences or perceived value of supports. The findings of this study fill a void in education field literature by containing evidence that can be used to identify whether, how, and to what extent data systems can help increase data analysis accuracy by providing analysis support within data systems and their reports.

Free abstract and interpretation guide templates, based on the formats proven effective in this study, can be downloaded at the researcher's personal website (see cover page). Data system and report providers, such as data system vendors and also district staff who maintain in-house data systems, can use these findings and templates to provide a footer, abstract, and interpretation guide for every report in the data system. Educators can use this study's findings to argue for these supports in their current and future data systems. The education research community can use these findings to support further research in determining how best to provide added "over-the-counter" data analysis support to educators.

Improvements data system and report providers make in light of this study have the potential to improve the accuracy with which educators analyze the data generated by their data systems. Considering only 48% accuracy in data interpretation was found at districts considered exemplars of data use (USDEOPEPD, 2009, 2011), and educators are already enlisting PD and staff supports within their control to improve data use (Strizek, Pittsonberger, Riordan, Lyter, & Orlofsky, 2006), these data system enhancements are especially warranted. More accurate data analyses will likely result in more accurate data-informed decision-making for the benefit of students. It is the strong conviction of this researcher that students deserve for stakeholders to apply *all* possible supports for improved data analysis accuracy in an effort to significantly reduce analysis errors that impact students' lives.

References

Aarons, D. (2009). Report finds states on course to build pupil-data systems. *Education Week*, *29*(13), 6. Retrieved from http://search.proquest.com/docview/202710770?accountid=28180

Bennett, R. E., & Gitomer, D. H. (2009). Transforming K-12 assessment: Integrating accountability testing, formative assessment and professional support. In C. Wyatt-Smith & J. J. Cumming (Eds.), *Educational assessment in the 21st century*, 43-61. New York, NY: Springer.

Brown-Brumfield, D., & DeLeon, A. (2010). Adherence to a medication safety protocol: Current practice for labeling medications and solutions on the sterile field. *Association of Operating Room Nurses*. AORN Journal, 91(5), 610-610-7. doi:10.1016/j.aorn.2010.03.002

Coburn, C. E., Honig, M. I., & Stein, M. K. (2009). What's the evidence on districts' use of evidence? In J. Bransford, D. J. Stipek, N. J. Vye, L. Gomez, & D. Lam (Eds.), *The role of research in educational improvement*, 67-88. Cambridge, MA: Harvard Education Press.

Data Quality Campaign (2009). *The next step: Using longitudinal data systems to improve student success*. Retrieved from http://www.dataqualitycampaign.org/find-resources/the-next-step/

Data Quality Campaign (2011). *Leveraging the power of state longitudinal data systems: Building capacity to turn data into useful information*. Retrieved from http://www.dataqualitycampaign.org/files/DQC-Research%20capacity%20May17.pdf

DeWalt, D. A. (2010). Ensuring safe and effective use of medication and health care: perfecting the dismount. *The Journal of the American Medical Association (JAMA), 304(23),* 2641-2642. doi: 10.1001/jama.2010.1844

Goodman, D. P., & Hambleton, R. K. (2004). Student test score reports and interpretive guides: Review of current practices and suggestions for future research. *Applied Measurement in Education, 17*(2), 145–220.

Hampton, T. (2007). Groups urge warning label for medical devices containing toxic chemical. *The Journal of the American Medical Association (JAMA), 298(11),* 1267. doi: 10.1001/jama.298.11.1267

Hattie, J. (2010). Visibly learning from reports: The validity of score reports. *Online Educational Research Journal*. Also: Paper presented at the annual meeting of the National Council for Measurement in Education (NCME), San Diego, CA. Retrieved from http://www.oerj.org/View?action=viewPaper&paper=6

Herbert, M. (2011). States on track to have top-notch data systems. *District Administration, 47(4),* 12. Norwalk, CT: Professional Media Group LLC.

Kahneman, D. (2003). Maps of bounded rationality: Psychology for behavioral economics. *The American Economic Review, 93(5),* 1449-1475.

Kahneman, D. (2011). *Thinking, fast and slow*. New York, NY: Farrar, Straus and Giroux.

Leeson, H. V. (2006). The mode effect: A literature review of human and technological issues in computerized testing. *International Journal of Testing, 6(1),* 1-24

Lukin, L. E., Bandalos, D. L., Eckhout, T. J., & Mickelson, K. (2004). Facilitating the development of assessment literacy. *Educational Measurement: Issues and Practice, 23(2),* 26–32.

Lyrén, P. (2009). Reporting subscores from college admission tests. *Practical Assessment, Research & Evaluation, 14(4)*, 3-12. Retrieved from http://pareonline.net/pdf/v14n4.pdf

McLaughlin, M., & Talbert, J. E. (2006). *Building school-based teacher learning communities: Professional strategies to improve student achievement*. New York, NY: Teachers College Press.

Minnici, A., & Hill, D. D. (May 9, 2007). *Educational architects: Do state education agencies have the tools necessary to implement NCLB?* Washington, D.C.: Center on Education Policy.

National Forum on Education Statistics. (2011). *Traveling through time: The forum guide to longitudinal data systems. Book Four of Four: Advanced LDS Usage* (NFES 2011–802). Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education.

Odendahl, N. V. (2011). *Testwise: Understanding educational assessment, Volume 1*. Lanham, MD: Rowman & Littlefield Education.

Qin, Y., Wu, M., Pan, X., Xiang, Q., Huang, J., Gu, Z., & ... Zhou, M. (2011, February 25). Reactions of Chinese adults to warning labels on cigarette packages: a survey in Jiangsu Province. *BMC Public Health, 11*(133). doi: 10.1186/1471-2458-11-133

Rodriguez, M. (2008, May 01). Learning how to incorporate technology in classrooms. *Inland Valley Daily Bulletin*, pp. n/a.

Sabbah, F. M. (2011). Designing more effective accountability report cards. *ProQuest Dissertations and Theses, AAT 3469488,* Retrieved from http://search.proquest.com/docview/893068662?accountid=28180

Sanchez, E.; Kline, D.; & Laird, E. (April, 2009). Data-driven districts: Building the culture and capacity to improve student achievement. *Data Quality Campaign: Using data to improve student achievement*. Retrieved from http://www.dataqualitycampaign.org/files/DQCbrief_FINAL-lowres_2_.pdf

Strizek, G.A., Pittsonberger, J.L., Riordan, K.E., Lyter, D.M., & Orlofsky, G.F. (2006). Characteristics of Schools, Districts, Teachers, Principals, and School Libraries in the United States: 2003-04 Schools and Staffing Survey (NCES 2006-313 Revised). *U.S. Department of Education, National Center for Education Statistics*. Washington, DC: U.S. Government Printing Office.

Thaler, R. H. & Sunstein, C. R. (2008). *Nudge: Improving decisions about health, wealth, and happiness*. New Haven, CT: Yale University Press.

Underwood, J. S., Zapata-Rivera, D., & VanWinkle, W. (2008) Growing Pains: Teachers Using and Learning to Use IDMS®. *ETS Research Memorandum. RM-08-07.* Princeton, NJ: ETS.

U.S. Department of Education Office of Planning, Evaluation and Policy Development (2009). *Implementing data-informed decision making in schools: Teacher access, supports and use*. United States Department of Education (ERIC Document Reproduction Service No. ED504191)

U.S. Department of Education Office of Planning, Evaluation and Policy Development (2011). *Teachers' ability to use data to inform instruction: Challenges and supports*. United States Department of Education (ERIC Document Reproduction Service No. ED516494)

Wayman, J. C., Cho, V., & Shaw, S. M. (2009, December). *First-year results from an efficacy study of the Acuity data system.* Paper presented at the Twenty-fourth Annual Texas Assessment Conference, Austin, TX.

Wayman, J. C., Snodgrass Rangel, V. W., Jimerson, J. B., & Cho, V. (2010). *Improving data use in NISD: Becoming a data-informed district*. Austin, TX: The University of Texas at Austin.

Zapata-Rivera, D., & VanWinkle, W. (2010). A research-based approach to designing and evaluating score reports for teachers (*ETS Research Memorandum No. RM-10-01*). Princeton, NJ: ETS.

Zwick, R., Sklar, J., Wakefield, G., Hamilton, C., Norman, A., & Folsom, D. (Summer, 2008). Instructional tools in educational measurement and statistics (ITEMS) for school personnel: Evaluation of three web-based training modules. *Educational Measurement: Issues and Practice, 27*, 14–27.