The Link Between Graphic Design and Report Use

Stephanie Evergreen[1], Jennifer Lyons[1], Julia Rollison[2]

[1]Evergreen Data
[2]Atlas Research

ABSTRACT

Background: Connections between graphic design and use of evidence have not been well-studied, particularly in policymaking settings. The aim of this study was to conduct an exploratory study of the relationship between research use and graphic design in a policymaking setting.

Methods: Using data from an initial study assessing citation data in U.S. Congressional hearings on teacher quality, analysts pulled a random sample of 88 reports and evaluated graphic design use with an Evaluation Report Layout Checklist. We conducted correlation analyses between frequency of use and individual checklist items, sub-section scores, and overall checklist scores.

Findings: Only three checklist items were statistically significant for partially meeting the graphic design criteria. No significant relationship was found between frequency of use and overall checklist scores.

Discussion: Decision-making and use of findings in a policymaking context is extremely complex and there may be mediating and moderating variables impacting perceptions of credibility and use, particularly in the context of U.S. Congressional hearings. Future research should consider measuring and including other potential factors like power and networks to better account for the dynamic nature of policy making.

Conclusion: This study showed little association between the extent to which a report implements basic graphic design principles and use in a real policy decision-making scenario and did not fit with the overarching theory of change behind many of the guidelines for better reporting practices. There is a pressing need for additional research into the impact the presentation of findings has on how evidence is used in policymaking settings.

Background

Policymakers use research-based evidence to define problems, shape stories, and ideally to drive policy directions and decisions (Stone, 2009). While the presentation of this evidence can take many forms, traditionally, the most common format is the written word (Mathison, 2009). The different forms of presentation generally depend on the audience to influence perceptions of the credibility of that evidence (Mathison, 2009). In fact, non-text-based formats, such as graphics, are often viewed as more accessible and powerful methods of communicating research, particularly when trying to persuade an audience to adopt a belief or action, such as in the context of policymaking (Mathison, 2009; Tufte, 2001). Researchers have shown that the addition of graphs, even trivial ones, increases the credibility of evidence because readers associate data visualization with appearing scientific (Tal & Wansink, 2016). Visualizing data is becoming an important part of the knowledge translation process (Petch, Lightowler, Pattoni, & Watson, 2014), which

Evergreen, S., Lyons, J. & Rollison, J. (2018). The link between graphic design and report use. Paper presented at the American Evaluation Association annual conference, Atlanta, GA, 26 October, 2016.

occurs in fields as broad as philanthropy, education, and policymaking.

The concept of applying graphical design strategies to statistical data was initially developed by William Playfair in the late 1800s and later popularized by Edward Tufte beginning in the 1980s in his publications on the visual presentation of information and evidence (Tufte, 2001). Since then, the fields of graphic design and data visualization have continued to rapidly grow and expand as a device for communicating evidence in various settings (Meeks, 2017; Miller & Hughes, 2017; Patil & Brynjolfsson, 2014; Smith, 2016). The idea of more accessible and powerful methods of communicating are especially relevant to policymakers who often have little time to digest long written reports or publications (Choi, 2005). Further, it is assumed increased accessibility of information may increase the impact and value for policymakers (Sullivan, Wells, & Coyle, 2015).

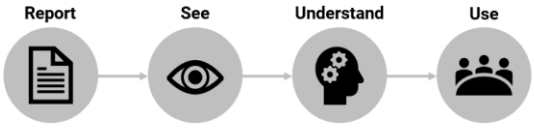
Guidance on the presentation of evidence purports to emphasize consistency in terms of clarity, comprehensiveness, and transparency, to help stakeholders like policymakers critically evaluate and interpret findings (Moher et al., 2010; Sullivan, Wells, & Coyle, 2015). Common guidance documents or checklists include the Consolidated Standards of Reporting Trials (CONSORT) statement (Schulz, Altman, & Moher, 2010) and the Standards for Reporting on Empirical Social Science Research (American Educational Research Association (AERA), 2006). However, neither of these commonly used tools call out specifics for graphic design beyond the need to ensure graphical displays protect anonymity (AERA, 2006).

There is an inherent logic in the concept that more visually appealing reporting of evidence will lead to increased actual use of the reports for some type of decision-making or action. This is particularly salient in scenarios where research and analytic studies are intended to influence policy, such as Federally-funded research or other social science studies (Davies & Powell, 2012; Murdock, Shariff, & Wilding, 2013; Dagenais et al., 2012). The gap between design and use, however, is large, with many intermediary steps. This gap is an important one to study and understand because otherwise, researchers and practitioners are spending a lot of time and energy to improve reporting evidence from studies based on assumptions. In some ways, existing advice on report design and visualization is misleading because it presumes use is the natural outcome without any evidence that such is the case.

Theory of Change

The movement toward better visual reporting is based on a theory of change about human cognition and action (see Figure 1) (Ware, 2008). The first step in the theory of change is that people can actually see the reporting. According to the theory, that should not be too difficult, because vision is the dominant sense of information among sighted individuals; it is our primary method of getting information (Stenberg, 2006). But in practice report authors do all kinds of things with text, colors, layout, and graph choice that get in the way of actually seeing.

Figure 1. Theory of change behind presenting data effectively.



Choosing the right fonts and colors can impair or support legibility. Serif fonts make paper-based reading easier, while sans serif fonts are better suited for the screen (Wheildon, 2005). Color contrast can also support or hinder reading. The combination of certain colors, such as red and green, also affects legibility for those who are colorblind (Wheildon, 2005). Researchers have

even studied which colors evoke what emotions (Gilbert, Fridlund, & Lucchina, 2016; Kuhbandner & Pekrun, 2013; Sutton & Altarriba, 2016), and the personalities carried by different fonts (Brumberger, 2003; Mackiewicz & Moeller, 2004).

In other words, we have some information to suggest that practitioners can take actions to make it more likely that readers will be able to see their reports. When done well, graphic design strategies like color, alignment, motion, orientation, and size will grab attention (Davies & Powell, 2012). Visual cognition researchers have shown repeatedly that grabbing initial attention boosts recall of information (Ware, 2008).

The next stage of change, "think," emphasizes the link between seeing and remembering. If users can comprehend the evidence, they are more likely to store it away in long term memory and take action, which is the ultimate goal in this theory of change and of most presentations of evidence (Evergreen, 2013).

Data visualization guidance focuses heavily on the middle of the theory of change: whether people can then comprehend what it says (Mason, 2014). Before information gets to the brain's long-term memory, it must pass through is working memory which is where we wrestle with information to process and understand it. But working memory is weak; research shows humans can only hold 3-5 chunks of information in working memory at any one time and that amount even varies based on the environment around us when we are trying to think (Cowan, 2000). However, graphic elements can reduce the cognitive load by doing some of the thinking for the reader. By visually organizing and emphasizing information, graphic design makes it more accessible for the reader, increasing engagement and comprehension.

Color impacts comprehension, as well as legibility. Researchers have supplied some evidence that the random assignment of colors, or even the use of a rainbow color scale, can cause confusion (So & Smith, 2002). Shades of a single color can increase accuracy of interpretation (Breslow, Ratwani, & Trafton, 2009). Using an action color on a key piece of the reporting will also boost recall later (Ware, 2008).

Well-structured titles for figures and graphs can also lead to increased comprehension. Borkin et al.'s (2016) study used eye-tracking to determine that the titles for data visualizations were attractive (the previous step in the theory of change) but that strong titles that tell an entire take away point from the data were better retained and recalled days after seeing the visual, indicating increased comprehension.

Various research studies have investigated ideal graph types. Dot plots are a versatile graph type for displaying discrete, comparison, and survey data. They present information in a way that overcomes some of the problems frequently encountered with other graphical displays (Jacoby, 2006). Bar charts are better at displaying comparative values and segment-to-segment judgements compared to pie charts and line graphs (Cleveland & McGill, 1984, Heer & Bostock, 2010, Kosara & Skau, 2016). When graphing time series data, line graph variations are the most common display and the easy for the brain to interpret (Zacks & Tversky, 1999). When graphs are displayed on tabletop and vertical surfaces, the perception of some graphical elements are more open to distortion than others (Wigdor, Shen, Forlines, & Balakrishnan, 2007).

Icon arrays, a set of identical icons with color applied to a portion of them, have been shown as effective interpretive supports for people with low numeracy skills (Galesic, Garcia-Retamero, & Gigerenzer, 2009). Some research has shown that simply displaying a number in a large font size is better for comprehension than the same data in pie charts, icon arrays, and stacked bars (Zikmund-Fisher, 2014).

Researchers have studied whether the addition of cartoon-like illustrations to a graph can garner more attention on the visual. Borkin et al (2013), Bateman et al (2010), and Haroz, Kasara, & Franconeri (2015) showed that chart embellishments can get attention as well as plain graphs and even produce better recall weeks after seeing the initial visual. However, research on PowerPoint has shown that extraneous illustrations distract the audience from the content being conveyed (Berk, 2011). Long-term recall, the true test of comprehension, is the next step in the theory of change, leading toward use. Again, here the research provides some evidence that graphic design and data visualization have a role. Graphics are particularly good at activating existing schemas, a key part of long-term recall. Mayer (2009) showed that when information is delivered only verbally, about 10 percent of the information is retained after about 48 hours. When visuals and graphics are added to the explanation, retention increases to 75 percent. Adding graphic design and visualization help better tap in to mental models so that new information can be more easily assimilated.

The growing body of research related to legibility and comprehension of design and data visualization suggests that practitioners, researchers, and policy analysts are becoming more self-aware of their visual strengths and shortcomings in the first few steps of the theory of change. However, there is limited research on the end of the theory of change: linking design to actual use of the better-designed evidence.

Improved Presentation of Evidence and Use

While the leap in logic that takes us from "appealing reports" to "actual use" feels like it makes sense, this area has not been very-well studied. Certainly, organizations funding studies have wondered about the use of their deliverables like study reports. The USAID, a U.S. Government agency focused on addressing global poverty, systematically asked their clients about the extent to which the deliverables were used (Hageboeck et al., 2016) but the study relies on self-report of use which can be unreliable. In addition, the visual nature of the report was not taken into account when measuring report quality. An additional recent study compared a cluttered diagram with a redesigned diagram that made better use of color and space. Their study population, random users of Mechanical Turk, rated the improved diagram as better on

Practitioner-scholars have been writing about the need to increase the use of research and data through a focus on user-friendly reporting (Davidson, 2007; Patton, 2011) Typically, discussion of use focuses on types of use (e.g., conceptual, process) and factors affecting use (e.g., relevance, timing), but graphic design is notably absent from these discussions (Weiss, 1998; Evergreen, 2011). Prominent authors in the field of data visualization and design commonly suggest that increased use is a natural outcome of good design (Duarte, 2012; Few, 2013; Knaflic, 2015; Schwabish, 2014) but do not cite any studies to support this assessment.

Further, report formatting and graphic design have been largely left out of the variable list when studying obstacles to use of findings. Authors usually restrict their discussion of use to knowing one's audience and tailoring report formats (i.e., brochures, newsletters, oral presentations) (Lawrenz, Gullickson, & Toal, 1997; Rossi, Freeman, & Lipsey, 2003; Russ-Eft, Atwood, & Egherman, 2002). Some use-of-evidence-oriented texts acknowledge the role of graphic design in reporting, but give it a cursory address, such as suggesting that one hire a graphic designer, or use white space (Patton, 2011; Stetson, 2008). Only a few texts have attempted to give guidance on graphic design, like providing direction on how to create charts or structure a report and they have become best sellers (Evergreen, 2016; Evergreen, 2017; Sue & Griffin, 2016). But even these texts only presume use will be the natural outcome of good design.

aesthetics and credibility. Response times were faster and responses were more accurate among those who viewed the improved diagram (Paige et al, 2017). But none of these metrics measure actual use.

In this paper, we present research on the impact of graphic design and data visualization concepts on actual use of evidence in a policy-making context (i.e., recorded congressional hearings). Rather than focusing on the specific subject matter, (in this case, the congressional hearings were related to teacher quality), we focused the analysis on how policy decisions are informed in an attempt to illuminate connections between evidence presentation and policy. The study approach allowed us to work backwards to examine the design of the reports that informed Congressional decisions without report authors knowing we were studying their products as well as ensure no biased self-report in calculations of report use (Burstein, 2014). We aimed to learn whether the theory of change around human cognition and action based on improved visual reporting of evidence is related to policy-related decision making.

Methods

Reckhow et al. (2015) identified a total of 197 Congressional hearings on teacher quality from 2000 to 2014. In those hearings, various parties attempt to influence the decision on the table by making their strongest arguments, backed by research, reports, and citations. The researchers in this initial study (Reckhow et al., 2015) identified every citation in the 197 hearings on teacher quality downloaded from the Government Printing Office from 2000 to 2014. This resulted in references to exactly 600 separate published research items, including academic articles, think tank reports, and government reports.

Working within this specific population of 600 cited reports and without seeing the data related to frequency of report usage, we launched an independent study. We used a random sample of 88 reports (for a 10% confidence interval). We selected all reports that were cited more than once (2% of all citations, or 10 reports) and randomly sampled from those cited only once until we had 88 reports. Then we evaluated their graphic design use with the Evaluation Report Layout Checklist.

The Evaluation Report Layout Checklist is the product of the first author's dissertation, which pulled from the best available research at the time – 2009 and 2010 – on how a report should be laid out and formatted to maximize legibility and comprehension (again, on the assumption that report use would naturally follow as the long-term outcome). The Checklist has been heavily vetted by subject matter experts in the field. It has been used by multiple raters on the same report with significant agreement (Evergreen, 2011). The Checklist provides a set of guidelines for graphic design of reports, divided into 4 main sections: Type, Arrangement, Graphics, and Color. Each section has several checkpoints, such as "Color reprints legibly in black and white" and "Body text is left or full justified." These research-backed guidelines help users format their reporting. Users are prompted to mark whether the report in question Fully Met, Partially Met, or Didn't Meet each guideline (coded as 2, 1, and 0, respectively). The full checklist is included in the appendix.

Two study authors calibrated our own scoring on a pilot sample of reports until we had 68 percent agreement (Krippendorf's alpha = .68), which is considered acceptable (Landis & Koch, 1977). We then recruited and trained seven volunteers who each spot-checked a selection of our scores against the original reports. We adjusted the final scores after discussions with spot checkers led to consensus.

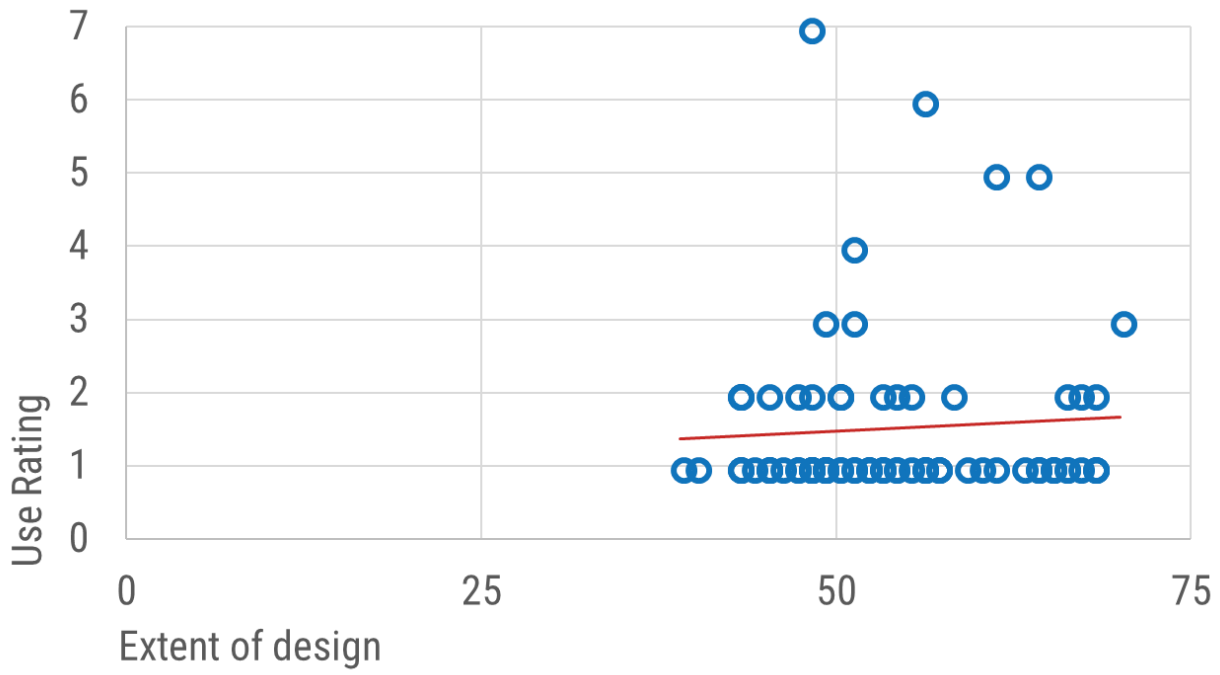
Prior to the analysis, we conducted an outlier analysis which led to the exclusion of one case (a

report cited 27 times), resulting in a total sample size of 87. We then matched up the graphic design ratings from the Checklist with the original data related to frequency of use (i.e., the number of times the report was cited in Congressional testimony) and conducted a correlation analysis. We analyzed each checklist item individually, using one-way ANOVA and a Tukey post hoc. In addition, we disaggregated the overall checklist score into subscores for each of the four main sections – Type, Arrangement, Graphics, and Color – and analyzed the subscore correlation against the number of times the report was used. We controlled for length of the publication and source (i.e., think tank, academic institution, other). Variables including funding level of report authors, extent of advocacy focus, and use of references were not investigated in this study as prior analyses found no relationship with subsequent use (Reckhow & Galey, 2017; Reckhow, Holden, Tompkins-Stange, 2015).

Findings

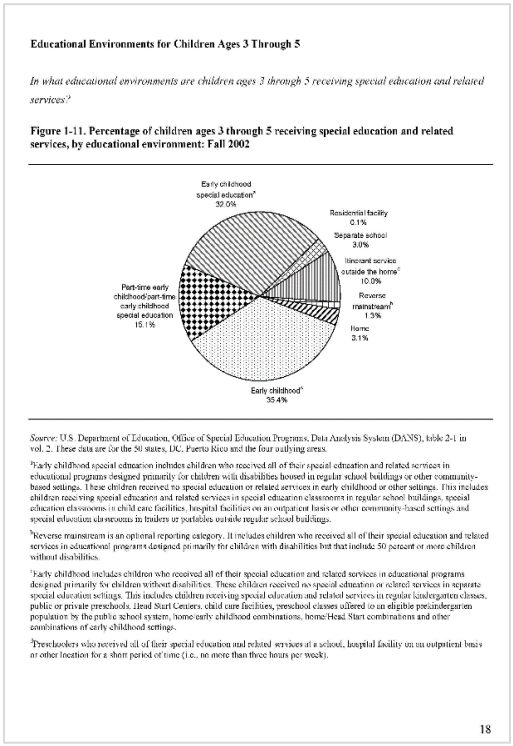
The extent to which the reports used graphic design ranged from 37 to 70, out of a possible 75 points, based on our scoring with the Checklist. The median score was 51. Each circle in Figure 2 is a report included in this study. The correlation is not statistically significant (r=0.077, n=88, p=0.48).

Figure 2. Extent of report graphic design and frequency of same report's subsequent use show little correlation.



Most of the reports (72 percent) in our sample were only cited in those Congressional hearings once. They tended to look like the report in Figure 3, with pattern-filled pie charts, an example that does not qualify as strong graphic design according to the Checklist. Others used colors but contained unclear diagrams (see example in Figure 4).

Figure 3. Report page reprinted from U.S. Department of Education, Office of Special Education and Rehabilitative Services, Office of Special Education Programs. (2004) 26th Annual Report to Congress on the Implementation of the Individuals with Disabilities Education Act, Vol. 1, Washington, D.C.



Educational Environments for Children Ages 3 Through 5

In what educational environments are children ages 3 through 5 receiving special education and related services?

Figure 1-11. Percentage of children ages 3 through 5 receiving special education and related services, by educational environment: Fall 2002

Source: U.S. Department of Education, Office of Special Education Programs, Data Analysis System (DANS), table 2-1 in vol. 2. These data are for the 50 states, DC, Puerto Rico and the four outlying areas.

[2]Early childhood special education includes children who received all of their special education and related services in educational programs designed primarily for children with disabilities housed in regular school buildings or other community-based settings. These children received no special education or related services in early childhood or other settings. This includes children receiving special education and related services in special education classrooms in regular school buildings, special education classrooms in child care facilities, hospital facilities on an outpatient basis or other community-based settings and special education classrooms in trailers or portables outside regular school buildings.

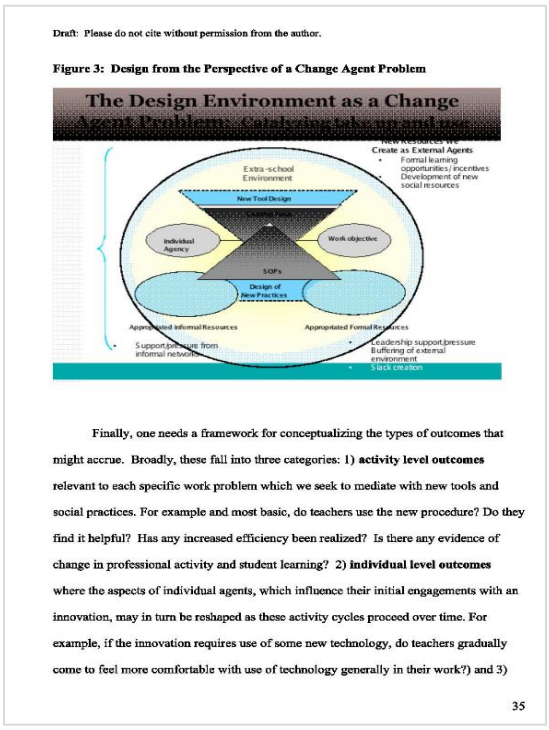
[5]Reverse mainstream is an optional reporting category. It includes children who received all of their special education and related services in educational programs designed primarily for children with disabilities but that include 50 percent or more children without disabilities.

[3]Early childhood includes children who received all of their special education and related services in educational programs designed primarily for children without disabilities. These children received no special education or related services in separate special education settings. This includes children receiving special education and related services in regular kindergarten classes, public or private preschools, Head Start Centers, child care facilities, preschool classes offered to an eligible prekindergarten population by the public school system, home/early childhood combinations, home/Head Start combinations and other combinations of early childhood settings.

[4]Preschoolers who received all of their special education and related services at a school, hospital facility on an outpatient basis or other location for a short period of time (i.e., no more than three hours per week).

18

Figure 4. Report page reprinted from Bryk, A. S., & Gomez, L. M. (2008). Ruminations on reinventing an R&D capacity for educational improvement. The future of educational entrepreneurship: Possibilities of school reform, 181-206.



Draft: Please do not cite without permission from the author.

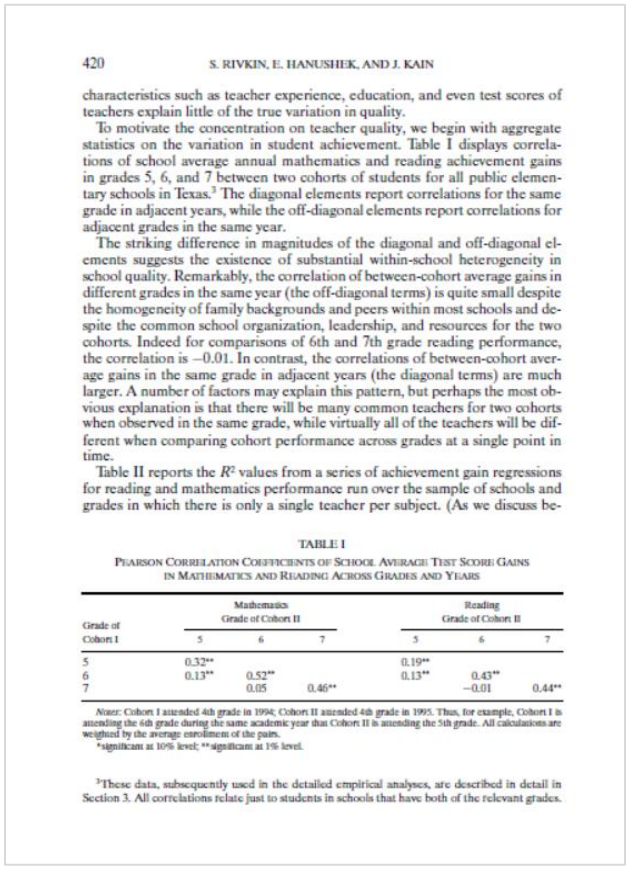
Figure 3: Design from the Perspective of a Change Agent Problem

Finally, one needs a framework for conceptualizing the types of outcomes that might accrue. Broadly, these fall into three categories: 1) activity level outcomes relevant to each specific work problem which we seek to mediate with new tools and social practices. For example and most basic, do teachers use the new procedure? Do they find it helpful? Has any increased efficiency been realized? Is there any evidence of change in professional activity and student learning? 2) individual level outcomes where the aspects of individual agents, which influence their initial engagements with an innovation, may in turn be reshaped as these activity cycles proceed over time. For example, if the innovation requires use of some new technology, do teachers gradually come to feel more comfortable with use of technology generally in their work?) and 3)

35

Eighteen percent (n=108) were used by the study subjects two times and ten percent of the reports in our sample (n=8) were cited in those hearings more than twice. For example, some were similar to a peer-reviewed journal article (see Figure 5 which scored 48) incorporating minimal graphic design. Others, like the sample in Figure 6 (a report that scored 52), included somewhat decent charts, columns, and stand out headings. At the higher end of the scale, some like the sample in Figure 7, contained more of a magazine appearance and scored a 64.

Figure 5. Report page reprinted from Rivkin, S. G., Hanushek, E. A., & Kain, J. F. (2005). Teachers, schools, and academic achievement. Econometrica, 73(2), 417-458.



420 S. RIVKIN, E. HANUSHEK, AND J. KAIN

characteristics such as teacher experience, education, and even test scores of teachers explain little of the true variation in quality.

To motivate the concentration on teacher quality, we begin with aggregate statistics on the variation in student achievement. Table I displays correlations of school average annual mathematics and reading achievement gains in grades 5, 6, and 7 between two cohorts of students for all public elementary schools in Texas.[3] The diagonal elements report correlations for the same grade in adjacent years, while the off-diagonal elements report correlations for adjacent grades in the same year.

The striking difference in magnitudes of the diagonal and off-diagonal elements suggests the existence of substantial within-school heterogeneity in school quality. Remarkably, the correlation of between-cohort average gains in different grades in the same year (the off-diagonal terms) is quite small despite the homogeneity of family backgrounds and peers within most schools and despite the common school organization, leadership, and resources for the two cohorts. Indeed for comparisons of 6th and 7th grade reading performance, the correlation is −0.01. In contrast, the correlations of between-cohort average gains in the same grade in adjacent years (the diagonal terms) are much larger. A number of factors may explain this pattern, but perhaps the most obvious explanation is that there will be many common teachers for two cohorts when observed in the same grade, while virtually all of the teachers will be different when comparing cohort performance across grades at a single point in time.

Table II reports the R^2 values from a series of achievement gain regressions for reading and mathematics performance run over the sample of schools and grades in which there is only a single teacher per subject. (As we discuss be-

TABLE I

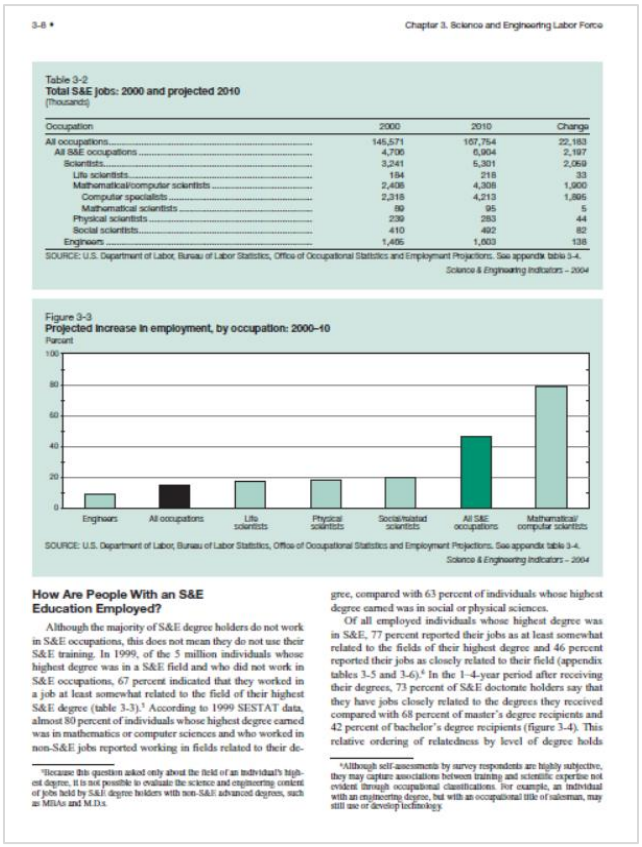
PEARSON CORRELATION COEFFICIENTS OF SCHOOL AVERAGE TEST SCORE GAINS IN MATHEMATICS AND READING ACROSS GRADES AND YEARS

Grade of Cohort I	Mathematics Grade of Cohort II			Reading Grade of Cohort II		
	5	6	7	5	6	7
5	0.32**			0.19**		
6	0.13**	0.52**		0.13**	0.43**	
7		0.05	0.46**		−0.01	0.44**

Note: Cohort I attended 4th grade in 1994; Cohort II attended 4th grade in 1995. Thus, for example, Cohort I is attending the 6th grade during the same academic year than Cohort II is attending the 5th grade. All calculations are weighted by the average enrollment of the pairs.
*significant at 10% level; **significant at 1% level.

[3]These data, subsequently used in the detailed empirical analyses, are described in detail in Section 3. All correlations relate just to students in schools that have both of the relevant grades.

Figure 6. Report page reprinted from Science and Engineering Indicators 2004: Chapter 3 Science and Engineering Labor Force.



3-8 ▪ Chapter 3. Science and Engineering Labor Force

Table 3-2
Total S&E jobs: 2000 and projected 2010
(Thousands)

Occupation	2000	2010	Change
All occupations	145,571	167,754	22,183
All S&E occupations	4,706	6,904	2,197
Scientists	3,241	5,301	2,059
Life scientists	194	218	33
Mathematical/computer scientists	2,408	4,308	1,900
Computer specialists	2,318	4,213	1,895
Mathematical scientists	89	95	5
Physical scientists	239	283	44
Social scientists	410	492	82
Engineers	1,465	1,603	138

SOURCE: U.S. Department of Labor, Bureau of Labor Statistics, Office of Occupational Statistics and Employment Projections. See appendix table 3-4.
Science & Engineering Indicators – 2004

Figure 3-3
Projected increase in employment, by occupation: 2000–10
Percent

SOURCE: U.S. Department of Labor, Bureau of Labor Statistics, Office of Occupational Statistics and Employment Projections. See appendix table 3-4.
Science & Engineering Indicators – 2004

How Are People With an S&E Education Employed?

Although the majority of S&E degree holders do not work in S&E occupations, this does not mean they do not use their S&E training. In 1999, of the 5 million individuals whose highest degree was in a S&E field and who did not work in S&E occupations, 67 percent indicated that they worked in a job at least somewhat related to the field of their highest S&E degree (table 3-3).[1] According to 1999 SESTAT data, almost 80 percent of individuals whose highest degree earned was in mathematics or computer sciences and who worked in non-S&E jobs reported working in fields related to their degree, compared with 63 percent of individuals whose highest degree earned was in social or physical sciences.

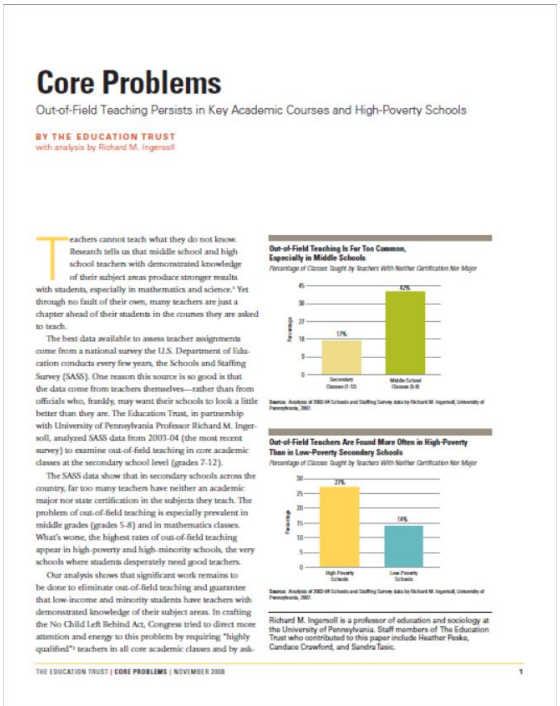
Of all employed individuals whose highest degree was in S&E, 77 percent reported their jobs as at least somewhat related to the fields of their highest degree and 46 percent reported their jobs as closely related to their field (appendix tables 3-5 and 3-6).[4] In the 1–4-year period after receiving their degrees, 73 percent of S&E doctorate holders say that they have jobs closely related to the degrees they received compared with 68 percent of master's degree recipients and 42 percent of bachelor's degree recipients (figure 3-4). This relative ordering of relatedness by level of degree holds

[1]Because this question asked only about the field of an individual's highest degree, it is not possible to evaluate the science and engineering content of jobs held by S&E degree holders with non-S&E advanced degrees, such as MBAs and M.D.s.

[4]Although self-assessments by survey respondents are highly subjective, they may capture associations between training and scientific expertise not evident through occupational classifications. For example, an individual with an engineering degree, but with an occupational title of salesman, may still use or develop technology.

Evergreen, S., Lyons, J. & Rollison, J. (2018). The link between graphic design and report use. Paper presented at the American Evaluation Association annual conference, Atlanta, GA, 26 October, 2016.

Figure 7. Report page reprinted from Ingersoll, R. (2008). Out-of-Field Teaching Persists in Key Academic Courses and High-Poverty Schools. Education Trust.



Core Problems

Out-of-Field Teaching Persists in Key Academic Courses and High-Poverty Schools

BY THE EDUCATION TRUST
with analysis by Richard M. Ingersoll

Teachers cannot teach what they do not know. Research tells us that middle school and high school teachers with demonstrated knowledge of their subject areas produce stronger results with students, especially in mathematics and science.* Yet through no fault of their own, many teachers are just a chapter ahead of their students in the courses they are asked to teach.

The best data available to assess teacher assignments come from a national survey the U.S. Department of Education conducts every few years, the Schools and Staffing Survey (SASS). One reason this source is so good is that the data come from teachers themselves—rather than from officials who, frankly, may want their schools to look a little better than they are. The Education Trust, in partnership with University of Pennsylvania Professor Richard M. Ingersoll, analyzed SASS data from 2003-04 (the most recent survey) to examine out-of-field teaching in core academic classes at the secondary school level (grades 7-12).

The SASS data show that in secondary schools across the country, far too many teachers have neither an academic major nor state certification in the subjects they teach. The problem of out-of-field teaching is especially prevalent in middle grades (grades 5-8) and in mathematics classes. What's worse, the highest rates of out-of-field teaching appear in high-poverty and high-minority schools, the very schools where students desperately need good teachers.

Our analysis shows that significant work remains to be done to eliminate out-of-field teaching and guarantee that low-income and minority students have teachers with demonstrated knowledge of their subject areas. In crafting the No Child Left Behind Act, Congress tried to direct more attention and energy to this problem by requiring "highly qualified"+ teachers in all core academic classes and by ask-

Out-of-Field Teaching Is Far Too Common, Especially in Middle Schools

Percentage of Classes Taught by Teachers With Neither Certification Nor Major

Source: Analysis of 2003-04 Schools and Staffing Survey data by Richard M. Ingersoll, University of Pennsylvania, 2007.

Out-of-Field Teachers Are Found More Often in High-Poverty Than in Low-Poverty Secondary Schools

Percentage of Classes Taught by Teachers With Neither Certification Nor Major

Source: Analysis of 2003-04 Schools and Staffing Survey data by Richard M. Ingersoll, University of Pennsylvania, 2007.

Richard M. Ingersoll is a professor of education and sociology at the University of Pennsylvania. Staff members of The Education Trust who contributed to this paper include Heather Peske, Candace Crawford, and Sandra Tasic.

THE EDUCATION TRUST | CORE PROBLEMS | NOVEMBER 2008 1

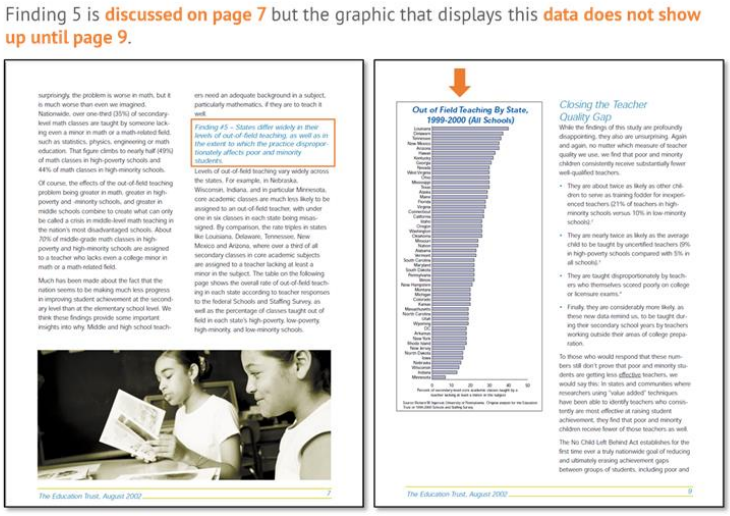
We found no correlation between the extent of graphic design and how often the reports were cited during the decision-making process in those Congressional hearings. We also found no influence of length of publication or publication source. When comparing the subscore correlation for the four main sections of the checklist against the number of times the report was used, we found no significant relationships. Differences emerged, however, on three specific checklist items representing several of the main sections of the checklist:

- Columns are 8-12 words in length (p=.008)
- Graphics are near associated text (p=.000)
- One or two emphasis colors are used (p=.025)

The significant differences were found in the Partially Met rating; the reports that were cited more only Partially Met these three checkpoints. Figure 8 provides an example of a report that would only Partially Meet the checkpoint "Graphic is near associated text" which shows an example where the visual was located right after the associated text. In the same report, text discusses a graphic and it is not shown until two pages later. These demonstrated inconsistencies rationalize the partial score.

Figure 8. Report page reprinted from Jerald, C. D., & Ingersoll, R. (2002). All Talk No Action: Putting and End to Out-of-Field Teaching.

Finding 1 is discussed on page 4 and the graphic is right below the text reference.



Finding 5 is **discussed on page 7** but the graphic that displays this **data does not show up until page 9**.

Evergreen, S., Lyons, J. & Rollison, J. (2018). The link between graphic design and report use. Paper presented at the American Evaluation Association annual conference, Atlanta, GA, 26 October, 2016.

Discussion

Only Partially Meeting some criteria (i.e., columns 8 to 12 words in length; graphics near associated text; and use of one or two emphasis colors) seemed to make a difference in report use. While these may have been spurious relationships, findings on these criteria are supported by the research. For example, comprehension studies show that readers can best track a text (i.e., finish reading one line and start another) when the length of the line is restricted to 8-12 words per line (Morrison, 1930; Samara, 2007; Wheildon, 2005). When the line is too long, readers lose track of which line to start with when finishing one line and returning to the left side of the page to continue reading.

Graphics located near associated text is a checkpoint that may stem more from common practice than research. Pictures or graphic elements should be placed within the foveal range of associated text because human eyesight has only a narrow range of focus (Few, 2006; Malamed, 2009; Tufte, 1990). The common practice of placing tables or graphs in an appendix or separated by pages from their corresponding text means that some information will be lost in the extended time a reader must take to flip back and forth to bring the text and the graphic together into a comprehensive whole.

Research on color supports the checkpoint that only one or two emphasis colors are used. For color to be deployed well, less critical parts of a page or a figure should be in a shade of gray so that chosen elements can appropriately stand out when selected emphasis colors are applied (Jamet, Garota, & Quaireau, 2008). The use of color for emphasis can impede comprehension if too many colors are used indiscriminately; readers expect that a change in color indicates a change in meaning and they will spend time and effort trying to understand the meaning shift (So & Smith, 2002). Using one or two selective colors for highlighting purposes brings attention to the germane parts of the page or figure.

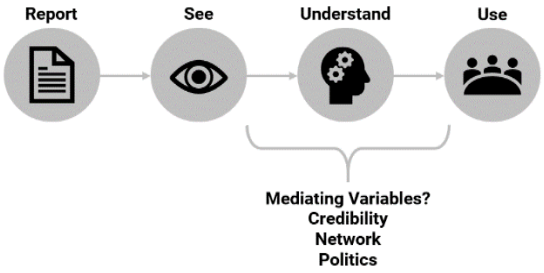
The findings on these Partially Met criteria may suggest that some attempt of incorporating graphic design principles in the presentation of evidence is useful. Additionally, perhaps a report that Fully Meets every checkpoint, and thus is perceived as too glossy, slick, or high-level, may be perceived as lacking credibility which could affect use when it comes to decision making in the policy context.

When examining graphic design by section or across the whole report, there were no significant relationships found between graphic design scores and the measure of use. This could simply indicate no relationship in this Congressional hearing context with graphic design and use, and would perhaps be somewhat unsurprising since there is limited research examining this relationship. However, more likely is that this Congressional hearing setting of use is extremely complex and there may be mediating and moderating variables impacting perceptions of credibility and use in this context.

The hearing recording may have reflected experts who self-cited and not necessarily Members of Congress or their staff. A network analysis by researchers in the initial study revealed “one large cluster of witnesses with many shared citations” (Reckhow, Holden, Tompkins-Stange, 2015, p. 18). Essentially, several of the think tanks continually cross-referenced one another, such that Congressional “witnesses in this cluster seem to exhibit remarkable consistency in the content of their testimony” (p. 18). In other words, actual use of the reports as citations in Congressional hearings may have been mediated by a network of individuals who cite each other, thus falsely inflating the use variable. In addition, it is possible the tight network nodes influenced the findings.

An outlier report, excluded from the analysis, was cited 27 times in Congressional hearings on teacher quality (by comparison, the second most cited report was only cited 7 times). This report was authored by a committee including representatives from large government contractors, pharmaceutical companies, and U.S. Federal agencies focused on intelligence and defense. This suggests other factors, such as power or social processes, may have been more critical to determining the credibility of evidence in the decision-making process than how the information in the report was presented. It may be unsurprising in the context of policy making that these factors heavily influence the final report (Stone, 2009). Smith and Joyce (2012) have suggested that networks, political divides, ideology, lobbying, inertia, and public opinion often outweigh the content of relevant research in policy making contexts. Research use is complex.

Figure 9. Modified theory of change behind presenting data effectively.



The inclusion of additional variables into multivariate analyses on use could help understand potential mediating and moderating relationships at play. While no relationships between use and certain variables (i.e., funding level of report authors; extent of advocacy focus; use of references) were found in previous studies with these same data (Reckhow & Galey, 2017; Reckhow, Holden, Tompkins-Stange, 2015), there are other factors which would be useful to consider measuring and including in future analyses to better account for the dynamic nature of policy making. For example, categorical variables on the relevancy of the content presented, type of Congressional committee, or report type would be especially useful to capture and incorporate into a multivariate analysis of use.

Alternatively, or in addition to, it is possible that a different instrument would better detect changes in report design and show a greater connection to report use in this context. These findings could also indicate that the Checklist is best suited for reports that have essentially no graphic design, the kind that are all black and white, Times New Roman, size 12, double spaced text. Perhaps the checklist overcorrects in this context of use in policymaking, where a modest level of graphic design (i.e., Partially Met) is sufficient.

Further, many of these reports may have been developed based on requirements for Section 508. These requirements help ensure accessibility standards for public reporting but can influence choices on how graphics and other visuals are presented. U.S. federal guidance around the compliance does not offer much regarding graphic design, figures, and data visualizations, so it is possible that report authors were mistakenly choosing weaker designs in an attempt to be compliant.

Strengths and Limitations

This study was the first of its kind to utilize existing data to examine an understudied relationship between graphic design and use in this policymaking setting. The team utilized a research-based checklist to assess implementation of good graphic design principles and incorporated a measure of use that did not rely on self-report, a noted limitation of other studies examining relationships between graphic design and use (Hageboeck et al., 2016; Paige et al, 2017).

This study focuses on a very specific instance of use: use of reports in Congressional hearings in

the United States. Specifically, this study incorporated a review of reports cited in Congressional hearings and developed from policy organizations, think tanks, and sometimes academic institutions. These reports stemmed from different types of funding mechanisms and often had varying intents and purposes. The reliance on citations as a measure of use could have been problematic given the range of factors and nuances involved in these Congressional hearings. In addition, the vast majority of reports were only cited once.

This study excludes common use scenarios where, for example, a school superintendent uses a think tank article on teacher quality to initiate changes at a district level. Decision making happens on a range of levels outside of Congressional hearings, even at the U.S. federal level. Use of research on teacher quality is broader than the specific context we have sampled from in this study.

Conclusion

This study showed little association between the extent to which a report implements basic graphic design principles and use in real policy decision-making scenarios. This finding does not fit with the overarching theory of change that undergirds many of the guidelines for better reporting practices. We know that clear visuals lead to more reader attention, comprehension, and retention and assume that retention will lead to use. However, this study cannot support the link from graphic design to use in the Congressional hearing context.

This study was a critical first step in exploring potential connections between incorporating principles of data visualization and use of findings by individuals informing, and developing, policy. It also provided insights into how to improve future research in this area to account for the complexities that mediate and moderate relationships of evidence and use of that evidence in the policymaking context. Given the rapid expansion of using graphic design and data visualization as devices for communicating evidence (Meeks, 2017; Miller & Hughes, 2017; Patil & Brynjolfsson, 2014; Smith, 2016), there is a clear need for continued research to provide more evidence-based insight into the presentation of findings - beyond content - to help inform and improve the eventual use of those findings.

Evergreen, S., Lyons, J. & Rollison, J. (2018). The link between graphic design and report use. Paper presented at the American Evaluation Association annual conference, Atlanta, GA, 26 October, 2016.

References

American Educational Research Association (AERA) (2006). Evaluation Standards for Reporting on Empirical Social Science Research in AERA Publications. Available at: http://www.sagepub.com/upm-data/13127_Standards_from_AERA.pdf.

Bateman, S., Mandryk, R. L., Gutwin, C., Genest, A., McDine, D., & Brooks, C. (2010). Useful junk? The effects of visual embellishment on comprehension and memorability of charts. CHI 2010.

Berk, R. A. (2011). Research on PowerPoint: From basic features to multimedia. International Journal of Technology in Teaching and Learning, 7(1), 24-35.

Borkin, M. A., Vo, A. A., Bylinskii, Z., Isola, P., Sunkavalli, S., Oliva, A., & Pfister, H. (2013). What makes a visualization memorable?. *IEEE Transactions on Visualization and Computer Graphics*, *19*(12), 2306-2315.

Borkin, M. A., Bylinskii, Z., Kim, N. W., Bainbridge, C. M., Yeh, C. S., Borkin, D., Pfister, H., Oliva, A. (2016). Beyond memorability: Visualization recognition and recall, IEEE Transactions on Visualization and Computer Graphics, 22, (1), 519-528.

Breslow, L. A., Ratwani, R. M., & Trafton, J. G. (2009). Cognitive models of the influence of color scale on data visualization tasks. Human Factors, 51(3), 321-338.

Brumberger, E. R. (2003). The rhetoric of typography: The awareness and impact of typeface appropriateness. *Technical communication*, *50*(2), 224-231.

Burstein, P. (2014). *American public opinion, advocacy, and policy in congress: What the public wants and what it gets*. Cambridge University Press.

Choi, B.K., Pang, T., Lin, V., Puska, P., Sherman, G., Goddard, M. et al. (2005). Can scientists and policy makers work together? *J Epidemiol Community Health*, 59, 632-637.

Cleveland, W. S., & McGill, R. (1984). Graphical perception: Theory, experimentation, and application to the development of graphical methods. *Journal of the American Statistical Association*, *79*(387), 531-554.

Cowan, N. (2000). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences, 24*, 87-185.

Dagenais, C., Lysenko, L., Abrami, P. C., Bernard, R. M., Ramde, J., & Janosz, M. (2012). Use of research-based information by school practitioners and determinants of use: a review of empirical research. *Evidence & Policy: A Journal of Research, Debate and Practice*, 8(3), 285-309.

Davidson, E. J. (2007). Unlearning some of our social scientist habits. Journal of Multidisciplinary Evaluation, 4(8), p. Iii-vi.

Davies, H. T., & Powell, A. E. (2012). Communicating social research findings more effectively: what can we learn from other fields? Evidence & Policy: A Journal of Research, Debate and Practice, 8(2), 213-233.Duarte, N. (2012). *HBR guide to persuasive presentations*. Harvard Business Press.

Evergreen, S. (2011). Death by boredom: The role of visual processing theory in written evaluation communication. Retrieved from ScholarWorks, 403.

Evergreen, S. (2013). Smart data presentation. Paper presented at the American Evaluation Association annual conference, Washington DC.

Evergreen, S. (2016). Effective data visualization. Thousand Oaks, CA: SAGE.

Evergreen, S. (2017). Presenting data effectively: Communicating your findings for maximum impact. Thousand Oaks, CA: SAGE.

Few, S. (2006). Visual communication: Core design principles for displaying quantitative information. Retrieved September 9, 2010 from http://www.perceptualedge.com/articles/Whitepapers/Visual_Communication.pdf

Few, S. (2013). *Information Dashboard Design: Displaying data for at-a-glance monitoring.* Analytics Press.

Galesic, M., Garcia-Retamero, R., & Gigerenzer, G. (2009). Using icon arrays to communicate

medical risks: Overcoming low numeracy. *Health Psychology*, 28(2), 210-216.

Gilbert, A. N., Fridlund, A. J., & Lucchina, L. A. (2016). The color of emotion: A metric for implicit color associations. *Food Quality and Preference*, *52*, 203-210.

Hageboeck, M., Frumkin, M., Heavenrich, J. L., Kasimova, L., Management Systems International, Mark, M. M., & Perez-Linan, A. (2016). *Evaluation Utilization at USAID*. Retrieved from http://pdf.usaid.gov/pdf_docs/pa00kxvt.pdf.

Haroz, S., Kosara, R., & Franconeri, S.L. (2015) Isotype visualization: Working memory, performance, and engagement with pictographs. In Proc. of CHI, pages 1191–1200. ACM, 2015.

Heer, J., & Bostock, M. (2010, April). Crowdsourcing graphical perception: using mechanical turk to assess visualization design. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 203-212). ACM.

Jacoby, W. G. (2006). The dot plot: A graphical display for labeled quantitative values. *The Political Methodologist*, *14*(1), 6-14.

Jamet, E., Garota, M. & Quaireau, C. (2008). Attention guiding in multimedia learning. Learning and Instruction, 18, 135-145.

Knaflic, C. N. (2015). *Storytelling with data: a data visualization guide for business professionals*. John Wiley & Sons.

Kosara, R., & Skau, D. (2016, June). Judgment error in pie chart variations. In *Proceedings of the Eurographics/IEEE VGTC Conference on Visualization: Short Papers* (pp. 91-95). Eurographics Association.

Kuhbandner, C., & Pekrun, R. (2013). Joint effects of emotion and color on memory. *Emotion*, *13*(3), 375.

Landis, J.R. & Koch, G.G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33(1), 159-174.

Lawrenz, F., Gullickson, A., & Toal, S. (1997). Dissemination: Handmaiden to evaluation use. *American Journal of Evaluation*, *28*(3), 275-291.

Mackiewicz, J., & Moeller, R. (2004, September). Why people perceive typefaces to have different personalities. In *Professional Communication Conference, 2004. IPCC 2004. Proceedings. International* (pp. 304-313). IEEE.

Malamed, C. (2009). Visual language for designers: Principles for creating graphics that people understand. Beverly, MA: Rockport.

Mason, S. (2014). *Probing the magic bullet: Are visual reports really more effective?* Presented at American Evaluation Association annual conference, Denver, CO.

Mathison, S. (2009). Seeing is believing: The credibility of image-based research and evaluation. In Donaldson, S.I., Christie, C.A., & Mark, M.M. (Eds.), *What Counts as Credible Evidence in Applied Research and Evaluation Practice?* (pp. 181-196). Thousand Oaks, CA: SAGE Publications, Inc.

Mayer, R. E. (2009). *Multimedia learning, 2nd edition*. New York: Cambridge University Press.

Meeks, E. (2017, March 14). *2017 Data Visualization Survey Results*. Retrieved from https://medium.com/@Elijah_Meeks/2017-data-visualization-survey-results-40688830b9f2

Miller, S. & Hughes, D. (2017) *The Quant Crunch: How the Demand for Data Science Skills is Disrupting the Job Market.* Burning Glass Technologies. Retrieved from https://www-01.ibm.com/common/ssi/cgi-bin/ssialias?htmlfid=IML14576USEN&

Moher, D., Hopewell, S., Schulz, K.F., Montori, V., Gøtzsche, P.C., Devereaux, P.J., Altman D.G. (2010). CONSORT 2010 Explanation and Elaboration: updated guidelines for reporting parallel group randomised trial. *BMJ*, 340:c869.

Morrison, S. (1930). First principles of typography. In S. Heller & P. B. Meggs (eds.) Texts on type: Critical writings on typography. New York: Allworth Press.

Murdock, A., Shariff, R., & Wilding, K. (2013). Knowledge exchange between academia and the third sector. Evidence & Policy: A Journal of Research, Debate and Practice, 9(3), 419-430.

Patil, T. H., & Brynjolfsson, A. M. (2014, November 05). *Data Scientist: The Sexiest Job of the 21st Century*. Retrieved from https://hbr.org/2012/10/data-scientist-the-sexiest-job-of-the-21st-century

Patton, M. Q. (2011). *Essentials of utilization-focused evaluation.* Sage Publications.

Petch, A., Lightowler, C., Pattoni, L., & Watson, I. (2014). Embedding research into practice through innovation and creativity: a case study from social services. Evidence & Policy: A Journal of Research, Debate and Practice, 10(4), 555-564.

Reckhow, S. & Galey, S. (2017, April). *Did Research Inform the National Policy Debate on Teacher Evaluation?* Paper presented at American Educational Research Association Annual Meeting, San Antonio, TX.

Reckhow, S., Holden, L., Tompkins-Stange, M. (2015, September). *Patrons of Ideas: How Advocacy Research Influences the Education Policy Debate*. Paper presented at American Political Science Association Annual Meeting, San Francisco, CA.

Rossi, P. H., Lipsey, M. W., & Freeman, H. E. (2003). *Evaluation: A systematic approach*. Sage publications.

Russ-Eft, D., Atwood, R., & Egherman, T. (2002). Use and non-use of evaluation results: Case study of environmental influences in the private sector. *American Journal of Evaluation*, *23*(1), 19-31.

Samara, T. (2007). Design elements: A graphic style manual. Beverly, MA: Rockport Press.

So, S. & Smith, M. (2002). Colour graphics and task complexity in multivariate decision making. Accounting, Auditing, & Accountability Journal, 15(4), 565-593.

Schwabish, J. A. (2014). An economist's guide to visualizing data. *The Journal of Economic Perspectives*, *28*(1), 209-233.

Schulz K.F., Altman, D.G., & Moher, D. (2010). CONSORT 2010 Statement: updated guidelines for reporting parallel group randomized trials. *Open Med*, 4(1), 60-68.

Smith, J. (2016, January 20). *The 25 best jobs in America right now, according to employees*. Retrieved from http://www.businessinsider.com/best-jobs-in-america-right-now-2016-1/#2-tax-manager-24

Smith, K. E., & Joyce, K. E. (2012). Capturing complex realities: understanding efforts to achieve evidence-based policy and practice in public health. Evidence & Policy: A Journal of Research, Debate and Practice, 8(1), 57-78.

So, S. & Smith, M. (2002). Colour graphics and task complexity in multivariate decision making. *Accounting, Auditing & Accountability Journal, 15*(4), 565-593.

Stenberg, G. (2006). Conceptual and perceptual factors in the picture superiority effect. European Journal of Cognitive Psychology, 18, 813–847.

Stetson, V. (2008). *Communicating and reporting on an evaluation: Guidelines and tools*. Baltimore, MD: Catholic Relief Services.

Stone, D. (2012), *Policy Paradox: The Art of Political Decision Making (3rd Edition)*. New York, NY: W.W. Norton & Company, Ltd.

Sue, V. M., Griffin, M. T. (2016). *Data Visualization & Presentation with Microsoft Office*. SAGE Publications.

Sullivan, S.M., Wells, G., & Coyle, D. (2015). What guidance are economists given on how to present economic evaluations for policymakers? A systematic review. *Value in Health*, 18, 915-924.

Sutton, T. M., & Altarriba, J. (2016). Color associations to emotion and emotion-laden words: A collection of norms for stimulus construction and selection. *Behavior research methods*, *48*(2), 686-728.

Tal, A. & Wansink, B. (2016). Blinded with science: Trivial graphs and formulas increase ad persuasiveness and belief in product efficacy. Public Understanding of Science, 25(1), 117-125.

Tufte, E. R. (1990). Envisioning information. Chesire, CN: Graphics Press.

Tufte, E. R. (2001). *The Visual Display of Quantitative Information (Second Edition)*. Chesire, CT: Graphics Press LLC.

Ware, C. (2008). Visual thinking for design. Burlington, MA: Morgan Kaufmann Publishers.

Weiss, C. (1998). Have we learned anything new about the use of evaluation? *American Journal of Evaluation*, *19*(1), 21-33.

Wheildon, C. (2005). *Type and layout: Are you communicating or just making pretty shapes?* Victoria, Australia: The Worsley Press.

Wigdor, D., Shen, C., Forlines, C., & Balakrishnan, R. (2007, April). Perception of elementary graphical elements in tabletop and multi-surface environments. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 473-482). ACM.

Zacks, J., & Tversky, B. (1999). Bars and lines: A study of graphic communication. Memory and Cognition, 27, 1073-1079.

Zikmund-Fisher, B.J. (2014). *Image Request IR5: Risks in a population: emergent disease case counts* [project summary]. Center for Health Communications Research, University of Michigan. Ann Arbor.

Appendix. **Evaluation Report Layout Checklist.**

Evaluation Report Layout Checklist

This checklist is meant to be used as a diagnostic guide to identify elements of evaluation reports that could be enhanced using graphic design best practices and/or the assistance of a graphic design expert. Suggestions are best suited for those using standard Microsoft Word software.

Instructions Rate each aspect of the report using the following rubric, by circling the most appropriate letter. Use Best Practice section as a guide for improvement.

F=Fully Met P=Partly Met N=Not Met

Type	Rating	Best Practice
Text fonts are used for narrative text	F P N	Use serif fonts. Nothing with lots of graphic detail.
Long reading is in 9-11 point size	F P N	Studies have shown that 11 point text is easiest to read at length, but it can depend on the typeface (font).
Body text has stylistic uniformity	F P N	Each text section has unbolded, normal text in sentence case (no all caps), except in short areas of intentional emphasis. This supports undistracted reading.
Line spacing is 11-13 points	F P N	For lines within paragraph, generally choose 1-2 points larger than the size of the body text.
Headers & callouts are emphasized	F P N	Header should be 150-200% of body text size. Sans serif or decorative is okay. Use sentence case. Contrast with body text by using different size, style, and/or color. Too similar looks unintentional.
No more than 3 fonts are used	F P N	A change in font will indicate a change in meaning. Use font changes to guide reader through information according to importance.
Bullets are slightly less thick than text	F P N	If bullets must be used, decrease their size to slightly less (70-80%) than the point size of the font. Otherwise, they are too strong and distracting. If good spacing is used in lieu of bullets, this best practice is Fully Met.

Notes

Nice serif choices include Garamond, Palatino, Cambria

Nice sans serif choices are Trebuchet, Verdana, Calibri

Sentence case is when the first letter of the line is capitalized and all others are lowercase, excepting proper nouns.

Body text is that which comprises the narrative of the report.

By contrast, *header text* is that which comprises your headlines and titles. Also known as display text.

- Default bullet size (too big)
- Appropriate bullet size

Evergreen, S., Lyons, J. & Rollison, J. (2018). The link between graphic design and report use. Paper presented at the American Evaluation Association annual conference, Atlanta, GA, 26 October, 2016.

Arrangement	Rating	Best Practice
Alignment is consistent	F P N	Alignment is a preattentive feature easily picked up by a reader, so be sure elements start in the same place on each page unless misaligned on purpose. Avoid centered elements.
Columns are 8-12 words in length	F P N	This is 50-80 characters, depending on font. Longer is difficult to track from line to line, shorter creates too many hyphenated words, distracting the reader. See?
Important elements are prominent	F P N	Most prominent position is top half of page and/or emphasized by size, color, orientation, etc. Supportive information is toned down.
Body text is left or full justified	F P N	Ragged right edge is more informal, but easier to read for average readers. Full justification is formal, easier for fluent readers, but creates design issues with "white rivers" or large gaps of white space between words.
Grouped items logically belong together	F P N	Grouped items are interpreted as one chunk. Place logical items together. Add space between groups. Minimize space between header and body text.
Empty area is allocated on each page	F P N	Leave plenty of space between paragraphs, around page margins, and between text and graphics. It gives eyes a rest.

Notes

Imagine each page divided into rows and columns. Draw imaginary lines to check that elements are aligned at the start of each row and top of each column.

Asymmetry is an easy way to create interest. Try placing a cool picture off to one side of the page.

Wide margins are a quick way to create empty area and manage line length.

Graphics	Rating	Best Practice
Pictures/graphic elements are present	F P N	Multimode learning increases chance at storage of info in long-term memory because it eases cognitive load of body text. Choose pictures or graphics related to your topic. Graphics include, but shouldn't be limited to, tables and charts. If there are no graphics, this section is all Not Met.
Graphics are near associated text	F P N	If readers must flip around to interpret between text and graphic, comprehension will be impaired.
Graphics are simple	F P N	Less visual noise leads to better assimilation. Eliminate gradation, textures, or graphics as backgrounds. Segment complex graphics into smaller chunks.
Size corresponds to changes in meaning	F P N	Use, for example, larger pictures on chapter start pages. In graphing, for example, be sure height of columns proportionately represents data.

Notes

Pictures and graphics related to your content will make your content more memorable.

Choose pictures from quality sources, like paid websites. Watermarks or fuzzy images are signs of an amateur.

Use a cover page at the beginning of a report. This is a good place for a very large graphic.

Evergreen, S., Lyons, J. & Rollison, J. (2018). The link between graphic design and report use. Paper presented at the American Evaluation Association annual conference, Atlanta, GA, 26 October, 2016.

Graphics direct toward text	F P N	Use the power of an image to direct the reader's gaze from the image to the associated text. Eyes in a photo, for example, should look inward at text.
Visual theme is evident	F P N	Pick a visual theme that can be used in different forms throughout report to give strong emotional connection.
Some elements are repeated	F P N	Repetition of some graphic elements adds unity to the piece, makes work more memorable. Careful not to overdo it – too many elements can add clutter or complication.

Color	Rating	Best Practice
Narrative text is dark grey or black	F P N	Black has highest comprehension levels, with low intensity colors taking a distant second place.
Background has white/subdued color	F P N	Reversed-out text (e.g., white text on black background) impairs information retention.
One or two emphasis colors are used	F P N	Subdued colors that still contrast with background should be used. When used, it should be to actually emphasize important information, like data in a graph. If more than one is selected, consider choosing along a color gradation so that order of importance is implicit.
Color changes mark meaning changes	F P N	Color changes signal a change in hierarchy of information. Be intentional with color changes so that a viewer doesn't get confused.
Color reprints legibly in black and white	F P N	Color looks different on a computer screen than on paper. Print on a black-and-white printer and then make a copy of that printout to check legibility.

Notes

Keep in mind various culture-laden *color connotations*. For example, pink is highly associated with feminine qualities in the USA. Make sure your color choices are appropriate for your audience.

Note that *people with colorblindness* have difficulty with red-green and yellow-blue combinations.

A safe bet is to *use your client's colors*.

Time to add up your points: _____

F = 1 point

P = ½ point

N = 0 points

Well-formatted reports score within 23-25 points. At this level, report readers are better able to read and retain content. For more easily-accessible resources, check out:

Robin Williams' book *The Non-Designer's Design Book*

Design Elem ready for more advanced material

Evergreen, S., Lyons, J. & Rollison, J. (2018). The link between graphic design and report use. Paper presented at the American Evaluation Association annual conference, Atlanta, GA, 26 October, 2016.