March 2017

Descriptive analysis in education: A guide for researchers

Susanna Loeb
Stanford University

Susan Dynarski
University of Michigan

Daniel McFarland
Stanford University

Pamela Morris
New York University

Sean Reardon
Stanford University

Sarah Reber
UCLA

Key Themes

- Descriptive analysis characterizes the world or a phenomenon—answering questions about who, what, where, when, and to what extent. Whether the goal is to identify and describe trends and variation in populations, create new measures of key phenomena, or describe samples in studies aimed at identifying causal effects, description plays a critical role in the scientific process in general and education research in particular.
- Descriptive analysis stands on its own as a research product, such as when it identifies socially important phenomena that have not previously been recognized. In many instances, description can also point toward causal understanding and to the mechanisms behind causal relationships.
- No matter how significant a researcher's findings might be, they contribute to knowledge and practice only when others read and understand the conclusions. Part of the researcher's job and expertise is to use appropriate analytical, communication, and data visualization methods to translate raw data into reported findings in a format that is useful for each intended audience.

U.S. Department of Education
Betsy DeVos, *Secretary*

Institute of Education Sciences
Thomas W. Brock, *Commissioner for Education Research*
Delegated the Duties of Director

National Center for Education Evaluation and Regional Assistance
Audrey Pendleton, *Acting Commissioner*
Elizabeth Eisner, *Acting Associate Commissioner*
Amy Johnson, *Project Officer*

NCEE 2017–4023

The National Center for Education Evaluation and Regional Assistance (NCEE) conducts unbiased large-scale evaluations of education programs and practices supported by federal funds; provides research-based technical assistance to educators and policymakers; and sup-ports the synthesis and the widespread dissemination of the results of research and evaluation throughout the United States.

March 2017

This report was prepared for the Institute of Education Sciences (IES) by Decision Information Resources, Inc. under Contract ED-IES-12-C-0057, Analytic Technical Assistance and Development. The content of the publication does not necessarily reflect the views or policies of IES or the U.S. Department of Education nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government.

Loeb, S., Dynarski, S., McFarland, D., Morris, P., Reardon, S., & Reber, S. (2017). *Descriptive analysis in education: A guide for researchers.* (NCEE 2017–4023). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.

This report is available on the Institute of Education Sciences website at http://ies.ed.gov/ncee/.

The authors would like to thank Tom Szuba, of Quality Information Partners, for his substantial role as a contributing writer.

Contents

Boxes

Figures

About This Document

Purpose

This document presents a guide for more effectively approaching, conducting, and communicating quantitative descriptive analysis, which is a critical component of the scientific process. Because understanding "what is" is essential to successful education research and effective policy and practice, this document also makes recommendations for improving the ways in which quantitative descriptive findings are communicated throughout the education and research communities.

Descriptive analysis characterizes the world or a phenomenon—identifying patterns in the data to answer questions about who, what, where, when, and to what extent.

Why Now?

Over the past 15 years, a focus on randomized control trials and the use of quasi-experimental methods (such as regression discontinuity) has improved the body of causal research in education. However, this emphasis on causal analysis has not been accompanied by an improvement in descriptive analysis. In fact, advances in the methodological precision of causal analysis may have made descriptive studies *appear* to be a less rigorous approach to quantitative research. In contemporary work, descriptive analysis is often viewed simply as a required section in a paper—motivating a test of effectiveness or comparing the research sample to a population of interest. This view of descriptive research is shortsighted: good descriptive analysis is often challenging—requiring expertise, thought, and effort—and can improve understanding about important phenomena. The potential for description to inform policy, practice, and research is even more significant, given the recent availability of large and complex datasets that are relevant for understanding education issues.

Intended Audience

Because description is common across the spectrum of empirical research, the audience for this document is broad and varied. The primary audience includes members of the research community who conduct and publish both descriptive and causal studies using large-scale data. This audience includes Regional Educational Laboratory (REL) researchers,[1] other education researchers, and scholars from a range of disciplines such as sociology, psychology, economics, public policy, and the social sciences broadly.

While our focus is on education research, the vast majority of this report applies much more broadly to quantitative and qualitative descriptive work in a wide range of fields.

Although social scientists are one audience of research studies, other members of the education community also rely on research to improve their understanding of the education system. Thus, an important secondary audience is the policymakers (at local, state, and national levels) and practitioners (such as teachers and school administrators) who read about or otherwise apply research findings throughout

[1] The Regional Educational Laboratory (REL) program, sponsored by the Institute of Education Sciences (IES) at the U.S. Department of Education, works in partnership with school districts, state departments of education, and others to use data and research to improve academic outcomes for students. Fundamentally, the mission of the RELs is to provide support for a more evidence-reliant education system. For more information about the REL program, visit http://ies.ed.gov/ncee/edlabs/.

the education system. The guide can be useful for these stakeholders because it identifies how description can be useful for policy decisions and because it can help them to distinguish relevant descriptive analyses from those that are ill-conceived or poorly implemented.

Organization

This document is organized into five chapters and related appendixes:

Chapter 1. Why Should Anyone Care about Descriptive Analysis? Raises awareness about the important role that descriptive analysis plays in the scientific process in general and education research in particular. It describes how quantitative descriptive analysis can stand on its own as a complete research product or be a component of causal research.

Chapter 2. Approaching Descriptive Analysis. Describes the iterative nature of the process of descriptive analysis, which begins with recognition of a socially meaningful phenomenon and advances through the identification of salient features, relevant constructs, and available measures. The process concludes (subject to iterative revision) when patterns in the data are observed and subsequently communicated in a format that is well suited to depict the phenomenon to a particular audience.

Chapter 3. Conducting Descriptive Analysis. Focuses on the specific components of description—including the research question, constructs, measures, samples, and methods of distillation and analysis—that are of primary importance when designing and conducting effective descriptive research.

Chapter 4. Communicating Descriptive Analysis. Reminds researchers (1) that no matter how significant their findings, those findings contribute to knowledge and practice only when others read and understand the conclusions and (2) that part of their job is to use appropriate communication and data visualization methods to translate raw data into reported findings in a format that is useful for each type of intended audience.

Chapter 5. Summary and Conclusions. Condenses the document's content into a concise summary of key messages.

Chapter 1. Why Should Anyone Care about Descriptive Analysis?

To understand what works in education, we need to identify causal relationships. For example, we might ask whether a specific academic intervention, such as a reading program, caused an effect, such as an increase in student performance, in a particular group of students. This type of causal analysis involves precise methods designed to isolate and measure the effects of specific variables hypothesized to be playing a significant role in a cause-effect relationship.[2]

While causal research garners substantial attention, most research (even most policy-relevant research) is descriptive. In order to know what types of interventions might be useful—what problems need to be solved—we must understand the landscape of needs and opportunities. Large-scale descriptive research provides this landscape. We focus here on quantitative description, in contrast to qualitative descriptive studies, which may have goals of identifying causal effects in specific contexts through ethnography or interpretive techniques. The goal of quantitative description is not deep understanding of personal perspectives of a phenomenon, but a more general understanding of patterns across a population of interest.

Quantitative descriptive analysis characterizes the world or a phenomenon by identifying patterns in data to answer questions about who, what, where, when, and to what extent. Descriptive analysis is data simplification. Good description presents what we know about capacities, needs, methods, practices, policies, populations, and settings in a manner that is *relevant to a specific research or policy question*. Thus, data alone are not descriptive research, because data are not purposeful: data dumps, all-purpose data dashboards, and generic tables of summary statistics may be useful for some purposes, but they do *not* qualify as descriptive *analysis*.

Causal research may be the "gold standard" for determining what works in education, but descriptive analysis is central to almost every research project and is a necessary component of high-quality causal analysis.

Descriptive analysis can stand on its own as a research product, such as when it identifies phenomena or patterns in data that have not previously been recognized. In many instances, however, quantitative description is part of a broader study that involves causal analysis. Causal research methods may yield strong evidence about the effects of an intervention, as implemented in a particular time and place, but descriptive research *explains* the conditions and circumstances of the cause.

A combination of causal and descriptive analysis is necessary for understanding "why" an intervention has a causal effect: a sound causal analysis can assess the effects of an intervention; and effective descriptive work can identify the characteristics of the population, the features of implementation, and the nature of the setting that is most relevant to interpreting the findings. When properly applied, description can help researchers understand a phenomenon of interest and use that knowledge to prioritize possible causal mechanisms, generate hypotheses and intervention strategies, interpret the findings of causal research, diagnose problems for practitioners and policymakers to address, and identify new issues to study.

[2] National Center for Education Evaluation and Regional Assistance. (2003). *Identifying and implementing educational practices supported by rigorous evidence: A user friendly guide* (NCEE EB2003). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance. Retrieved from http://ies.ed.gov/ncee/pdf/evidence_based.pdf.

When policymakers (at local, state, and national levels) and practitioners (such as teachers and school administrators) make good decisions about how to improve education, it is often because they have access to a broad body of information that is the product of both causal studies *and* descriptive analysis—pointing toward causal understanding of real phenomena occurring in our classrooms, schools, and school districts.

Descriptive Analysis and the Scientific Method

Application of the scientific method advances knowledge through observing phenomena, identifying questions, generating hypotheses, testing hypotheses, and then producing new observations, questions, and hypotheses. Descriptive analysis is a fundamental component of this process because of the role it plays in helping us to observe the world or a phenomenon and, subsequently, in identifying research questions and generating hypotheses based on what has been observed (see Box 1).

Box 1. Descriptive Analysis Is a Critical Component of Research

Descriptive analyses are central to almost every research project. Whether the goal is to identify and describe trends and variation in populations, create new measures of key phenomena, or simply describe samples in studies aimed at identifying causal effects, descriptive analyses are part of almost every empirical paper and report. Some studies provide excellent descriptive analyses that are clearly focused on relevant aspects of a phenomenon. Unfortunately, other descriptive studies do little to provide relevant information, instead presenting a range of facts only tangentially related to the topic at hand. To be useful as an application of the scientific method both the goals and the findings of descriptive work should be clear.

Descriptive Analysis as Stand-Alone Research

There are times when descriptive analysis stands on its own as research—particularly when findings focus on identifying undocumented phenomena, identifying hidden patterns in large datasets, or diagnosing real-world needs that warrant policy or intervention.

Descriptive analysis is relevant to all types of research. It can stand alone as a complete research project or supplement causal analyses.

This type of descriptive study can be especially informative when we do not yet have a basic understanding of a phenomenon. For example, when virtual classrooms were initially introduced in schools, policymakers, practitioners, and causal researchers wanted to assess its effect on teaching and learning. However, descriptive analysis was needed first to clarify our basic understanding of the key aspects of the new phenomenon. Descriptive research was used to answer questions like:

- *Who was enrolled in virtual education?* For example, was it homebound students for a finite period of time, students who took one or two virtual classes to supplement their traditional school experience, or full-time online students? Understanding who took online courses is useful for properly assessing their potential merit. The potential implications are different if students taking virtual classes have or don't have access to similar material in face-to-face settings.
- *When was virtual instruction occurring?* For example, was it during a specific class period during a school day or was it self-paced to permit students to work at their convenience? If the courses were largely synchronous (at a specific time), then they would probably not add flexibility to students' schedules, but if the courses were asynchronous (on-demand), they might add flexibility for students who need it. Thus, looking at the effects separately for those most likely to benefit from flexibility has merit.

- *How was time spent during the virtual course?* For example, did students interact more or less with instructors in the online setting? What about with peers? Are they exposed to a more diverse set of peers or to more effective instructors in one of the settings? Do they spend more time listening to lectures or working actively on problems? By understanding the differences between settings and the variation within settings, researchers shed light on when, where, and how the settings could affect student learning and experience.

Descriptive research can be particularly valuable in today's age of large datasets in which the volume of information may otherwise obscure recognition of basic relationships. Countless pieces of data are collected each day about our education system—each student's attendance, classroom participation, assessment results, grades, and disciplinary incidents; each school's enrollment, curriculum, class schedules, staff characteristics, and facilities; and every state's number and types of schools, revenues and expenses, and academic achievement.[3] Descriptive research can be used to distill these datasets into meaningful dimensions to uncover patterns and inform and improve decision-making.

Descriptive analysis can also be used to diagnose issues that warrant the immediate attention of policymakers, practitioners, and researchers. For example, a descriptive study that reveals a previously unknown obstacle to college enrollment, such as the "summer melt" (see Box 2), helps stakeholders understand that there is a problem and, subsequently, target and test interventions for the population in need. In such a case, the descriptive study informs practitioners about what is actually happening in their world—problems, opportunities, or other aspects of their system that they had not previously understood.

Box 2. Examples of Using Descriptive Analyses to Diagnose Need and Target Intervention on the Topic of "Summer Melt"

Diagnosing need

Arnold, K., Fleming, S., DeAnda, M., Castleman, B. L., & Wartman, K. L. (2009). The summer flood: The invisible gap among low-income students. *Thought and Action*, Fall: 23–34.

"Summer melt" refers to the phenomenon of high school students who expect, in the spring, to attend college when they graduate from high school but who fail to enroll in college the following fall. This paper was the first to recognize the phenomenon after a descriptive analysis was made of counselor records, exit surveys from graduating seniors, counselor-student interviews, and reports of actual college enrollments for students in the Big Picture Longitudinal Study (BPLS). The authors reported that with slight variations across schools, 95–100 percent of BPLS students were accepted into college, but even under best-case scenarios, one-third of the students reconsidered their college plans over the summer after graduation, and at least one in five decided not to begin college at all—the "summer melt."

[3] The development of statewide longitudinal data systems (SLDSs) represents a particularly large and rich source of education data. These systems, which are partially sponsored by the IES SLDS Grant Program, are intended to enhance the ability of states, districts, schools, and educators to efficiently and accurately analyze and use education data, including individual student records. SLDSs are expected to help policymakers and practitioners make better data-informed decisions in order to improve student learning outcomes. They will also facilitate research that will contribute to efforts to increase student achievement and close achievement gaps. For more information about the SLDS Grant Program, visit http://nces.ed.gov/programs/SLDS/.

Targeting and Testing Intervention

Castleman, B. L., Arnold, K. D., & Wartman, K. L. (2012). Stemming the tide of summer melt: An experimental study of the effects of post-high-school summer intervention on college enrollment. *The Journal of Research on Educational Effectiveness, 5*(1): 1–18.

Castleman, B. L., Page, L. C., & Snowdon, A. L. (2012). *Summer melt handbook: A guide to investigating and responding to summer melt*. Cambridge, MA: Harvard University Center for Education Policy Research. Retrieved March 25, 2015, from http://cepr.harvard.edu/cepr-resources/files/news-events/sdp-summer-melt-handbook.pdf.

The diagnosis of this previously unrecognized problem resulted in education policymakers and practitioners having an urgent need to develop interventions that would counter summer melt. An example of an intervention resource that emanated from the descriptive work and the resultant development and testing of an intervention is the *Summer melt handbook: A guide to investigating and responding to summer melt* (the preceding reference), which states:

"This guide is written for people who want to understand and confront the summer melt problem. It is intended to provide specific guidance for designing and implementing a summer counseling initiative to mitigate summer melt in your district or student community. It will be useful to people in a variety of roles, including: school district administrators, school counseling staff and leaders, high school leaders (that is, headmasters, principals and vice-principals), and community-based organizations focused on providing resources to high school students."

Similarly, this identification of summer melt as a problem led to the testing of interventions to reduce the melt, such as a text messaging program for potentially at-risk high school graduates.

Descriptive Analysis as a Component of Causal Research

Causal research methods may generate strong evidence about the effects of an intervention, but descriptive research explains the conditions and context of the cause. In causal research, answering the question of whether an intervention "worked" is typically only the beginning of the research effort. The natural and necessary extension of such analysis is "why" or, perhaps more often in the field of social science experimentation, "why not." Descriptive data can facilitate answering these "why" and "why not" questions by providing real-world data that help to frame, contextualize, and interpret causal study.

When descriptive analysis uncovers facts that appear to confirm or reject particular theories or causal stories, some researchers refer to this as moving "toward a causal understanding."

During the planning phase of causal research, descriptive analysis can create or contribute to the rationale for undertaking a study. Descriptive identification of a phenomenon can help identify the potential benefits of interventions that can then be tested. Moreover, although descriptive methods cannot be used to assert a causal relationship between variable X and outcome Y, they can be used to exclude explanations that are not consistent with observation—thereby serving as evidence that refutes proposed causal mechanisms that do not reflect what has been seen in the data. This evidence can, in turn, be integrated into the researcher's hypothesis and subsequent planning of a study's design, interventions, and methods (see Box 3).

Box 3. An Example of Using Descriptive Analysis to Evaluate Plausible Causes and Generate Hypotheses

Scott-Clayton, J. (2012). What explains trends in labor supply among U.S. undergraduates? *National Tax Journal, 65*(1): 181–210.

Using October's Current Population Survey (CPS) data from the U.S. Census, the paper in the preceding citation describes an increase in paid work among college students over time (1970–2000). While description did not enable Scott-Clayton to confirm the cause of this change, she was able to examine the plausibility of a number of explanations through a series of descriptive analyses. In particular, by separately showing the trends for students with different background characteristics, she was able to reveal plausible causes: demographic changes in college student populations did not entirely explain the trend and the expansion of the federal work-study program appeared to coincide with the shifts in employment rates. This example demonstrates that while descriptive work cannot assert causal effect, it can contribute to our understanding of causal mechanisms, largely by ruling out some potential causes and subsequently influencing the generation and prioritization of other plausible hypotheses.

Interpreting "why" an intervention did or did not produce a causal effect is strengthened when a researcher has access to descriptive details that accurately and thoroughly characterize the context and conditions of a study. Such data fuel the design of more robust research efforts and advance the scientific method and ongoing discovery.

For example, recent attention to implementation fidelity within causal studies highlights the usefulness of information not only on what works, but also on how variation in the delivery of an intervention can influence demonstrated treatment effects.[4] Descriptions of implementation often focus on how implementation may have varied across settings or across populations and provide valuable insight into variation in study findings. Similarly, many causal studies rely on descriptive analysis to portray the quality of, and identify anomalies in, data. Analysis of this type is not always presented in public reporting, but it is nonetheless critical to understand when assessing the reproducibility and generalizability of findings (see Box 4 for two examples).

Quality description of the context and conditions of a study influences the interpretation of study findings and advances the scientific method and ongoing discovery.

Box 4. An Example of Using Descriptive Analysis to Interpret Causal Research

Glazerman, S., Isenberg, E., Dolfin, S., Bleeker, M., Johnson, A., Grider, M., & Jacobus, M. (2010). *Impacts of comprehensive teacher induction: Final results from a randomized controlled study* (NCEE 2010–4028). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.

To evaluate the impact of comprehensive teacher induction relative to low-intensity induction support generally given to beginning teachers, the research team conducted a randomized experiment in a set of districts that were not already implementing a comprehensive induction program. Key findings showed that exposure to one or two years of comprehensive induction did not affect retention or other teacher workforce outcomes.

In order to better understand this lack of effects, the authors reviewed various aspects of implementation. One part of this analysis described differences in the hours of induction received by teachers in both the treatment and control groups. While the treatment group received more support than the control group, control teachers still received substantial mentoring. For example, in fall 2005, control teachers in districts doing mentoring for one year received 67 minutes of support, while treatment teachers received 87 minutes of support. This simple description points to the possibility that the lack of effect may not have been due to a lack of the value in the mentoring but, instead, to a weak treatment-control contrast.

With the appropriate caveats that this type of description should be interpreted with caution (because it is not causal), this application of descriptive analysis extended the research team's understanding of the complex dynamics occurring within their experimental work and clearly demonstrated the value of description as a tool for enhancing the interpretation of causal findings.

---

[4] Bianco, S. D. (2010, June). Improving student outcomes: Data-driven instruction and fidelity of implementation in a Response to Intervention (RTI) model. *TEACHING Exceptional Children Plus, 6*(5).

Quint, J., Zhu, P., Balu, R., Rappaport, S., & DeLaurentis, M. (2015). *Scaling up the Success for All model of school reform: Final report from the Investing in Innovation (i3) evaluation.* New York: MDRC.

This report describes the evaluation of the U.S. Department of Education's Investing in Innovation program scale-up of the Success for All (SFA) school reform model. Thirty-seven schools were randomly assigned either to receive SFA (19 schools) or an alternative reading program (18 schools). After 3 years, students in SFA schools scored significantly higher, on average, on a measure of phonics skills but not on tests of reading fluency or comprehension. The results were more positive for students who entered school with low preliteracy skills, who did gain reading fluency relative to the control group.

In order to understand the source of the effects, the research team collected principal and teacher surveys and the School Achievement Snapshot, a form used by SFA coaches, as well as interviews and focus groups with school personnel focused on implementing SFA program elements. They used this information to understand the extent to which each school implemented each element of the program. They find that all but two of the 19 SFA schools were able to implement SFA with at least adequate fidelity and that the breadth and depth of implementation improved over time, particularly between the first and second year of the program. The one element of SFA that the researchers identified as having been used less (in less than half the schools) is the computerized tutoring program for struggling students, Team Alphie. This descriptive part of the analysis shed light on whether the effects were driven by only a few schools implementing the program or by implementation of only some aspects of the program. Overall, it provides evidence that the program was implemented with fidelity and, thus, the results represent the program undertaken relatively fully, though the results might be slightly stronger for a cohort that began after the first year, since implementation was weaker earlier on.

The simplicity maxim often ascribed to Albert Einstein is relevant to anyone conducting and presenting descriptive research:

"Everything should be made as simple as possible, but not simpler."

The Researcher's Role

A range of empirical techniques supports effective descriptive analyses. Simple statistics that describe central tendencies and variation (for example, means, medians, and modes) are the most common tools of descriptive work and can be very helpful for describing data; however, more sophisticated techniques for data manipulation have improved our ability to describe phenomena. Geographic information systems (GIS), for example, offer tools for describing geographic variation. Similarly, network analysis provides methods for identifying patterns of interactions among individuals and organizations.

With exponential increases in the amount of available data and the power of new technologies to analyze large datasets, researchers are sometimes tempted to rely too heavily on sophisticated, but unnecessary, analytical methods. Far too many research reports are plagued by this "disease of complexity" in which complicated methods and presentation are assumed to imply greater scientific rigor or value. When descriptive research is conducted or presented, complexity is not better or more robust than simplicity—and it certainly isn't more useful as a tool for communicating findings to a reader.

The descriptive researcher's job is to reduce the body of data to a format that is useful for the audience. This data reduction does not imply that all aspects of a setting or phenomenon should be weighted equally. Instead, it focuses on the most salient features of the phenomenon as it *really* exists and, more broadly, the real-world context in which a research study is to be interpreted.

Appropriately presented descriptive analysis can help a reader[5]

- View the data in the correct context (real-world and research settings).
- Identify relevant information in the data.
- Assess the quality of the data, such as bias in data source(s).
- Recognize the assumptions, limitations, and generalizability of the findings.

Whether the researcher's goal is to describe trends in populations, create new measures of key phenomena, or simply describe methods used to identify causal effects, descriptive analysis is a valuable research tool (see Box 5). When approached correctly, it can contribute substantially to a wide range of studies, both descriptive and causal in nature.

Box 5. Common Uses of Descriptive Accounts in Education Research and Practice

- Establishing the characteristics of a place, population, policy, procedure, or phenomenon.
- Explaining how a system and its constituent components operate.
- Diagnosing real-world problems that need to be addressed by policies or interventions.
- Prioritizing potential causal mechanisms (what do the data corroborate, and what do they rule out?).
- Planning research rationale, design, and methods.
- Generating data-supported hypotheses and exploratory directions.
- Describing fidelity of implementation.
- Assessing and describing data quality.
- Simplifying data for improved understanding by researchers and other audiences (for example, distilling a body of geographic data into a single metric).

[5] National Forum on Education Statistics. (2012). *Forum guide to taking action with education data* (NFES 2013–801). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. Retrieved from http://nces.ed.gov/pubs2013/2013801.pdf.

Chapter 2. Approaching Descriptive Analysis

When approaching descriptive work, researchers should first endeavor to recognize a phenomenon of interest (something that is occurring in the world). Once a phenomenon has been identified, a researcher must fully consider the phenomenon in question, determine which features are most salient, and make choices about data collection and methods. The combination of conceptualization, design, planning, and analysis is iterative (see Box 6).

Box 6. Steps in a Descriptive Analysis—An Iterative Process

Step 1. Identify a phenomenon.

Step 2. Consider which features of the phenomenon are most salient.

Step 3. Identify the constructs (measures) that best represent these features.

Step 4. Determine whether there are observable patterns in the data.

Step 5. Communicate the patterns in the data that describe the realities of the phenomenon.

Step 6. Rethink and repeat as needed.

Approaching Descriptive Analysis as an Iterative Process

Step 1. Identify a phenomenon. What is the phenomenon? The first task in high-quality descriptive research is for the researcher to fully consider the phenomenon in question independently from the data and methods that he or she will eventually collect and use. Sometimes a researcher has a preconceived idea about a research question or phenomenon; at other times, the researcher discovers the question or phenomenon in the process of data collection or analysis.[6] In either case, once a question or phenomenon has been identified, a researcher should consider it thoughtfully before making choices about data and methods.

Step 2. Consider which features of the phenomenon are most salient. Which aspects, concepts, or categorizations of reality are necessary to describe the phenomenon? Which are not? The answers to these questions are likely to reveal what types of data collection will be most applicable for study. The researcher's conceptualization of the phenomenon will determine how to best organize the data for analysis and reduce the data for description. For example, consider a researcher trying to describe variation in classroom engagement. Classroom engagement entails a variety of interrelated factors, including students' motivations when they enter the classroom, teachers' instructional practices, the number and nature of peers, and the content of the material. The researcher must choose which of these concepts are most salient and worth exploring. The phenomenon may be complex and previously unmeasured. Without clear hypotheses of what constitutes the phenomenon, it will be difficult or impossible to describe and study.

Step 3. Identify the constructs (measures) that best represent the most salient features. What measures will be most effective to systematically observe important features of the phenomenon? Consider type(s) of data (for example, the unit of analysis) and method(s) of data collection that will produce the appropriate level of abstraction and quantification for analysis. After carefully considering the phenomenon, researchers should identify ideas, attributes, or concepts that will be measured. Choosing what to measure depends on what the data collector believes is relevant and feasible.

[6] See Chapter 3 for a discussion of "exploring the data" as it relates to descriptive analysis.

Whether collecting data or using pre-existing data sources, researchers must have a clear idea of what they want to measure so that they can identify any gaps in their data, some of which might be overcome by other data, modified methods, or thoughtful analyses. Sometimes, researchers will collect their own data. At other times, researchers are limited to existing datasets (and are not able to capture the unmeasured constructs). In either case, a researcher's goal at this stage of the process is to transform reality into data that are available for analysis.

Step 4. Determine whether there are observable patterns in the data. The researcher should use methods that will identify patterns in the data if such patterns are present. If observed, these patterns might provide a more developed narrative and holistic depiction of a phenomenon being studied. Once the constructs are clear and the data have been collected, researchers may use a wide range of investigative and statistical methods to scrutinize the data for patterns and relationships that describe the key concepts of the phenomenon. Sometimes, researchers review the data with an expectation that a certain pattern exists; at other times, patterns emerge from a more general exploration of the data. In descriptive work, unlike causal research, these efforts to identify patterns in the data do not always have to be limited to a pre-existing hypothesis.

Step 5. Communicate the patterns in the data that describe the realities of the phenomenon. What type of data presentation is best suited to depict the phenomenon? The answer to this question usually depends heavily on the intended audience and the types of data representations that they are comfortable interpreting. After identifying any patterns in the data, the researcher's job is to reduce the body of data to a format that is best suited to a particular audience. A research effort may include thousands of analyses, but good presentation is distilled and targeted to succinctly capture the essences of the phenomenon and to help the reader think more productively about a particular topic. This translation, from raw to reported findings, is undertaken specifically to meet the information needs of practitioners, policymakers, other researchers, or other audiences (see Chapter 4). Such translation, however, does not need to compromise the thoroughness of the descriptive research. In order to ensure that high-quality descriptive research continues to be valued in education, this research needs to be reported transparently, with the details easily available to those who are interested.

Step 6. Rethink and repeat as needed. The process of descriptive analysis is iterative, with each step building upon others and requiring reconsideration and modification as the researcher's understanding of the phenomenon, relevant theory, and the study advances. In most cases, the completion of a study does not complete our understanding of a phenomenon. Few studies capture all aspects of a phenomenon, so the researcher's goal is often to describe relevant elements in convincing and compelling ways—contributing to a larger body of research that advances knowledge and fuels future research.

Meaningful Descriptive Analysis Reveals Socially Important Patterns

Descriptive research becomes relevant when it identifies patterns in data that convey meaningful information. This information may be obviously or immediately meaningful to practitioners, policymakers, or other researchers, or the researcher may reveal its importance through the interpretation of the description. In order to be meaningful, the patterns must be socially important, not simply present. Consider the case of describing a dataset by providing the means of each variable for two groups—males and females. This description provides information and may even show some patterns (males have higher means on some variables and lower on others), but it isn't a useful description unless it explicitly identifies patterns that are socially meaningful (see Box 7).

Box 7. Data Summaries Are Not Descriptive Analysis

The National Center for Education Statistics (NCES) publishes numerous valuable reports that summarize the data that it collects. For example, the *Digest of Education Statistics* contains data about a variety of subjects in the field of education statistics, including the number of schools and colleges, teachers, enrollments, and graduates, in addition to educational attainment, finances, federal funds for education, libraries, and international education.[7] This annual report broadly describes the education context in the United States and is a valuable tool for both policy and research, but it is not intended to illuminate particular phenomena or communicate their importance. Other NCES reports include analysis and interpretation that are somewhat more descriptive in nature (for example, *The Condition of Education*).[8] Academic researchers have also used NCES data to uncover patterns more intentionally through descriptive analyses.[9] This difference in focus and motivation distinguishes summary statistics from descriptive analysis.

Examples of Descriptive Studies That Reveal Consequential Phenomena

The following examples of descriptive studies highlight how patterns in data can describe and reveal phenomena that are important to education practitioners, policymakers, and researchers.

- In "The Widening Academic Achievement Gap Between the Rich and the Poor: New Evidence and Possible Explanations," Reardon (2011) describes how the achievement gap between children from high- and low-income families is approximately 30–40 percent larger among children born in 2001 than among students born 25 years earlier—and appears to have been growing for the past 50 years.[10] The study clearly defines the phenomenon as well as the unique methodological contribution relative to previous research. In particular, the study brings to bear a more comprehensive combination of datasets to trace the evolution of socioeconomic achievement gaps over time. While 50 years ago, the achievement gap between black and white children was larger than the gap between the tenth and ninetieth percentile of the income distribution, the achievement gap is now almost twice as large on the basis of income rather than race. The study exemplifies productive descriptive analysis because the use of data is carefully laid out and justified, the results are believable, and the implications affect policy and practice (for example, the need to prioritize efforts to overcome

---

[7] The primary purpose of the *Digest of Education Statistics* is to provide a compilation of statistical information covering the broad field of American education from prekindergarten through graduate school. For more information about the *Digest of Education Statistics*, visit http://nces.ed.gov/programs/digest/.

[8] *The Condition of Education* summarizes important developments and trends in education and uses the latest available data. The report presents 42 indicators on the status and condition of education, which represent a consensus of professional judgment on the most significant national measures of the condition and progress of education for which accurate data are available. For more information about *The Condition of Education*, visit http://nces.ed.gov/programs/coe/.

[9] See, for example, Master, B., Sun, M., & Loeb, S. (in press). Teacher workforce developments: Recent changes in academic competiveness and job satisfaction of new teachers. *Education Finance and Policy.* That study uses three Baccalaureate and Beyond studies from NCES to describe changes in the characteristics of new college graduates who enter teaching.

[10] Reardon, S. F. (2011). The widening academic achievement gap between the rich and the poor: New evidence and possible explanations. In R. Murnane & G. Duncan (Eds.), *Whither opportunity? Rising inequality and the uncertain life chances of low-income children.* New York: Russell Sage Foundation Press.

the achievement gap for children from lower-income families) as well as research (for example, the need to identify interventions that can benefit these children).

- "Teacher sorting and the plight of urban schools: A descriptive analysis" (Lankford, Loeb, & Wyckoff, 2002) is an example of descriptive research that assesses the distribution of educational resource availability rather than just variation in education outcomes.[11] The study presents a new, rich set of data about teachers in order to evaluate variation in the attributes of teachers across schools. Data show large differences in the qualifications of teachers across schools, with urban schools and schools with low-income, low-achieving, and non-white students, in particular, having less-qualified teachers. The study has implications for both research and practice, pointing toward a need to better understand why teacher characteristics are sorted (distributed), what interventions might reduce sorting, and the importance of trying to change these patterns even when limited by imperfect information about what has been proven to work in a real-world setting.
- In "The missing 'one-offs': The hidden supply of high-achieving, low-income students," Hoxby and Avery (2013) present an example of how a descriptive study can identify previously unrecognized phenomena.[12] The authors rely on ACT and College Board data for every student in the high-school graduating class of 2008 who took either the ACT or the SAT I, including information on where they applied to college. With these data, the authors were the first to identify a problematic phenomenon in education—that high-achieving, low-income students, especially those in small districts without selective public high schools, were not applying to selective colleges, even though they had academic records that could qualify them for admissions. The study creates actionable information for policymakers and practitioners by identifying geographic areas with academically eligible, low-income students who would potentially benefit from appropriate intervention strategies.
- In "Teaching students what they already know? The (mis)alignment between instructional content in mathematics and student knowledge in kindergarten," Engel, Claessens, and Finch (2013) report that although most children enter kindergarten already able to count and recognize geometric shapes, teachers still spent substantial class time on this material.[13] The study was useful because it identified the phenomenon of the misaligned kindergarten curriculum. Moreover, its publication in a journal that focused on education evaluation and policy analysis ensured that the appropriate audience would learn about the phenomenon and conceivably respond with efforts to better align kindergarten curriculum with the needs of kindergarten students.
- In "Curricular flows: Trajectories, turning points, and assignment criteria in high school math careers," McFarland (2006) uses network analytic techniques to depict the actual

[11] Lankford, H., Loeb, S., & Wyckoff, J. (2002). Teacher sorting and the plight of urban schools: A descriptive analysis. *Educational Evaluation and Policy Analysis, 24*(1): 37–62.

[12] Hoxby, C., & Avery, C. (2013). The missing "one-offs": The hidden supply of high-achieving, low-income students. *Brookings Papers on Economic Activity, 46*(1): 1–65.

[13] Engel, M., Claessens, A., & Finch, M. A. (2013). Teaching students what they already know? The (mis)alignment between instructional content in mathematics and student knowledge in kindergarten. *Educational Evaluation and Policy Analysis, 35*(2): 157–178.

course-taking patterns as a network of students flows across courses.[14] Using that representation, he identifies flow patterns and reveals different career paths and points where tracks intersect and overlap to show that it is in these junctures that most students are susceptible to "track changes" that determine future course taking. His description of the course-taking patterns reveals certain courses and types of students as being potentially more susceptible to interventions aimed at pushing them into college-bound courses.

Each of these studies focuses on a primary phenomenon, clearly defining its boundary and describing relevant features. For example, Reardon (2011) does not simply show trends in achievement separately for students with different characteristics. Rather, he specifically describes income and race gaps over time. Similarly, Engel et al. (2013) do not uniformly describe students' knowledge and the curriculum they experience, but instead, they specifically characterize the misalignment in great detail.

Informative descriptive studies often bring to bear new data that provide more convincing evidence about a phenomenon. Acquiring such data and applying them in a novel, but appropriate, manner often require a researcher to think originally or from a new perspective. In each of the examples above, researchers drew on new data or introduced unique approaches to using existing data. Reardon (2011) is innovative in his approach to combining multiple datasets to describe trends over time in a way that could not be achieved with a single dataset. Lankford et al. (2002) use state administrative data to characterize all teachers across a state and, subsequently, make comparisons across schools (within and across geographic areas) in ways that nationally representative, but incomplete, data could not be used. Similarly, Hoxby and Avery (2013) use data collected for practitioners (not specifically designed for research), but nonetheless are able to provide new insights. Finally, Engel et al. (2013) apply publicly available data from the National Center for Education Statistics' Early Childhood Longitudinal Study to a new research question, illustrating how a currently existing dataset can be used to provide new insights.

Descriptive Analysis to Support Causal Understanding

Although descriptive analysis can stand on its own as a research product, in some instances, description is a precursor to explanation and cause. Causal research methods may yield strong evidence about the effects of an intervention, but understanding "why" an intervention had a causal effect often necessitates a combination of causal and descriptive work: sound causal analysis to assess the effects and effective descriptive work to identify the characteristics of the population, the features of implementation, and the nature of the setting most relevant to interpreting the findings.

Descriptive analysis plays a central role in many aspects of causal research, including planning an intervention strategy, targeting interventions, contributing to the interpretation of causal study, assessing variation in treatment impact, and prioritizing potential causal mediators (when description provides evidence concerning which alternative hypotheses are more or less consistent with observed reality) (see Box 8).

[14] McFarland, D. A. (2006). Curricular flows: Trajectories, turning points, and assignment criteria in high school math careers. *Sociology of Education, 79*(3): 177–205.

Box 8. An Example of Using Descriptive Analysis to Support or Rule Out Explanations

Murnane, R. J. (2013). U.S. high school graduation rates: Patterns and explanations. *Journal of Economic Literature, 51*(2): 370–422.

Murnane (2013) uses multiple data sources and measurement approaches to document high-school graduation rates in the U.S. from 1970–2010—an important contribution, given that no single dataset measures national high-school graduation rates over time. Through this descriptive analysis, he describes six striking patterns, including stagnation over the last three decades of the twentieth century and increases in graduation rates over the first decade of the twenty-first century. The study then interprets salient patterns in the data in light of more or less plausible explanations for these trends:

- "Increases in high school graduation requirements during the last quarter of the twentieth century increased the nonmonetary cost of earning a diploma for students entering high school with weak skills. By so doing, they counteracted the increased financial payoff to a diploma and contributed to the stagnation in graduation rates over the last decades of the twentieth century. Of course, this raises the question of why high school graduation rates increased during the first decade of the twenty-first century, a period in which high school graduation requirements were not reduced, and in some states were increased." [page 47]
- "Evidence from the NAEP indicates an improvement over the last ten to fifteen years [leading into the first decade of the twenty-first century] in the reading and mathematics skills among entering freshman at the bottom of the skills distribution. This may have translated into lower nonmonetary costs of completing high school graduation requirements." [page 48]

Planning an Intervention Strategy

In the planning phase, descriptive analyses are valuable when determining whether an intervention is needed and how to design a needed intervention to be most effective. An example of this use of description comes from the Foundations of Learning Demonstration site in Newark, New Jersey.[15] Using descriptive analysis, researchers learned that (a) a substantial proportion of children in preschool classrooms in low-income communities have behavioral challenges; (b) such behavioral challenges were impeding teachers' ability to deliver effective instruction; (c) teachers reported that these issues were a primary concern, so staff could be expected to be open to strategies that address the issue; and (d) teachers were currently receiving very little training on these issues. As a result of the information that researchers learned from the descriptive analyses, the study's interventions were intentionally designed to target teachers' abilities to effectively set limits and support children's behavior regulation in the classroom. See Chapter 1 for introductory discussion on these and related topics.

Targeting Interventions

Descriptive data can also provide assessments of target populations that are more or less likely to have a positive response to an intervention. For example, Gennetian, Castells, and Morris (2010) used description to determine that when interventions that provided earning supplements were targeted at parents who were likely to work in the absence of the program, the supplements were less likely to achieve their purpose of increased self-sufficiency.[16] Similar effects were observed when in-

---

[15] Morris, P., Raver, C., Millenky, M., Jones, S., & Lloyd, C. (2010). *Making preschool more productive: How classroom management training can help teachers*. New York: MDRC. Retrieved from http://files.eric.ed.gov/fulltext/ED514648.pdf.

[16] Gennetian, L., Castells, N., & Morris, P. (2010). Meeting the basic needs of children: Does income matter? *Children and Youth Services Review, 32*(9): 1138–1148.

terventions targeted parents who had substantial barriers to their employment. But when intervention programs targeted parents who would not have otherwise worked, the financial incentive was more likely to meet the objective of increasing family income and self-sufficiency.

Contributing to the Interpretation of Causal Study

When a randomized trial is conducted to assess the effects of an intervention, the question of whether the intervention "worked" (had an effect) is usually only the beginning of analysis and interpretation. Descriptive data can be used to assess the subsequent (and critically important) "why" and "why not" questions. For example, when an effect is seen, descriptive analysis can help to identify plausible mechanisms; when an effect is not observed, description can support efforts to distinguish between theory failure and implementation failure.

One of the central challenges facing researchers is when their study finds no significant differences between control groups and treatment groups in spite of an intervention (which is termed "null effects"). Two competing explanations are typically at play: implementation failure and theory failure (Rosenbaum, 1986; Wandersman, 2009).[17] Implementation failure refers to an intervention that fails to meet its stated objectives because it was not effectively implemented, which would result in low-dosage or low-quality delivery of the program to the stated participants. For example, sometimes ineffective implementation occurs because of poor attendance (as happens when parents don't show up to a parenting program), or sometimes it occurs because of problems with the program itself (such as when internet or hardware issues affect the implementation of computer assisted instruction). In contrast, theory failure occurs when an intervention changes a participant's intended behavior, but that new behavior doesn't result in intervention outcomes as expected. Thus, it is a failure of the "theory of change" that led to null effects.

For example, in Jacob, Goddard, Kim, Miller, and Goddard (2015), the Balanced Leadership principal program resulted in participants reporting that they felt more efficacious, used more effective leadership practices, and created a better instructional climate than control group principals.[18] However, teachers indicated that the instructional climate of the schools did not change. Thus, the intervention resulted in new behavior as intended, but these changes did not lead to expected improvements in teachers' perspectives on instructional climate or student achievement, despite expectations about the value of strong leadership for effective schools—an example of theory failure. This example of description illustrates the problem with the theory behind the program—changes in principals' perceptions did not lead to changes in teachers' perspectives. More-intensive descriptive research (for example, ethnography) would be needed to understand why the teachers did not experience a change in the principals' behavior.

Null effects may occur for a third reason, unrelated to implementation failure or theory failure. In a randomized control trial, one group of participants is assigned to the intervention condition and

---

[17] Rosenbaum, D. P. (1986). *Community crime prevention: Does it work?* Beverly Hills, CA: Sage Publications. Wandersman, A. (2009). Four keys to success (theory, implementation, evaluation, and re-source/system support): High hopes and challenges in participation. *American Journal of Community Psychology, 43*(1–2), 3–21.

[18] Jacob, R., Goddard, R., Kim, M., Miller, R., & Goddard, Y. (2015). Exploring the causal impact of the McREL Balanced Leadership Program on leadership, principal efficacy, instructional climate, educator turnover, and student achievement. *Educational Evaluation and Policy Analysis, 37*(3): 314–332.

another is assigned to the control condition (the "counterfactual" without intervention). Program impact in randomized trials is produced by this contrast, which is the difference between services received by the treatment group (participants in the intervention) and those received by the control group (which does not receive the intervention).[19] Null results or variation in results across settings could be driven by the control conditions. In some cases, the control group may experience alternative interventions that mask the true effects of the intervention under study. Researchers can use descriptive data to assess and better understand the "counterfactual story" and other factors within treatment and control groups that may influence variation in the observed impact of an intervention.

For example, a recent analysis of 28 studies of Head Start conducted between the program's inception and 2007 found that much of the variation in Head Start's impact on child achievement and cognitive development could be explained by differences in the types of preschool services used by the control group (Shager et al., 2013).[20] Similarly, analysis of variation across sites suggested that program impacts were smaller for Head Start centers that draw more children from center-based programs rather than from home-based care (Walters 2014).[21] In these examples, description of the counterfactual condition (and other variation) can be useful when interpreting causal studies designed to predict or determine differences based on treatment effects.

Assessing Variation in Treatment Impact

Variation in treatment impact has been the focus of a number of efforts to synthesize research and uncover what works (and how much and under what conditions effects might be most pronounced). This analysis of moderators (or interactions of the treatment effect with pre-treatment characteristics) as a component of an impact study can be used to shed light on the circumstances in which larger and smaller impacts occur.[22] However, moderator variables should be chosen with reason, not simply because they are asked of every participant. Looking at many moderator effects may yield one or more that are "statistically significant" but statistical tests can yield "false positives." Conducting more tests yields more false positives, a point we return to below in discussing data "fishing."

---

[19] For more information about describing treatment contrast and the counterfactual condition, see Dynarski, M., & Kisker, E. (2014). *Going public: Writing about research in everyday language* (REL 2014–051). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Analytic Technical Assistance and Development. Retrieved from http://ies.ed.gov/ncee/pubs/REL2014051/pdf/REL_2014051.pdf.

[20] Shager, H. M., Schindler, H. S., Magnuson, K. A., Duncan, G. J., Yoshikawa, H., & Hart, C. M. (2013). Can research design explain variation in Head Start research results? A meta-analysis of cognitive and achievement outcomes. *Educational Evaluation and Policy Analysis, 35*(1): 76–95.

[21] Walter, C. (2014). *Inputs in the production of early childhood human capital: Evidence from Head Start* (NBER Working Paper No. 2014.01). Cambridge, MA: National Bureau of Economic Research.

[22] For more on analysis variation in treatment effects, see Schochet, P., Puma, M., & Deke, J. (2014). *Understanding variation in treatment effects in education impact evaluations: An overview of quantitative methods.* Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Analytic Technical Assistance and Development. Retrieved from https://ies.ed.gov/ncee/pubs/20144017/pdf/20144017.pdf.

Bloom, Hill, and Riccio (2003) present an innovative and influential example of empirical analysis that effectively brings together implementation and impact efforts.[23] In this synthesis, a series of randomized trials of welfare-to-work initiatives were conducted in 59 locations. In each location, data were collected to discern variation in approaches to program implementation. The authors were able to link these descriptive data about implementation to data about program impact at each location—showing which aspects of implementation were most strongly associated with impacts on earnings (the key outcome of these welfare-to-work experiments). Findings suggested that those program locations with smaller caseloads, close personal attention to participants, and a strong emphasis on getting a job quickly were the most effective at boosting earnings among program group participants.

Prioritizing Potential Causal Mediators

In many randomized trials, it is difficult to determine the precise cause of intervention impacts, which sometimes raises concerns about the "black box" nature of experiments. Explicitly proving that a particular mechanism is the basis for the causal effect of an intervention (that is, the "central causal mediator") may not be possible given the design of randomized experiments. However, descriptive information sometimes enables a researcher to triangulate on more and less likely candidates for a causal mechanism given available information (see Box 8 above).

For example, in the welfare reform arena, there has been much debate about the effectiveness of policy-induced increases in family income as a means for improving school achievement in children.[24] In the design of policies, practices, and intervention strategies, it is critical to understand how much, if any, of the association between parents' income and children's achievement is causal. Why, exactly, did preschool children benefit when their parents participated in welfare-to-work support programs? Was it the increased employment that their parents experienced as a result of the program, or was it their reduced reliance on welfare? At the family-process level, do children benefit because parents invested more financial resources in their children or because the parents were less stressed financially and, subsequently, more engaged in positive interactions with their children? Duncan, Morris, and Rodrigues (2011) revisited data on a number of possible mediators in a set of welfare and antipoverty experiments conducted in the 1990s.[25] Their descriptive findings point to the importance of income for boosting a child's achievement and suggest that family income has a policy-relevant, positive impact on the eventual school achievement of preschool children more broadly. While this descriptive analysis does not provide causal proof, and the authors cautioned against extrapolating beyond their specific findings, the study has compelling implications on both theory and policy—with findings suggesting the possibility that interventions that increase family income might be more effective mechanisms for improving student achievement than interventions in schools.

---

[23] Bloom, H. S., Hill, C. J., & Riccio, J. A. (2003). Linking program implementation and effectiveness: Lessons from a pooled sample of welfare-to-work experiments. *Journal of Policy Analysis and Management, 22*(4): 551–575.

[24] Magnuson, K., & Votruba-Drzal, E. (2009). Enduring influences of childhood poverty. *Focus, 26*(2): 32–37.

[25] Duncan, G., Morris, P., & Rodrigues, C. (2011). Does money really matter? Estimating impacts of family income on young children's achievement with data from random-assignment experiments. *Developmental Psychology, 47*(5): 1263–1279.

Approaching Descriptive Analysis: Summary

Descriptive analysis is a valuable research tool. It can contribute to a wide range of studies, both descriptive and causal in nature. When approaching descriptive work, researchers should endeavor to first recognize a phenomenon of interest. Once a phenomenon has been identified, the researcher must fully consider the phenomenon in question, determine which features are most salient, and define relevant constructs (measures) that represent these features. Analysis should focus on identifying patterns in the data that are most important to "telling the story." The researcher's job includes presenting the information in a format that is readily comprehensible for a particular audience or audiences. This approach to descriptive analysis is iterative, with each step building upon others and requiring reconsideration and modification as the researcher's understanding of the phenomenon and the study unfolds.

Chapter 3. Conducting Descriptive Analysis

The process of descriptive analysis begins with a phenomenon in question. Sometimes the phenomenon emerges from the data; sometimes it arises from experience or anecdote; and sometimes it comes from gaps in the extant research. Not all phenomena or questions about phenomena are well articulated or can be answered with existing or collected data, and often, a researcher needs to rethink, reconsider, and reevaluate the study question until it is well articulated, conceptually clear, and methodologically feasible. However it materializes, the questions and their importance must be precise and apparent—independent of the data and methods that will be used. The researcher must understand the phenomenon in question and the concepts that are central to it.

The researcher must then bring data to bear on the question of interest. Sometimes, multiple datasets are available and together provide better description than a single dataset, but even highly related data may not always be suitable to fully describe a phenomenon or answer a study question. While existing datasets may be appropriate for some aspects of the study, sometimes the data are not collected at the right time or at the right level of granularity, or they simply may not be available at all. When existing datasets are not appropriate, the researcher may choose to conduct a custom data collection. Whatever the data source, the researcher should ensure that the data sufficiently match the question, relevant constructs, appropriate measures, and available methods.

Descriptive research does not describe data—it uses data to describe the world for the purpose of identifying and improving our understanding of socially important phenomena.

Once the data are in hand, the researcher needs to understand what the observed facts are and how they relate to the study question. Doing so requires the researcher to select appropriate analytical methods for answering the research question. An important part of this descriptive analysis is characterizing the uncertainty of the observed phenomenon—no data are perfect and, by definition, any findings that emerge from the data are an inexact description of the world. Uncertainty is a critical factor for ascertaining how precise (or imprecise) even high-quality, highly relevant data are for their intended purposes.

Key Terminology and Methodological Considerations

The following concepts and terms are of primary importance when designing and conducting effective descriptive research: *research questions*, *constructs*, *measures*, *samples*, and *methods* of synthesis and analysis.

Research Questions

A compelling research question addresses a persistent or persuasive problem or intellectual tension such as when competing theories suggest different courses of action.[26] For education research, such questions will be socially important and highly relevant to improving our understanding of education processes, distribution (access), effects, and quality. Research questions that relate to these aspects of the education system often will inform decision-making about policies and practices in some regard. For example, identifying gaps in student outcomes or educational opportunities can point

[26] Booth, W. C., Colomb, G. G., & Williams, J. M. (2008). *The craft of research,* 3rd ed. (NFES 2013–801). Chicago, IL: The University of Chicago Press.

toward ways to more effectively target resources. Without a compelling research question, researchers may improve understanding about a phenomenon, but for what purpose?

The role of the researcher is to choose and refine the study question. Question development is iterative. The researcher might be generally interested in a particular topic, but preliminary observation or data analysis might suggest ways to refine the question of interest. For example, a researcher may initially be concerned about assessing variation in educational opportunities that are driven by differences in school quality. Further discernment, however, might reveal that even when spending (perhaps the most basic measure of equity) is the same across schools, differences in staff skills and experience might systematically disadvantage some students. With this logic, the researcher might be motivated to ask a more precise question: how do the skills and experience of teachers vary across schools that serve students of different types, as characterized by income, race, and ethnicity? In such a case, a fairly general topic of exploration has been distilled into a more specific and actionable research question.[27]

We present this review of key terms and methodological considerations as an overview of these issues as they relate to descriptive analysis. We do not fully capture the complexities of all topics.

Instead, we suggest that you refer to graduate-level coursework and texts to appropriately address the many facets of research design and methods.

Constructs

With a research question in hand, the researcher should identify the key constructs of interest (that is, those ideas, attributes, or concepts that will be measured). Good descriptive work requires a clear conceptualization of the constructs that one wants to describe, but achieving such clarity can be a challenging task (see Box 9).

The motivation for the research question often helps to distill and clarify the constructs of interest. For example, consider the broad topic of the distribution of teachers across schools. If the underlying motivation for a research question was to assess differences in educational opportunities for students, then a "teacher characteristics" construct within that broad topic would likely reflect the teachers' ability to provide educational opportunities. Relevant constructs might include teaching practices and subject-matter knowledge. Other constructs, such as gender, may be less appropriate for the purpose of that particular research question. If, however, the motivation for the research question originated in an interest in labor market demographics and occupational choices in the education workforce, a "teacher characteristics" construct might include constructs related to gender, age, and race-ethnicity rather than teaching practices and subject matter knowledge.

---

[27] Logic models can aid in this process of refinement. For more in-depth treatment of this approach, see Lawton, B., Brandon, P. R., Cicchinelli, L., & Kekahio, W. (2014). *Logic models: A tool for designing and monitoring program evaluations* (REL 2014–007). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Pacific. Retrieved from https://ies.ed.gov/ncee/edlabs/projects/project.asp?ProjectID=404; or Shakman, K., & Rodriguez, S. M. (2015). *Logic models for program design, implementation, and evaluation: Workshop toolkit* (REL 2015–057). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Northeast & Islands. Retrieved from https://ies.ed.gov/ncee/edlabs/projects/project.asp?ProjectID=401.

Box 9. An example of the Complexity of Describing Constructs

How many eighth grade English teachers are in your schools?

Component	Issues to be clarified
How many	Does "how many" refer to a head count or full-time equivalent (FTE) count?
Eighth grade	Does "eighth grade" include classes with seventh-, eighth-, and ninth-grade students; or just classes with only eighth graders?
English	Does "English" include reading and writing classes? Special education English language classes? Other language arts classes? English as a Second Language classes?
Teachers	Do "teachers" include only certified teachers? Only certified English teachers? Certified teaching assistants? Only teachers assigned to teach classes/students this grading period?
Are	At what point in time should the answer be valid? At the beginning or end of the current or previous school year?
In	Does this include teachers of students cross-enrolled in virtual settings? What if someone teaches English in more than one school—does he or she get counted more than once?
Your	Does this mean only schools under the authority of the state or local education agency, or does it include all schools within the boundaries of the state or locality?
Schools	Are special education schools included? Correctional institutions that grant educational degrees? Other residential facilities? Cross-enrolled virtual settings? Private schools?

How many eighth-grade English teachers are in your schools? This "simple" question illustrates the complexity of the real world and the importance of clearly conceptualizing and precisely measuring a research construct. On one end of the spectrum, there may not be any full-time certified English teachers teaching an English class to only eighth-grade students in the single middle school in a school district this semester. At the same time, 50 or more full- or part-time teachers may be leading reading, writing, or language classes with at least one eighth-grade student at some point during the academic year. Clearly, the "right" answer depends on the context of the question.[28]

Measures

Measures should be valid and reliable. A valid measure captures the concept of interest. For example, we might aim to measure how engaged and active parents are in their child's school. We might consider measuring this engagement with the number of times a teacher reports that he or she has interacted with the parents at school. If the teacher never initiated the meetings, this measure might be a valid measure of engagement. However, if most meetings are initiated by teachers, for example in response to student misbehavior, then the number of interactions would not be a valid measure of the construct of interest, instead measuring student misbehavior. A valid measure, thus measures the concept of interest. A reliable measure captures the construct of interest accurately, not just on average. For example, we might be interested in students' vocabulary. Asking each student all possible words is infeasible. Measuring overall vocabulary by asking students to report the meaning of only one or two words, is more feasible but not at all accurate. Some students will know many words but not the specific ones chosen, while others will know very few but just happened to know the

[28] National Forum on Education Statistics. (2005). *Forum guide to metadata: The meaning behind education data* (NFES 2009–805). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. Retrieved from http://nces.ed.gov/forum/pub_2009805.asp.

ones chosen. A reliable measure will capture the underlying construct with some accuracy, though almost all measures have some error.[29]

Once a researcher has a definition of the construct to be described that he or she can justify explicitly, it is necessary to operationalize the construct and determine how it can be measured. Different research questions call for, and different researchers prefer to use, different types of data. Sometimes available datasets are sparse, and the researcher is left with little choice but to use what is accessible (assuming that it is relevant). At other times, datasets are rich, or the researcher will be conducting an original (primary) data collection and will need to refine and distill data and methods to generate the measures of interest. In either case, the researcher must choose the measures. This decision-making should integrate the researcher's intuition, an awareness of prior research, and an examination of available data and data collection opportunities.

Some measures of interest simply may not be available in a dataset or may be too burdensome in terms of effort or cost to collect. For example, a dataset with measures of observed teaching practice on a large scale has only recently become accessible.[30] Prior to the availability of this dataset, collecting such data was beyond the capacity of most individual researchers.

In the absence of ideal data—an absence that almost always occurs—researchers must rely on proxy data, which refer to data that relate to, but are not a direct measure of, a construct. Proxy data are, by definition, imperfect, but by necessity, are commonly used when they are a reasonable substitute or approximation. Many descriptive studies use multiple proxies for measuring a construct of interest. When this occurs, each proxy may be imperfect, but together they can provide more robust insight into the phenomenon. However, when using proxy data, the researcher should describe (1) why the proxies are relevant for answering the research question and (2) their imperfections as measures of the underlying construct. If proxy measures have strong construct validity—that is, they capture the underlying concept or process of interest—then they can be used to answer the research questions. If not, they are not valid and cannot be used. For example, consider a researcher interested in measuring parent involvement in school. He or she may have access to the number of times a parent entered the school, but this data might not be a valid proxy because it includes the times parents are called in to deal with student misbehavior. A measure limited to the number of times a parent volunteers to help at school, attends school-wide meetings, or participates in PTA events is likely to be a far more valid, though still imperfect, proxy of parent involvement.

In addition to construct validity, measures must be sufficiently reliable for their intended purpose. Some measures may be appropriate for capturing an underlying construct, on average, but may be too imprecise to adequately measure other changes or differences. If the purpose of the measure is to assess change over time or differences between groups, then an imprecise or time-limited measure is not sufficiently reliable, even though it may be acceptable for other purposes. Concern about the precision of a measure is particularly relevant when dealing with small samples or trend analyses.

---

[29] Many textbooks and articles cover these and other aspects of measurement. See, for example, Crocker, L., & Algina, J. (2006). *Introduction to classical and modern test theory.* Belmont, CA: Wadsworth Publishing Co.; Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational measurement*, 3rd ed. (pp. 13–103). New York: Macmillan; and Feldt, L. S., & Brennan, R. L. (1989). Reliability. In R. L. Linn (Ed.), *Educational measurement,* 3rd ed. (pp. 105–146). New York: Macmillan.

[30] The Measures of Effective Teaching Longitudinal Database (MET LDB), available at http://www.icpsr.umich.edu/icpsrweb/METLDB/.

For example, although common measures of improvement (change) in students' test scores may describe school effectiveness, on average, over time, these "value-added" measures may not be sufficiently precise for gauging individual school performance, especially those schools with smaller student populations.

Samples

Motivation and availability, the two driving factors behind determining *constructs* and *measures*, are also important when choosing samples. For example, when research is motivated by the purpose of improving our understanding of how well schools are serving America's students, relevant graduation-rate data might reflect a sample that is restricted to individuals who attended schools in the United States. If, instead, the researcher's interest in graduation rates is driven by an effort to assess skills available in the job market, it will make sense to include all individuals in the U.S. job market, regardless of where they attended school.

Researchers often are constrained by their choice of samples—either because they are using secondary data (that has already been collected and may or may not be entirely relevant) or because they are collecting primary data but have limited time, funding, or access to respondents and collection systems. As is the case with *constructs* and proxy *measures*, good research explicitly and thoroughly explains both the strength and limitations of a sample as they relate specifically to the study at hand. A relationship clearly observed in a sample (*internal validity*) may or may not be observed in the population of interest (*external validity*) depending on how representative the sample is of the population.

As a rule of thumb, researchers can answer their research questions best when they use as much data as possible—as long as those data are of good quality. Incorporating irrelevant or low-quality data only detracts from validity. But when multiple sources of good data provide independent measures of the same construct, agreement increases the credibility of the results, and disagreement suggests (1) the need for caution when interpreting findings and (2) a call for additional research (or thought) about why different data point to different understandings.

Good descriptive research relies primarily on low-inference, low-assumption methods that use no or minimal statistical adjustments.

Graphical methods and more complicated statistical adjustments have their place in descriptive analyses as well but need to be used with caution.

Using Data to Answer Research Questions

Good descriptive research relies primarily on low-inference, low-assumption methods that use no or minimal statistical adjustments.[31] Therefore, measures of central tendency (such as mean, median, and mode), measures of variation (such as range and standard deviation), and basic frequency analyses are particularly useful statistical tools for description.

[31] We use the term "statistical adjustments" to describe methods, such as regression analysis, that can separate the variation in one variable that overlaps with variation in other variables. For example, an "adjusted" measure of height might not include the variation in height that is related to age or gender and, instead, provide a measure of how tall a person is relative to other people of the same age and gender. Regression analysis can be used to create such a measure.

Measures of central tendency are simple and compelling ways to describe and compare measures of interest but are best used in combination with measures of variation. For example, it may be true that the average scores for black students on a particular assessment are lower than the average scores for white students, but this statement masks the fact that many black students score higher than many white students.[32] Thus, the use of central tendencies, without corresponding measures of variation, is susceptible to not accurately characterizing the data, measure, construct, question, and phenomenon. Good descriptive analyses also incorporate assessments of heterogeneity—by location, time, institution, etc.—when presenting central tendency and variation. The black-white test score gap, for example, varies across time, place, grade level, and test subject.[33] Descriptive research can incorporate other types of statistical methods, some of which are discussed below, but more complex treatments often present challenges and should be used with caution.

Statistical Adjustments

Statistical adjustments are appropriate and helpful when a researcher tries to ensure that measures are comparable. Suppose, for example, that test scores from one group were collected at ages 8–10 and from another group at ages 9–11. A researcher might use a statistical model to adjust the average scores in each group to account for the differences in ages of the samples.

However, statistical adjustments—particularly those relying on regression models with multiple variables and those that rely heavily on linearity or other assumptions—can present challenges to researchers because these complex adjustments can mask underlying relationships in the data. For example, if a researcher is interested in racial or ethnic differences across groups but includes controls for income and residential characteristics, such adjustments might mask actual differences in the groups that reflect economic or social segregation.

Much of what a researcher may wish to do with regression can also be accomplished by calculating unadjusted statistics (such as averages) separately by groups or by plotting data in order to visualize full distributions and bivariate relationships (instead of relying on statistical adjustments). The benefit of simpler approaches is that they do not require assumptions that are common in regression-based approaches. Often, more controls and more assumptions about the structure of relationships between variables (for example, linearity) do not improve analysis, and, in fact, they frequently cloud our understanding of what is really happening in the data and the real world. Thus, researchers can use statistical adjustments to more clearly identify the phenomena in question, but need to be careful that the adjustments are, in fact, clarifying and not muddling the relationships in question.

---

[32] Stark, P., & Noel, A. M. (2015). *Trends in high school dropout and completion rates in the United States: 1972–2012* (NCES 2015–015). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. Retrieved July 2015 from http://nces.ed.gov/pubs2015/2015015.pdf.

[33] Vanneman, A., Hamilton, L., Anderson, J. B., & Rahman, T. (2009). *Achievement gaps: How black and white students in public schools perform in mathematics and reading on the National Assessment of Educational Progress* (NCES 2009–455). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics.

Comparisons

Much descriptive research is comparative—that is, it compares data across time, place, and population or group. Comparative research is essential for understanding the world but has some common difficulties that researchers should work to overcome.

One common problem with comparative study is that constructs of interest may change over time or may differ across places or populations. For example, a researcher who wanted to measure the disparity in education outcomes between children of parents with higher and lower levels of educational attainment over the course of the twentieth century would have to consider the fact that the characteristics that represent high and low levels of education have changed dramatically over the 100-year window of study. Thus, a comparison between the outcomes of children whose parents have less than a high school degree and those whose parents have a college degree would reflect two very different populations in 2015 from what it would have 30 or 50 or 80 years ago. This problem is a conceptual one—if the social or education definition of categories such as "educational attainment," "race," "wealth," or "achievement" change over time, this loss of consistency means that the underlying trends in these categories will be blurred by changes in definition.

Another common problem when conducting comparative studies is the realization that data from different times, places, or populations are not always collected in the same way. One survey may have collected data about both parents' highest academic degree earned; another about the number of years of education either parent received; and a third about the highest degree in a household by any resident adult. This type of data problem can sometimes be solved but often proves to be insurmountable without the use of innovative methods and statistical approaches (see Box 10). Thus, while simple techniques are usually preferable, more complicated techniques may be necessary to overcome data limitations.

Box 10. Example of Descriptive Research that Compares Academic Achievement Gaps by Socioeconomic Status over Time

Reardon, S. F. (2011). The widening academic achievement gap between the rich and the poor: New evidence and possible explanations. In R. Murnane & G. Duncan (Eds.), *Whither opportunity? Rising inequality and the uncertain life chances of low-income children*. New York: Russell Sage Foundation Press.

Reardon (2011) set out to study academic achievement gaps by socioeconomic status across time.

Unfortunately, some measures of socioeconomic status, like parental education and occupational status, are difficult to compare over time, and the relative value of income varies over time as well. But Reardon noticed that many rank-based measures of income-achievement gaps have a common interpretation over time. Thus, the gap in test scores between children whose family income is at the tenth and ninetieth percentiles of the (contemporaneous) income distribution was a well-defined construct.

A second comparability concern facing Reardon was that income was not measured the same way across the 12 studies (datasets) available to him. In some cases, income was reported by students; more often, it was reported by parents. Regardless of the source, in most of the studies, income had been reported in a set of 5–16 ordered categories that were not consistent over time. Reardon took two steps to improve the comparability of these measures: (1) he smoothed the category-based income-achievement association in order to estimate average achievement levels at the tenth and ninetieth percentiles; and (2) he adjusted the resulting measures to account for the fact that student-reported income is generally less reliable than parent-reported income.

By creating a construct whose interpretation was constant over time, and by using differently collected data to estimate this common parameter, Reardon reported plausibly comparable measures of income-achievement gaps over time, which is useful across a broad spectrum of research- and policy-related work.

Groupings, Networks, and Clusters

While much descriptive work relies heavily on central tendencies and variation, other analytical methods are emerging to identify and explore more recently observed phenomena, such as networks, clustering, and other interactions within education organizations and interpersonal relationships. In the context of education, peer networks sometimes refer to the phenomenon of students interacting with other students to attain education goals.[34] For example, peer groups of students may tend to organize themselves into certain course-taking patterns.[35] Clusters, on the other hand, are groups of units (such as students, teachers, classrooms, or schools) not that interact with each other but that have similar traits. Members of a cluster share multiple characteristics, such as personality traits or learning behaviors, which distinguish them from other clusters.[36]

Network analysis and cluster analysis are two common approaches for studying grouping phenomena (see Box 11). Network analysis describes systems of links among people or other subjects of study, whereas cluster analysis describes sets of actors with the same characteristics, whether they are courses or behavioral indexes.[37] Descriptive analysis can contribute to the identification and depiction of these types of patterns from within large sets of data that measure individual dynamics and interactions.[38] Once better understood, such complex social and organizational patterns may shed light on the emergence of skill gaps and other achievement outcomes, as well as on possible points for productive interventions in the social and interpersonal associations of students and other individuals in a classroom or school setting.

Box 11. Example of Descriptive Research that Uses Network and Cluster Analysis as Descriptive Tools

Daly, A. J., & Finnigan, K. (2011). The ebb and flow of social network ties between district leaders under high stakes accountability. *American Education Research Journal, 48*(1): 39–79.

This study examined the underlying social networks of a district leadership team engaged in systemic reform efforts in response to multiple schools being designated as “in need of improvement” under the No Child Left Behind (NCLB) Act of 2001. Using a survey to collect data on social networking relationships between central office administrators and school building leaders, the researchers developed a longitudinal case study focused

---

[34] O'Donnell, A. M., & King, A. (1999). *Cognitive perspectives on peer learning*. New York: Routledge.

[35] McFarland, D. A. (2006). Curricular flows: Trajectories, turning points, and assignment criteria in high school math careers. *Sociology of Education, 79*(3): 177–205.

[36] See, as examples, for more information on network analysis: Kadushin, C. (2012). *Understanding social networks*. Oxford, UK: Oxford University Press; Knoke, D., & Yang, S. (2008). *Social network analysis,* 2nd ed. Los Angeles: Sage Publications; Prell, C. (2011). *Social network analysis: History, theory and methodology,* 3rd ed. Los Angeles: Sage Publications; Scott, J. (2012). *Social network analysis*. Los Angeles: Sage Publications; and Wasserman, S., & Faust, K. (1994). *Social network analysis: Methods and applications*. Cambridge, UK: Cambridge University Press. See, as examples, for more information on cluster analysis: Aldenderfer, M. S., & Blashfield, R. K. (1984). *Cluster analysis* (Quantitative Applications in the Social Sciences). Los Angeles: Sage Publications; Everitt, B. S., Landau, S., Leese, M., & Stahl, D. (2011). *Cluster analysis,* 5th ed. New York: Wiley; and King, R. S. (2014). *Cluster analysis and data mining: An introduction*. Dulles, VA: Mercury Learning & Information.

[37] Wasserman, S., & Faust, K. (1994). *Social network analysis: Methods and applications*. Cambridge, UK: Cambridge University Press.

[38] Rawlings, C. M., & McFarland, D. A. (2010). Influence flows in the academy: Using affiliation networks to assess peer effects among researchers. *Social Science Research, 40*(3): 1001–1017.

on three reform-related networks: knowledge, advice, and innovation. The findings suggested that the networks became more centralized over time, which may have limited the transfer of knowledge, advice, and innovation and impeded the effectiveness of reform efforts. The descriptive network analysis provided practical insight into which individuals "moved" knowledge and practice throughout the system. From a policy perspective, the findings showed that successful reform activities were associated with intentional partnerships (network connections) among central office staff and onsite school administrators.

Cautions Regarding Uncertainty and Fishing

All sampling methods introduce some error and uncertainty, but many other potential sources of uncertainty affect data collection and analysis, including missing data, participant attrition, and variation in the delivery of an intervention (that is, implementation fidelity). Reporting standard errors and other descriptions of uncertainty (for example, evidence of clustering or dependence in the data) influence how data are analyzed and findings are interpreted and authors should make these available to audiences in either raw form (for researchers) or descriptive form (for practitioners and policymakers).

In today's era of big data, "fishing" is a concern in both descriptive and causal analyses. Fishing refers to analysis in which the primary motivation is to look for any and all patterns in a dataset without the boundaries of preconceived hypotheses, research questions, and constructs. If an investigator tests for any and all possible differences between two groups for which a large set of data is available, it is possible (and even likely) that some patterns may be observed, even if they are without merit and will not hold for the population more broadly. Researchers may search for "statistically significant" relationships and rely too heavily on statistical significance to signal an important result.[39] Statistical significance, however, is only a measure of how unlikely a relationship is to exist in a population given the size of the observed relationship in the sample. If researchers search for too many "unlikely" relationships in a sample, they are bound to find one that looks as if it is big enough to exist in the population, even if it does not. Limiting the number of allowable comparisons to be tested buffers against this possibility.[40] Thorough reporting of research methods, including the number of statistical tests, also reduces the temptation to search for and focus on statistical significance.[41]

[39] Ioannidis, J. P. A. (2005). Why most published research findings are false. *PLoS Medicine, 2*(8): e124.

[40] A healthy debate considers the role of "exploring the data" (or more pejoratively "fishing" or "cherry-picking results") when it comes to descriptive analysis. Unlike causal research that evaluates whether evidence is consistent with a specific hypothesis, descriptive analysis is intended to describe what is real in the world based on data. When a researcher reviews data for descriptive purposes, previously unseen patterns in the data sometimes arise. The fact that the analysis was not designed to uncover that truth does not detract from the value of identifying it. There is latitude for "exploring the data" in descriptive analysis, subject to the caveat that such "findings" should be consistent with sound theory and rigorous methods. It may be appropriate to use these findings as the foundation for generating hypotheses, prioritizing causal mechanisms, or otherwise pointing toward causal understanding, but not as evidence for causation.

[41] For a discussion of these difficulties and a 20-word solution to the problem of transparency around methods and results of quantitative studies, see Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science, 22*(11), 1359–1366. Retrieved from http://journals.sagepub.com/doi/10.1177/0956797611417632.

In contrast to fishing, data "mining" can be a productive avenue for description *if approached cautiously*. Data mining utilizes every observation and feature (even those that lack a basis in a concept) for the purpose of describing what is occurring in the world. Online or virtual education settings specifically, and education information systems more generally, provide rich opportunities for data mining as the data capture complex interactions at both large and detailed scale. Though traditional data mining techniques are often not appropriate for some types of issues and characteristics relevant to education,[42] data mining can lead to the identification of new phenomena or hidden patterns in data that had not previously been recognized. There can be no claim of causation with data-mined findings, but they can serve as foundations for generating hypotheses, prioritizing possible causal mechanisms, or otherwise pointing toward causal understanding.

Conducting Descriptive Analysis: Summary

In education research, a robust study question is socially important and highly relevant to improving our understanding of education processes, distribution (access), effects, and quality. With a research question in hand, the researcher must identify the key ideas, attributes, or concepts (constructs) that will be measured. Existing datasets can be appropriate for some aspects of a study. When they are not, the researcher may choose to conduct a custom data collection or rely on proxy data, which are imperfect but commonly used when they are a reasonable substitute or approximation. Data mining can be a productive approach for description if approached cautiously.

Good descriptive research relies primarily on low-inference, low-assumption methods that use no or minimal statistical adjustments. Measures of central tendency, variation, and basic frequency analyses are particularly useful tools. Although there can be no claim of causation, any descriptive finding that uncovers a socially relevant "truth" in the data can serve as the foundation for generating hypotheses, prioritizing possible causal mechanisms, or otherwise pointing toward causal understanding.

[42] Romero, C., & Ventura, S. (2013). Data mining in education. *WIREs Data Mining and Knowledge Discovery, 3*: 12–27.

Chapter 4. Communicating Descriptive Analysis

A great book has literary value only when it can be read and understood. The same is true for great research studies: no matter how significant the findings, they contribute to knowledge and practice only when others read and understand the conclusions.

A researcher could communicate findings simply by displaying all available raw data so that readers are able to manipulate and interpret it in any way they prefer. While possible, this approach is clearly less than ideal: presenting a large amount of data without description increases the likelihood of misinterpretation by the audience.

The researcher's job and expertise is to reduce the body of data to a format that is useful for the audience. This translation, from raw to reported findings, is undertaken specifically to meet the information needs of the readers.

Communicating Data Visually

Most researchers understand the utility of statistics to reduce raw data into information that is meaningful to an audience (for example, presenting the mean, median, mode, and other measures of central tendencies for a large dataset). Data visualization, or the graphical display of information, is another useful tool for communicating research findings.[43]

Data visualization is part of the research and the communications processes—helping both the researcher and the reader identify patterns in the data.

Because data are generally more easily processed and understood when they are visualized, graphs can help both researchers and their audiences identify and understand patterns in data. In the process of descriptive research, data visualization can also clarify patterns or problems in the data, which may or may not be shared with the audience.

When considering visualization, it is helpful to focus on three key features: function, familiarity, and simplicity.

- **Function.** Above all other considerations, graphics should accurately reflect the data and appropriately illustrate the story of your findings. The overarching message that you are trying to communicate should be clear. Before including a figure, always consider whether efforts to visualize the data have increased the risk of a reader misunderstanding or misinterpreting the information.
  Visualizing data typically requires more space than presenting the same information in a table. Because of this reality, figures should provide benefits beyond what tables can offer. The tradeoff between space and information will need to be evaluated in light of data type, physical constraints, and audience need.
- **Familiarity.** People tend to process information more quickly when it is presented in a familiar manner, so visualization choices often depend on how an audience expects to see a

[43] Few, S. (2014). Data visualization for human perception. In M. Soegaard & R. F. Dam (Eds.), *The encyclopedia of human-computer interaction*, 2nd ed. Aarhus, Denmark: The Interaction Design Foundation. Retrieved March 2015 from https://www.interactiondesign.org/encyclopedia/data_visualization_for_human_perception.html.

particular type of data presented. If there is a standard way of presenting the data, a researcher needs a strong reason to depart from the norm. For example, time trends are often presented as line graphs with time on the x-axis and the variable of interest on the y-axis. Readers are accustomed to this format and, as a rule, can readily interpret these types of visualizations.

There are times, however, when the frequently recognized mode of presentation is so ineffective that its costs outweigh the benefits of its familiarity with audiences. The pie chart is a prime example: while familiar, pie charts are a cumbersome approach to presenting simple information and make it difficult for a reader to accurately perceive meaning in the data.

- **Simplicity.** Good communicators tend to recognize the benefits of simplicity. Sometimes a researcher may wish to add visual interest to a graph in order to make a report appear more attractive or noticeable. In all circumstances, cosmetic upgrades should be secondary to the goal of effective communication. The use of pictures, colors, extraneous symbols, or other imagery often serves only to distract from the fundamental message that needs to be conveyed. Data visualization should *always* preserve the integrity of the message, and efforts to visualize data should *never* detract from accuracy or integrity.

All efforts to improve the visual appeal of data should also improve the accuracy or integrity of the message.

The Process of Communicating the Message

Like other aspects of descriptive analysis, choosing how to communicate findings is best done iteratively. Begin by deciding what the story is (the *message*) and to whom it will be told (the *audience*). Only then can the researcher make sound decisions about the most effective way to communicate the story (*customization*).

Often, the process of customizing the message for an intended audience results in presentations that differ by audience type. A report published in a peer-reviewed academic journal will likely look different from the same research story written for a policymaker or practitioner audience. The research audience may be used to graphical representations with numerous axes and multiple statistics (for example, standard deviations as well as averages). The practitioner audience, on the other hand, may expect fewer but more focused graphics that tell the story more succinctly. In each case, the complementary final products contain only a subset of the analysis undertaken by the researcher. The research principles, standards, and methods do not change, but because of varying information needs, the "final" presentation is modified for intended audiences.

What you do to develop your conclusions and what you present to communicate your message may be different.

How to Frame Visualization Needs

Entire books have been written about communications, report writing, research methods, and data visualization (see appendix A). They are worth reading and understanding. While we cannot go into such detail here, some of the principles driving the process of customizing a research message for a particular audience can be distilled into the following four questions.

Core Message: What is your story (the message)?

- Carefully consider the motivation for the research and the significance of the findings (for example, their ramifications on education practice, policies, or research). Doing so will help to prioritize the features of the message that need to be effectively highlighted and accurately communicated.
- Emphasize the most important patterns in the data for answering your research question. Limit presentation to those data that are clearly focused (on your specific research question and findings) rather than to a range of facts only tangentially related to the core message.

Potential Audiences: Who would benefit from hearing this story (the audience)?

- Identify your intended audience(s). One approach for doing so is simply to ask yourself (or knowledgeable colleagues) which audiences would benefit from learning about your findings. Before you narrow your targets, reflect broadly on potentially interested parties, including researchers in other disciplines, teachers, administrators, policymakers (local, state, and federal), leaders of national organizations, and policy advocates.

Customizing the Message: For each target audience, what are the audience's specific information needs and how can you most effectively communicate your story (customization)?

- Assess which aspects of your findings are most relevant to each specific audience. For example, a practitioner or policy audience may be interested in a newly diagnosed problem that requires their attention. Alternatively, a researcher audience may need to understand your data sources and methods so that they can plan a causal study that grows out of your findings.
- Consider other information that might be helpful to frame or contextualize your story for each audience. Existing research, previously recognized problems, or opportunities facing practitioners, and contemporary policy debate can be meaningful to the discussion.
- Determine what type of presentation will best reach each of your intended audiences—knowing that each audience will likely need a slightly or substantially different presentation of your story to correctly interpret your message. Identify the types of visualizations that each audience expects for each type of data and determine whether there is a standard (familiar) way of presenting a particular type of data. While customizing communications takes effort, recall that your work contributes to knowledge and practice only when others read and understand your conclusions.
- Choose the publication mechanism that is most appropriate for each audience. For example, you may wish to target a particular academic journal for education researchers, a different journal for economists, a third for psychologists, a public policy quarterly for state legislators, and a practitioner-friendly periodical for school principals and teachers.[44]

Note that different audiences have different information needs.

Some audiences require details about methods and statistical significance, while others benefit from summary findings that focus on a "take home" message.

[44] While producing separate, customized documents for specific audiences may be the ideal form of communication, some publications are able to reach multiple audiences. For example, within a single document,

Iterate: How can you improve efforts to communicate your message to each specific target audience?

- Try different approaches to visualization and presentation. Share a draft with a representative of your target audience. Even a single voice can help to confirm or redirect your efforts.
- Ask yourself whether you can further simplify the presentation. First drafts are often more complicated than they need to be. Show as much as you need to make the point, and no more. Be realistic about how much data your audience can absorb in one figure.

The product of your efforts to visualize your data should be evaluated against the three key features of data visualization described in preceding text: function, familiarity, and simplicity:

- *Function*. Does the resulting graphic accurately reflect the data and the overarching message you are trying to communicate? Have efforts to visualize the data increased the risk of misunderstanding or misinterpreting the information?
- *Familiarity*. Will the graphical approach be familiar to your audience? Do the benefits of an innovative approach to visualization outweigh the cost of novelty?
- *Simplicity*. Can you further simplify the graphic? As a rule, simpler is better. Purely cosmetic changes for the purpose of increasing visual appeal rarely increase accuracy or reduce complexity and, therefore, are not recommended.

In addition, all figures should stand alone and be easy to read when printed.

- *Stand-alone*. Can the information contained in the graphic (including the caption) be understood completely without additional explanation? If not, given how frequently audiences refer to images without reading associated text, what should be added to the image to construct a stand-alone piece of information?
- *Color and Printability*. A substantial portion of your audience prints reports *only* in black and white. Moreover, nearly 8.0 percent of men and 0.5 percent of women of Northern European ancestry have some version of red-green color blindness.[45] Make your images distinguishable on the basis of contrast rather than color, and confirm that they are distinguishable by printing them without color to ensure that your graphics are distinct in gray scale and in black and white.

Common Approaches to Data Visualization

Researchers have many options when considering how to approach data visualization to best meet the information needs of their audience(s). After all, the same statistical fact can be expressed in multiple ways, such as a table, line graph, or scatter graph (see Box 12). For example, a conditional mean may be communicated by (1) running a regression that includes indicator variables and their interactions; (2) showing subgroup means in a table; or (3) plotting subgroup means in a graph. For the vast majority of readers, tables and graphs (options 2 and 3) are a more transparent and easily interpreted way to describe data.

---

an executive summary can target policymaking audiences in need of high-level messaging, main text reaches a general audience, and appendices provide data and methodological details for a researcher audience.

[45] American Academy of Ophthalmology. (2011). *Genetics home reference: Color vision deficiency.* Bethesda, MD: U.S. Department of Health and Human Services, National Institutes of Health, National Library of Medicine. Retrieved March 2015 from http://ghr.nlm.nih.gov/condition/color-vision-deficiency.

Box 12. Visualization as Data Simplification

TABLE 1. Science and engineering research space in academic institutions, by field: FY 1988–2003

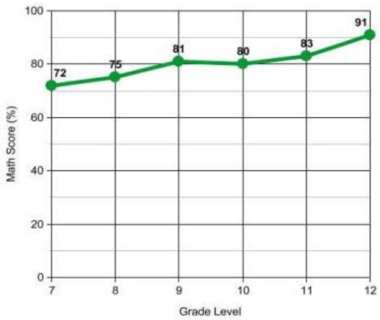
(Net assignable square feet in millions)

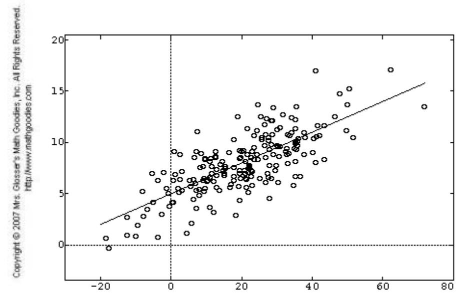
Field	1988	1990	1992	1994	1996	1998	1999	2001	2003
All fields	112	116	121	127	136	143	148	155.1	172.6
Agricultural sciences	18	21	20	20	22	25	24	26.7	26.4
Biological sciences	24	27	28	28	30	31	31	33.4	36.0
Computer sciences	1	1	2	2	2	2	2	2.4	3.1
Earth, atmospheric, and ocean sciences	6	6	7	7	7	8	8	8.1	8.9
Engineering	16	17	21	21	22	23	24	25.5	27.4
Mathematics	1	1	1	1	1	1	1	1.0	1.5
Medical sciences	19	20	23	23	25	25	26	27.8	34.9
Physical sciences	16	16	17	17	18	18	19	19.2	20.4
Psychology	3	3	na	3	3	3	4	3.6	4.4
Social sciences	3	3	na	3	4	5	3	4.5	5.7
Other sciences	4	2	2	2	2	3	3	3.0	3.8
Animal research space	na	na	na	11	12	12	13	na	16.7

na = not available.

NOTES: Details may not add to totals due to rounding. Animal research space is included in the space totals for individual fields where appropriate.

SOURCES: National Science Foundation/Division of Science Resources Statistics, Survey of Science and Engineering Research Facilities, Fiscal Years 1988-2003.





The same statistical facts can usually be expressed with a table, in a graph, or via regression, but tables and graphs are a much more transparent way to describe findings for most audiences.

Tables

Tables tend to be most effective as a presentation format when displaying only a few data points (values), presenting secondary findings, or reporting standard errors, confidence intervals, and significance levels. In these cases, tables may be sufficient for the audience and will likely require less space than a graph.

Tables may also be preferable on the rare occasions when a researcher needs to present a large amount of data so that the reader will be able to replicate the work or match the data to a specific source. However, this goal can also be addressed by presenting data and source notes in an appendix. Using an appendix to share detailed notes about sources, values, and methods is an effective approach to communicating the message because it visually displays data in the text, yet preserves the information needed for replicability and transparency for those readers who are interested.

When presenting data in tables, consider the following tips:

- Be thoughtful about table titles. A good title will convey meaningful information to your *specific* audience.
    - For a researcher audience, the title should describe the specific information presented in the table (for example, "Table 3. Predictors of court involvement and school segregation, 1961").
    - For a policy or practitioner audience, consider using a title that summarizes the main point of the table. If your data support a specific message, the title could convey that message, such as "Urbanicity was the most important predictor of court involvement and school segregation in 1961." This strategy, if undertaken, has to be done carefully and accurately so as not to be misleading. Sometimes findings are more complicated than a title can convey accurately.

For the vast majority of audiences, graphs and tables are much more easily understood than regression coefficients and other complex statistical presentations.

- Assign a meaningful descriptive title to each variable presented in the table (such as, First Grade Students, Second Grade Students, etc., rather than Group 1, Group 2, etc.). Then group the variables thematically or logically in your table, rather than alphabetically or randomly. An exception to this is a long list of variables that do not reflect a theme or logic, such as a list of state names.

- Include detailed notes in a caption so that your table is self-contained. This gives the viewer the option to "just read the table," yet still view all of the information needed to understand your message.

Graphs

As a rule of thumb, any data that can be accurately presented in graphic form should be presented in graphic form. The goal of visualization, after all, is to help readers better identify and understand important patterns in the data—and graphs generally are more easily processed and understood than tables.

Design choices should be based on meaningful criteria, such as data type, message, and audience need. Unfortunately, too many published graphs appear to be designed haphazardly, as though what is easiest to do in a spreadsheet program will meet audience need regardless of data type and message (hence the abundance of apparently arbitrarily constructed pie charts, area graphs, bar and column charts, scatter plots, and line graphs). Other than pie charts, these types of graphs can be suitable visualization choices, but not because they are easy to create in a spreadsheet or analytical application.

Space permitting, and with rare exception, whatever data can be presented graphically should be presented graphically.

Reviewing all graph types is far beyond the scope of this project, but a few examples help illustrate the range of options.[46] It is important to present as much content as necessary for a graphic to fully meet the information needs of the audience and, as practical, to stand alone as a complete piece of information. To varying degrees, the following examples include descriptive titles, source notes, and explanations of graphed data in captions to help ensure that the figure stands alone as a useful communication tool.

Time trend data: Line graphs are appropriate for continuous data and are an effective way to present and compare slopes (or levels, if the y axis starts at zero). For example, figure 1 effectively illustrates how line graphs can be used to demonstrate differences in time trends (for characteristics of three groups of teachers over 35 years).

---

[46] For more information on visual presentations, see Cleveland, W. (1994). *The elements of graphing data*, 2nd ed. Summit, NJ: Hobart Press; Few, S. (2009). *Now you see it: Simple visualization techniques for quantitative analysis*. Oakland, CA: Analytics Press; Tufte, E. R. (2001). *The visual display of quantitative information,* 2nd ed. Cheshire, CT: Graphics Press; and Wong, D. M. (2010). *The Wall Street Journal guide to information graphics: The dos and don'ts of presenting data, facts, and figures.* New York: W. W. Norton & Company, Inc. A number of statistical packages have graphing capabilities. Stata, for example, includes numerous graphing options, as described in Mitchell, M. N. (2012). *A visual guide to Stata graphics*, 3rd ed. College Station, TX: Stata Press. A number of online resources are also available including, among others: http://www.stata.com/support/faqs/graphics/gph/stata-graphs/ http://www.ats.ucla.edu/stat/stata/library/GraphExamples/ and http://data.princeton.edu/stata/graphics.html.

Figure 1. Line graphs showing time trends for three groups of teachers.

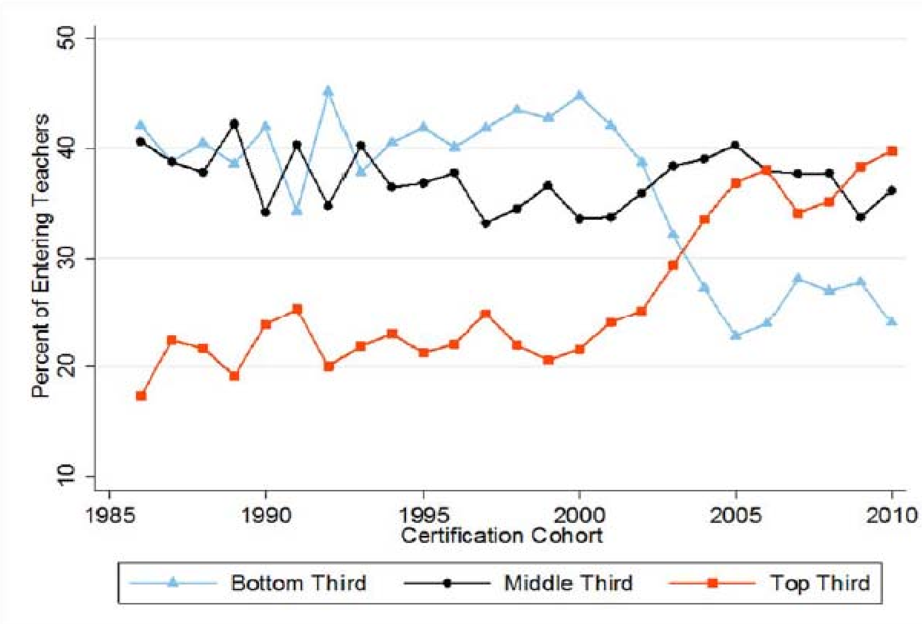


FIGURE 3. *Percent of entering teachers in New York City schools drawn from the bottom, middle, and top thirds of the statewide score distribution by certification cohort, 1986 to 2010*

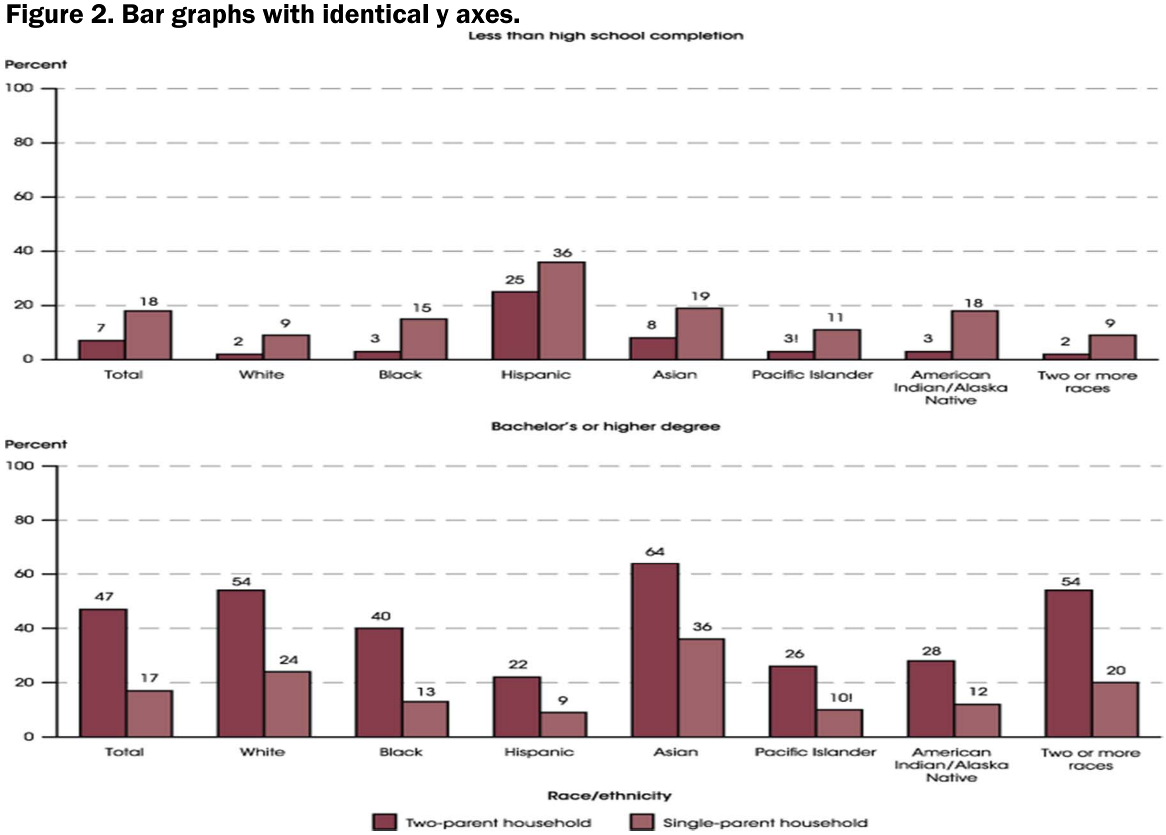
Source: Lankford, H., Loeb, S., McEachin, A., Miller, L. C., & Wyckoff, J. (2014). Who enters teaching? Encouraging evidence that the status of teaching is improving. *Educational Researcher, 43*(9), 444–453.

Discrete data with frequencies or proportions: Bar graphs are appropriate for discrete data and can be used to distinguish magnitude and proportion. Each bar provides the value for a group and a comparison across bars yields information on relative size. When using bar graphs to communicate proportions, the y-axes should be identical so that any differences in the heights of the bars can be accurately compared (see figure 2).

While it is often appropriate for bar graphs to show the entire y-axis (0–100), doing so sometimes results in so much white space that it is difficult to see data values. In all cases, however, when comparing values across more than one figure, axes should be identical to ensure accurate representation.

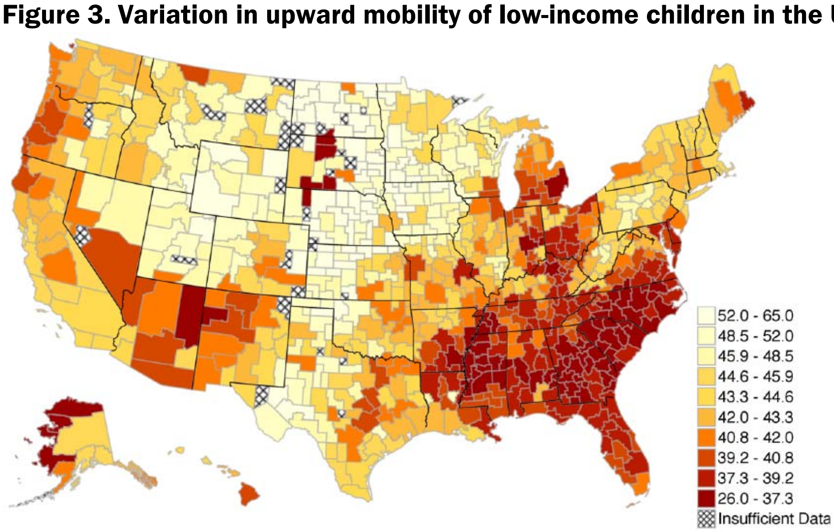
Geographic data: Maps can be useful for describing geographic variation. For example, figure 3 shows variation in the upward mobility of low-income children across the United States. The darker colors stand out to illustrate the low social mobility in those geographic regions and serve as an example of how researchers can use visualization tools to creatively present findings in a clear, audience-friendly manner.

Figure 2. Bar graphs with identical y axes.



Source: https://nces.ed.gov/programs/coe/indicator_saa.asp.

Figure 3. Variation in upward mobility of low-income children in the United States

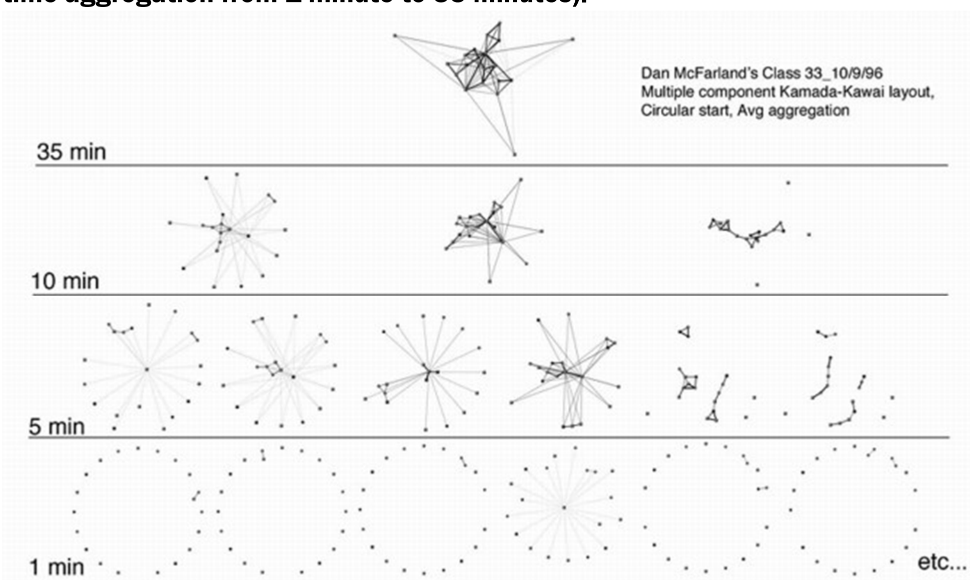


Note: This figure presents a heat map of measures of intergenerational mobility by commuting zone. It is based on 1980–82 birth cohorts. Children are assigned to commuting zones based on the location of their parents (when the child was claimed as a dependent), irrespective of where they live as adults. In each commuting zone, child income rank is regressed on a constant and parent income rank. Using the regression estimates, absolute upward mobility is defined as the predicted child rank given parent income at the 25th percentile. The maps are constructed by grouping commuting zones into ten deciles and shading the areas so that lighter colors correspond to higher absolute mobility.

Source: Chetty, R., Hendren, N., Kline, P., & Saez, E. (2014). Where is the land of opportunity? The geography of intergenerational mobility in the United States. *The Quarterly Journal of Economics, 129*(4): 1553–1623.

Graphs for researchers. Researchers often demand much more detailed information than other readers. Like other audiences, researchers need to understand the main message of a graph but often will also benefit from additional details, such as standard error estimates. A range of visualization choices are available to highlight key findings *and* present additional details. Figure 4 is an example of an information-packed graph that shows the emergence of networks within a classroom.[47] Figure 5 illustrates how researchers can succinctly include critical descriptions of graphed data within a single figure caption so that the figure stands alone as a useful communication tool.

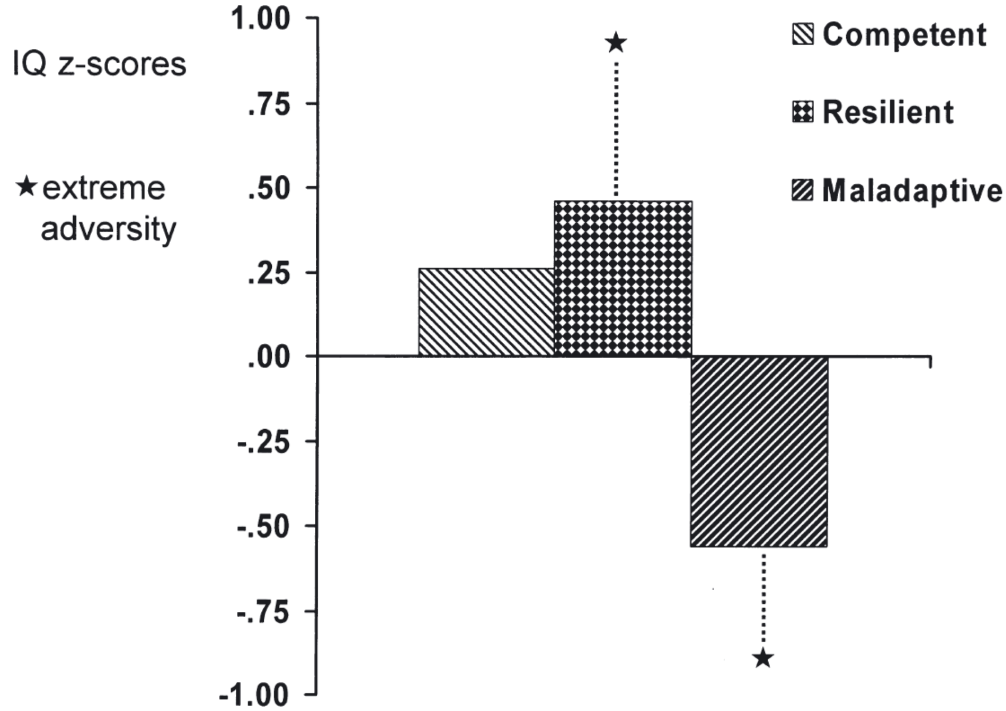
Figure 4. An information-packed graph showing the emergence of networks within a classroom (with time aggregation from 1 minute to 35 minutes).



Source: Bender-deMoll, S., & McFarland, D. A. (2006). The art and science of dynamic network visualization. *Journal of Social Structure, 7*(2).

[47] For more information on graphing networks, see Bender-deMoll, S., & McFarland, D. A. 2006. The art and science of dynamic network visualization. *Journal of Social Structure, 7*(2). Retrieved from http://www.cmu.edu/joss/content/articles/volume7/deMollMcFarland/.

Figure 5. A detailed title is used to convey information to ensure that the graphic stands alone and complete



Note: A comparison of three groups of youth "diagnosed" around the age of 20 years as competent (good adaptation and low adversity history), resilient (good adaptation and high adversity history), and maladaptive (poor adaptation and high adversity history). Mean IQ scores from childhood (measured 10 years earlier) are shown for the three groups. Means for the subgroups of resilient and maladaptive youth with extremely high lifetime adversity are shown by stars.

Source: Masten, A. S., & Obradović, J. (2006). Competence and resilience in development. *Annals of the New York Academy of Sciences, 1094*: 13-27. doi: 10.1196/annals.1376.003.

Communicating Descriptive Analysis: Summary

One of a researcher's jobs is to translate raw data into reported findings that are useful for the intended audience. It is often most effective to begin by deciding what the story (the message) is and to whom (the audience) it will be told. Only then can the researcher make sound decisions about the most appropriate way to communicate the story (customization).

Unnecessarily complex presentation is an obstacle to effective communication. For the vast majority of audiences, graphs and tables are much more easily understood than regression coefficients, three-dimensional imagery, or other complex presentations. It is critical that graphics present just as much information as necessary to fully meet the needs of the audience. Keep figures simple and neat, label axes, and avoid formats that are likely to confuse readers. The benefits of an appealing image are not worth the risk of misrepresenting data or message (see Box 13).

Box 13. Summary of Data Visualization Tips

Chapter 4 has recommended some rules of thumb for data visualization, but very few hard and fast recommendations are broadly applicable other than the following tips:

- Unnecessarily complex presentation is an obstacle to understanding and is detrimental to research.
- The benefits of an appealing image are not worth misrepresenting data or an important message.
- With rare exception, whatever *can* be presented graphically *should* be presented graphically.

- Good figures are designed to contain all of the information necessary to be understood. Although it is not always practical for a figure to stand alone, if extensive text is needed to convey meaning, reconsider presentation to see if an alternative approach to visualizing the data is possible.
- The same data can be presented differently for different purposes or for different audiences, depending on what you wish to convey and to whom.
- Three-dimensional (3D) graphs should not be used to represent two-dimensional (2D) data. While there may be rare exceptions, *most* readers will view your findings on a 2D computer monitor or a piece of paper, and 3D renderings can negatively affect perception and understanding.

Chapter 5. Summary and Conclusions

Descriptive analysis characterizes the world or a phenomenon; it identifies patterns in data to answer questions about who, what, where, when, and to what extent. Good descriptive analysis presents what we know about capacities, needs, methods, practices, policies, populations, and settings in a manner that is relevant to a *specific* research or policy question.

This document presents suggestions for more effectively considering, conducting, and communicating descriptive analysis, which is a critical component of the scientific process.

Descriptive analysis can stand on its own as a research product, such as when it identifies patterns in data that have not previously been recognized. Sometimes, description is a precursor to causal analysis (pointing toward causal understanding) or a follow-up to causal analysis (explaining why effects were or were not observed and suggesting more likely and less likely mechanisms). Whether the goal is to describe trends and variation in populations, create new measures of key phenomena, or describe samples for causal effects, descriptive analyses are part of almost every empirical paper and report.

When approaching descriptive work, researchers should endeavor first to recognize a phenomenon or question of interest (something that is occurring in the real world). Once a phenomenon of interest has been identified, a researcher should consider the phenomenon in question comprehensively and determine which features are likely to be salient before defining the constructs that represent these features. Choices about data collection and methods flow from these decisions.

During the data analysis phase, researchers use analytical and statistical methods to uncover observable patterns in the data. Sometimes, researchers review data with an expectation that a certain pattern exists. At other times, patterns emerge from a more general exploration of the data. In order to be meaningful, the patterns must be important, not simply present. Data dumps, all-purpose data dashboards, and generic tables of summary statistics do not qualify as sound descriptive analyses.

Descriptive analysis is an iterative process, with each step building upon others and requiring reconsideration and modification as the researcher's understanding of the phenomenon and the study unfolds.

Good descriptive research relies primarily on low-inference, low-assumption methods that use no or minimal statistical adjustments. Measures of central tendency (such as mean, median, and mode), measures of variation (such as range and standard deviation), and basic frequency analyses are useful statistical tools for description. Graphical methods and more complicated statistical adjustments have their place in descriptive analyses as well, but they should be used with caution because adjustments can mask relationships in the data.

At its core, descriptive analysis is data simplification. The researcher's job is to reduce the body of data to a format that helps an audience (such as practitioners, policymakers, or other researchers) to better understand a phenomenon. Often, this process of simplifying and customizing findings results in presentations that differ by audience type. A report published in a peer-reviewed academic journal may look very different from the same research story written for a policymaker or practitioner audience. Researchers may need access to data for the purposes of reproducibility, while practitioners may need a more comprehensive description of the study population to assess generalizability. In either case, the complementary final products contain only a subset of the analyses undertaken by the researcher. The underlying findings do not change, but because of varying information needs, the "final" presentation is modified for specific intended audiences. The translation, from raw to reported findings, is undertaken specifically to meet the information needs of the readers.

Some of the overarching principles that drive the process of customizing a research message for a particular audience can be distilled into the following questions, which, when answered by a researcher, frame the presentation of descriptive findings:

- What is your story (the message)?
- Who needs to hear this story (the audience)?
- For each target audience, what are the audience's specific information needs and how can you effectively communicate your story (customization)?
- How can you further improve your efforts to communicate your message to each specific target audience (iteration)?

Description plays an important role in identifying both needs and solutions in day-to-day and long-term efforts to improve teaching, learning, and the administration and management of our education system. Because understanding "what is" is a first step to effective practice and policy, educators, policymakers, and researchers should all be able to recognize good descriptive analysis. Although conducting descriptive analysis requires substantial expertise, its core principles reflect sound research and communications practices relating to clarity, comprehensiveness, accuracy, comprehension, and relevance (see Box 14).

Box 14. How to Recognize Good Descriptive Analysis

- Good description is clear about what it is trying to describe, its justification of methods and measures, and how data were transformed into a description of the phenomenon of interest.
- Good description provides detail and breadth that fully capture a phenomenon without being unnecessarily complex with respect to concept, data, methods, or presentation. It is neither too narrow nor too broad in focus; it is not unsuitably coarse or unnecessarily fine-grained.
- Good description is accurate. It reflects key concepts, incorporates a variety of perspectives and approaches, does not distort data or lend itself to misinterpretation, and will be accepted by broad communities of researchers, practitioners, and policymakers because it reflects real-world observation.
- Good description is both sensible and comprehensible. It uses appropriate concepts and methods, relies on relevant measures, and is presented in a manner that relates the salient features of a phenomenon in a way that can be readily interpreted by its intended audience.
- Good description focuses on socially important phenomena (and research questions).

Over the past 15 years, a focus on randomized control trials and the use of quasi-experimental methods has improved the body of causal research in education. However, this emphasis on causal analysis has not been accompanied by an improvement in descriptive analysis. In contemporary scholarship, descriptive analysis is too frequently viewed simply as a responsibility of the publication protocol. This misinterpretation of the scientific method should be corrected if the education community is to benefit from the unique value of descriptive research, which is a powerful tool for fueling discovery and advancing knowledge.

Appendix A. Resources Related Especially to Communications and Visualization

This appendix presents examples of additional resources that have been published about communications and data visualization.

Cleveland, W. (1993). *Visualizing data*. Lafayette, IN: Hobart Press.

Evergreen, S. D. H. (2013). *Presenting data effectively: Communicating your findings for maximum impact*. Los Angeles, CA: SAGE Publications.

Few, S. (2009). Now you see it: Simple visualization techniques for quantitative analysis. Oakland, CA: Analytics Press.

Few, S. (2012). *Show me the numbers: Designing tables and graphs to enlighten*, 2nd ed. Oakland, CA: Analytics Press.

Few, S. (2014). Data visualization for human perception. In M. Soegaard & R. F. Dam (Eds.). *The encyclopedia of human-computer interaction*, 2nd ed. Aarhus, Denmark: The Interaction Design Foundation. Retrieved March 2015 from https://www.interaction-design.org/encyclopedia/data_visualization_for_human_perception.html.

Tufte, E. R. (2001). *The visual display of quantitative information,* 2nd ed. Cheshire, CT: Graphics Press.

Wong, D. M. (2010). The Wall Street Journal guide to information graphics: The dos and don'ts of presenting data, facts, and figures. New York: W. W. Norton & Company, Inc.

Appendix B. References

Aldenderfer, M. S., & Blashfield, R. K. (1984). *Cluster analysis* (Quantitative Applications in the Social Sciences). Los Angeles: Sage Publications.

American Academy of Ophthalmology. (2011). *Genetics home reference: Color vision deficiency.* Bethesda, MD: U.S. Department of Health and Human Services, National Institutes of Health, National Library of Medicine. Retrieved March 2015 from http://ghr.nlm.nih.gov/condition/color-vision-deficiency.

Arnold, K., Fleming, S., DeAnda, M., Castleman, B. L., & Wartman, K. L. (2009). The summer flood: The invisible gap among low-income students. *Thought and Action,* Fall: 23–34.

Bender-deMoll, S., & McFarland, D. A. (2006). The art and science of dynamic network visualization. *Journal of Social Structure, 7*(2).

Bianco, S. D. (2010, June). Improving student outcomes: Data-driven instruction and fidelity of implementation in a Response to Intervention (RTI) model. *TEACHING Exceptional Children Plus,* 6(5).

Bloom, H. S., Hill, C. J., & Riccio, J. A. (2003). Linking program implementation and effectiveness: Lessons from a pooled sample of welfare-to-work experiments. *Journal of Policy Analysis and Management, 22*(4): 551–575.

Booth, W. C., Colomb, G. G., & Williams, J. M. (2008). *The craft of research,* 3rd ed. (NFES 2013–801). Chicago, IL: The University of Chicago Press.

Castleman, B. L., Arnold, K. D., & Wartman, K. L. (2012). Stemming the tide of summer melt: An experimental study of the effects of post-high-school summer intervention on college enrollment. *The Journal of Research on Educational Effectiveness, 5*(1): 1–18.

Castleman, B. L., Page, L. C., & Snowdon, A. L. (2012). *Summer melt handbook: A guide to investigating and responding to summer melt.* Cambridge, MA: Harvard University Center for Education Policy Research. Retrieved March 25, 2015, from http://cepr.harvard.edu/cepr-resources/files/news-events/sdp-summer-melt-handbook.pdf.

Chetty, R., Hendren, N., Kline, P., & Saez, E. (2014). Where is the land of opportunity? The geography of intergenerational mobility in the United States. *The Quarterly Journal of Economics, 129*(4): 1553–1623.

Cleveland, W. (1993). *Visualizing data*. Lafayette, IN: Hobart Press.

Cleveland, W. (1994). *The elements of graphing data*, 2nd ed. Summit, NJ: Hobart Press.

Crocker, L., & Algina, J. (2006). *Introduction to classical and modern test theory.* Belmont, CA: Wadsworth Publishing Co.

Daly, A. J. and Finnigan, K. (2011). The ebb and flow of social network ties between district leaders under high stakes accountability*. American Education Research Journal, 48*(1): 39–79.

Duncan, G., Morris, P., & Rodrigues, C. (2011). Does money really matter? Estimating impacts of family income on young children's achievement with data from random-assignment experiments. *Developmental Psychology, 47*(5): 1263–1279.

Dynarski, M., & Kisker, E. (2014). *Going public: Writing about research in everyday language* (REL 2014–051). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Analytic Technical Assistance and Development. Retrieved July 21, 2015, from http://ies.ed.gov/ncee/pubs/REL2014051/pdf/REL_2014051.pdf.

Engel, M., Claessens, A., & Finch, M. A. (2013). Teaching students what they already know? The (mis)alignment between instructional content in mathematics and student knowledge in kindergarten. *Educational Evaluation and Policy Analysis, 35*(2): 157–178.

Evergreen, S. D. H. (2013). Presenting data effectively: Communicating your findings for maximum impact. Los Angeles, CA: SAGE Publications.

Everitt, B. S., Landau, S., Leese, M., & Stahl, D. (2011). *Cluster analysis,* 5th ed. New York: Wiley.

Feldt, L. S., & Brennan, R. L. (1989). Reliability. In R. L. Linn (Ed.), *Educational measurement,* 3rd ed. (pp. 105–146). New York: Macmillan.

Few, S. (2009). Now you see it: Simple visualization techniques for quantitative analysis. Oakland, CA: Analytics Press.

Few, S. (2012). *Show me the numbers: designing tables and graphs to enlighten*, 2nd ed. Oakland, CA: Analytics Press.

Few, S. (2014). Data visualization for human perception. In M. Soegaard & R. F. Dam (Eds.), *The encyclopedia of human-computer interaction*, 2nd ed. Aarhus, Denmark: The Interaction Design Foundation. Retrieved March 2015 from https://www.interaction-design.org/encyclopedia/data_visualization_for_human_perception.html.

Gennetian, L., Castells, N., & Morris, P. (2010). Meeting the basic needs of children: Does income matter? *Children and Youth Services Review, 32*(9): 1138–1148.

Glazerman, S., Isenberg, E., Dolfin, S., Bleeker, M., Johnson, A., Grider, M., & Jacobus, M. (2010). *Impacts of comprehensive teacher induction: Final results from a randomized controlled study* (NCEE 2010–028). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.

Hoxby, C., & Avery, C. (2013). The missing "one-offs": The hidden supply of high-achieving, low-income students. *Brookings Papers on Economic Activity, 46*(1): 1–65.

Ioannidis, J. P. A. (2005). Why most published research findings are false. *PLoS Medicine, 2*(8): e124.

Jacob, R., Goddard, R., Kim, M., Miller, R., & Goddard, Y. (2015). Exploring the causal impact of the McREL Balanced Leadership Program on leadership, principal efficacy, instructional climate, educator turnover, and student achievement. *Educational Evaluation and Policy Analysis*, *37*(3): 314–332.

Kadushin, C. (2012). *Understanding social networks*. Oxford, UK: Oxford University Press.

King, R. S. (2014). *Cluster analysis and data mining: An introduction*. Dulles, VA: Mercury Learning & Information.

Knoke, D., & Yang, S. (2008). *Social network analysis,* 2nd ed. Los Angeles: Sage Publications.

Lankford, H., Loeb, S., McEachin, A., Miller, L. C., & Wyckoff, J. (2014). Who enters teaching? Encouraging evidence that the status of teaching is improving. *Educational Researcher, 43*(9): 444–453.

Lankford, H., Loeb, S., & Wyckoff, J. (2002). Teacher sorting and the plight of urban schools: A descriptive analysis. *Educational Evaluation and Policy Analysis, 24*(1): 37–62.

Lawton, B., Brandon, P. R., Cicchinelli, L., & Kekahio, W. (2014). *Logic models: A tool for designing and monitoring program evaluations* (REL 2014–007). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Pacific. Retrieved from https://ies.ed.gov/ncee/edlabs/projects/project.asp?ProjectID=404.

Magnuson, K., & Votruba-Drzal, E. (2009). Enduring influences of childhood poverty. *Focus, 26*(2): 32–37.

Masten, A. S., & Obradović, J. (2006). Competence and resilience in development. *Annals of the New York Academy of Sciences, 1094*: 13-27. doi: 10.1196/annals.1376.003.

Master, B., Sun, M., & Loeb, S. (in press). Teacher workforce developments: Recent changes in academic competiveness and job satisfaction of new teachers. *Education Finance and Policy*.

McFarland, D. A. (2006). Curricular flows: Trajectories, turning points, and assignment criteria in high school math careers. *Sociology of Education, 79*(3): 177–205.

Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational measurement*, 3rd ed. (pp. 13–103). New York: Macmillan.

Mitchell, M. N. (2012). *A visual guide to Stata graphics*, 3rd ed. College Station, TX: Stata Press.

Morris, P., Raver, C., Millenky, M., Jones, S., & Lloyd, C. (2010). *Making preschool more productive: How classroom management training can help teachers*. New York: MDRC. Retrieved from http://files.eric.ed.gov/fulltext/ED514648.pdf.

Murnane, R. J. (2013). U.S. high school graduation rates: Patterns and explanations. *Journal of Economic Literature, 51*(2): 370–422.

National Center for Education Evaluation and Regional Assistance. (2003). *Identifying and implementing educational practices supported by rigorous evidence: A user friendly guide* (NCEE EB2003). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance. Retrieved from http://ies.ed.gov/ncee/pdf/evidence_based.pdf.

National Forum on Education Statistics. (2005). *Forum guide to metadata: The meaning behind education data* (NFES 2009–805). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. Retrieved from http://nces.ed.gov/forum/pub_2009805.asp.

National Forum on Education Statistics. (2012). *Forum guide to taking action with education data* (NFES 2013–801). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. Retrieved from http://nces.ed.gov/pubs2013/2013801.pdf.

O'Donnell, A. M., & King, A. (1999). *Cognitive perspectives on peer learning*. New York: Routledge.

Prell, C. (2011). *Social network analysis: History, theory and methodology*, 3rd ed. Los Angeles: Sage Publications.

Quint, J., Zhu, P., Balu, R., Rappaport, S., & DeLaurentis, M. (2015). Scaling up the Success for All model of school reform: Final report from the Investing in Innovation (i3) evaluation. New York: MDRC.

Rawlings, C. M., & McFarland, D. A. (2010). Influence flows in the academy: Using affiliation networks to assess peer effects among researchers. *Social Science Research, 40*(3): 1001–1017.

Reardon, S. F. (2011). The widening academic achievement gap between the rich and the poor: New evidence and possible explanations. In R. Murnane & G. Duncan (Eds.), *Whither opportunity? Rising inequality and the uncertain life chances of low-income children*. New York: Russell Sage Foundation Press.

Romero, C., & Ventura, S. (2013). Data mining in education. *WIREs Data Mining and Knowledge Discovery, 3*: 12–27.

Rosenbaum, D. P. (1986). *Community crime prevention: Does it work?* Beverly Hills, CA: Sage Publications.

Schochet, P., Puma, M., & Deke, J. (2014). *Understanding variation in treatment effects in education impact evaluations: An overview of quantitative methods.* Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Analytic Technical Assistance and Development. Retrieved from https://ies.ed.gov/ncee/pubs/20144017/pdf/20144017.pdf.

Scott, J. (2012). *Social network analysis*. Los Angeles: Sage Publications.

Scott-Clayton, J. (2012). What explains trends in labor supply among U.S. undergraduates? *National Tax Journal, 65*(1): 181–210

Shager, H. M., Schindler, H. S., Magnuson, K. A., Duncan, G. J., Yoshikawa, H., & Hart, C. M. (2013). Can research design explain variation in Head Start research results? A meta-analysis of cognitive and achievement outcomes. *Educational Evaluation and Policy Analysis, 35*(1): 76–95.

Shakman, K., & Rodriguez, S. M. (2015). *Logic models for program design, implementation, and evaluation: Workshop toolkit* (REL 2015–057). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Northeast & Islands. Retrieved from https://ies.ed.gov/ncee/edlabs/projects/project.asp?ProjectID=401.

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science, 22*(11), 1359–1366. Retrieved from http://journals.sagepub.com/doi/10.1177/0956797611417632.

Snyder, T. D., & Dillow, S. A. (2013). *Digest of education statistics 2012* (NCES 2014–015). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics.

Stark, P., & Noel, A. M. (2015). *Trends in high school dropout and completion rates in the United States: 1972–2012* (NCES 2015–015). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. Retrieved July 2015 from http://nces.ed.gov/pubs2015/2015015.pdf.

Tufte, E. R. (2001). *The visual display of quantitative information,* 2nd ed. Cheshire, CT: Graphics Press.

Vanneman, A., Hamilton, L., Anderson, J. B., & Rahman, T. (2009). Achievement gaps: How black and white students in public schools perform in mathematics and reading on the National Assessment of Educational Progress (NCES 2009–455). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics.

Walter, C. (2014). *Inputs in the production of early childhood human capital: Evidence from Head Start* (NBER Working Paper No. 2014.01). Cambridge, MA: National Bureau of Economic Research.

Wandersman, A. (2009). Four keys to success (theory, implementation, evaluation, and resource/system support): High hopes and challenges in participation. *American Journal of Community Psychology, 43*(1–2), 3–21.

Wasserman, S., & Faust, K. (1994). *Social network analysis: Methods and applications*. Cambridge, UK: Cambridge University Press.

Wong, D. M. (2010). The Wall Street Journal guide to information graphics: The dos and don'ts of presenting data, facts, and figures. New York: W. W. Norton & Company, Inc.