



RACHEL DUNIFON

How to Read a Research Article

The goal of this Research Brief is to provide information that will make reading a research article more illuminating. For example, you might want to learn about whether any research has been done showing that children benefit when they are read to more frequently. To answer this question, start with a literature search on the web (see "Resources for Doing Web Research"¹) and then identify the relevant articles on this topic.

One such article is entitled "Quality of Adult Book Reading Affects Children's Emergent Literacy," by Elaine Reese and Adell Cox.² This article discusses how different styles of reading to children (which they call the describer style, the comprehender style and the performance-oriented style) impact children's literacy, and will be used as an example throughout this Research Brief. It is hoped that, by the end of this brief, you will be able to understand and critique this article, and others like it.

What is a "research article"? A research article is a paper written by authors who either collected and analyzed their own data (primary data analysis), or analyzed data that had been collected by someone else (secondary data analysis). A research article consists of new, original work, which no one else has done before. It is not simply a summary of research that others have done—instead, it consists of the presentation of new analyses. In a research article, the authors provide background information on similar work that has been done in the area, illustrate why the current paper is important, describe the data and methods they use, present their results, and discuss the results and limitations of the study. A research article should contain the following sections— Literature Review, Data Description, Methods, Results, and Discussion.

Why should I read a research article? Educators today are inundated by information— some of it good, and some of lesser quality. It is important to discern what

information is evidence-based, and what is not. Oftentimes, reports in the popular media suggest that something is "proven," when actually, the research is not so clear (for example, the idea that listening to Mozart in utero is associated with higher IQ scores among children). Research articles provide the best source of rigorously-tested, evidence-based information that can inform your work and help you assess if what you have read in the media really is true. Sometimes educators rely on synopses of research (such as a Research Brief), rather than reading the original research articles. This can be useful when attempting to gain a general sense of knowledge about a topic. However, if you want to gain an in-depth understanding of a topic, it is important to read the original research articles themselves. Frequently synopses of research are filtered through someone else's perspective about which articles are important and how they should be described. By finding articles and reading them yourselves, you are empowered to determine what is relevant to your work and what the research means for you.

The article by Reese and Cox, for example, could tell you what type of reading style is most beneficial for children, which could in turn inform the types of literacy programs that Cornell Cooperative Extension offers.

Step 1: The Source—what am I reading? The first thing to examine when reading a research article is where the article was found. Is it a peer reviewed journal? Was it published on the web? Is it from a government source? (See "Resources for Doing Web Research" for help in evaluating articles found on the web). Typically, peer-reviewed journal articles are considered to be the highest quality, because they have undergone a rigorous review process prior to publication. Most academic journals are peer-reviewed, while many research briefs and reports found on the web are not. The Reese and Cox article² was published in *Developmental Psychology*, a signature journal published by the American Psychological Association.

¹ Available at

<http://www.parenting.cit.cornell.edu/How%20to%20do%20Web3.pdf>

² Available at

<http://www.parenting.cit.cornell.edu/Reese%20and%20Cox%20article.pdf>

Step 2: The Literature Review —what has been done before? All research articles begin with a review of the other research that has taken place on this topic. A good literature review should:

- describe work done by other scholars, not just the author of the paper; and
- mainly discuss articles from peer-reviewed journals.

The overall goal of the literature review is to provide the reader with an integrative summary of other research findings and the questions that remain unanswered or require additional research.

Reese and Cox begin their paper with a review of previous research on reading to children and identify research that is needed to determine how different reading styles impact children.

Step 3: The Research Question—what are they doing? Authors will often follow the literature review by setting forth their research question. It should not simply replicate what someone else has done before, but instead offer something new. Frequently authors will state the research question as a hypothesis by offering a prediction of what they think they will find, and will test that hypothesis to see if it holds true (example: "children who are read to more frequently will have higher test scores"). However, many times an author will simply state the general question they seek to answer, without offering a hypothesis (example: "is reading to children associated with improved test scores?"). The research question should be both clearly stated and answered by the end of the article.

Reese and Cox describe their research question on page 21 ("our primary goal was to assess experimentally the relative benefits of these naturally occurring reading styles"), and their hypotheses on page 22 ("we predicted that children of higher initial skill levels would advance more with higher demand styles, and less skilled children would advance more with lower-demand styles").

Step 4: Data—who is being studied? Research articles should provide a good description of the data used in the analysis. Some important aspects of the data that require consideration are:

- **Sample size**—researchers often face a trade-off between obtaining less detail on a large number of people, vs. obtaining lots of detail on a small number of people. There are no absolute standards for the ideal number of subjects in a study, but keep in mind that this trade-off exists. Additionally, the smaller the sample size, the more difficult it will be to conduct

statistical analyses and the less reliable and generalizable these analyses will be.

- **Representativeness**—the description of the data should help clarify whom the authors hope to represent or describe in their sample. For example, a sample could represent college students at a University, mothers applying for WIC in a small town in upstate New York, or the entire United States.

Three concepts are central to understanding representativeness:

1. The *population* is the group of people whom the data intended to describe. Some studies are representative of all households in the United States. Therefore, their *population* is all households in the United States. Other studies may be representative of all children attending a specific elementary school.
2. Rather than collecting data on everyone in the targeted population, researchers often use *a sample* that is intended to represent the entire population. So, most studies whose population is all households in the U.S. do not actually collect data from each household in the U.S.; instead, they might use a sample of a smaller number of households that is representative (in terms of age, race, family structure, etc.) of all households in the U.S. The authors should be clear about how their sample was obtained and offer information allowing the reader to know to what extent it is representative of the population that is being studied.
3. Finally, it is important to know the *response rate* of the sample. If someone did a survey of children in an elementary school, but only half of them completed the survey, the response rate is 50%. This leads to concerns that the children who completed the survey are different from those who did not (for example, more motivated, outgoing, or have more unoccupied time at school), and as a result, the sample would not be an accurate representation of all children in that school. There is no "right" number for a response rate, but knowing what the rate is can help you judge how representative the data are.
4. Understanding the representativeness of the data is important in order to determine *the generalizability* of the results. Results from a study that is representative of the entire U.S. can be generalized to the entire U.S. If the study found that reading to children daily

improves children's test scores, then we can be confident that this finding applies to the *average* American child. Results from a study that is representative of a more narrow population cannot be generalized beyond that population. For example, results from a study of mothers applying for WIC in Ithaca cannot be generalized to all mothers in the state of New York.

- **Descriptive statistics**—the authors should provide descriptive statistics (means and standard deviations, defined below) to describe their sample. This should include the average ages, race, educational level, income, and other characteristics of the people in their data.

Reese and Cox have a sample of 50 four-year olds recruited from preschools in one town in New Zealand. They do not appear to be a random sample of children from these schools; rather, they are children whose parents agreed to let them participate. Therefore, they are not representative of all children at these schools. The authors also do not provide information on the response rate, so we do not know how many children refused to participate. The authors provide some descriptive statistics in the text of the paper, on page 22, telling us the ages of the children, their socioeconomic status (SES), race, and test scores.

Step 5: Measures—the authors should tell you how each variable used in their analyses was measured and defined. For example, if they say that they are measuring children's cognitive ability, how is this defined? Are they using a widely-used scale? Are they using several items or a single item? Did they create their own measure, and if so, how did they do it?

When deciding how to measure things, such as "cognitive ability," researchers can either use assessments that have already been developed and tested by others, or they can create their own measures. If they use an assessment that has already been developed and tested, they simply cite the fact that it is widely-used and has been shown to measure what it really claims to be measuring (this is called *validity*). If researchers create their own measure, they need to prove that their instrument actually measures what they say it does. Frequently they do this by comparing their measure to others that are already proven to be valid to show that the two measures are capturing similar phenomena.

Reese and Cox describe their measures under the Procedures section on pages 22-23. They used well-known tests for some measures (such as using the PPVT-R

for vocabulary), but created their own for others (such as story comprehension).

Step 6: Methods and Results—what did they do and what did they find? Here are some definitions of terms that you might come across in the Methods and Results sections of an article.

- **Mean**—this is the average value of all of the responses to one question. For example, if the researchers asked 92 parents to report how many days per week they read to their children, the average (also known as the mean) reading frequency would be the sum of all 92 responses, divided by 92.
- **Standard deviation**—this is how the variables are distributed around the mean. A larger number indicates that the individual responses (i.e., individual reports of reading frequency) deviate significantly from the mean some parents read much more often than others. By contrast, a low number indicates that most individual reports are close to the mean.
- **P-value**—this is a numerical representation of how "trustworthy" the results are. The researchers apply a formula to determine whether their results are due to chance as opposed to an actual correlation. For example, if a study found that children who were read to daily score 10 points higher on a cognitive test than children who were not, the p-value would tell us how likely it is that this result is due to chance as opposed to a real correlation between reading and test scores. Researchers usually agree that a p-value of .05 or less is statistically significant. So, for example, if someone says that the difference in test scores between two groups of children is significant at the $p < .05$ level, this means that we are 95% sure that there is a real difference between the two groups and that it is not due to chance.
- **Regression**—a regression is a type of analysis that uses one or more independent variables to predict one dependent variable. So, for example, children's test scores (the variable under observation, i.e., the *dependent variable*) could be influenced by many factors (such as parental reading frequency, household income, and parental education—the *independent variables*). Using all of the data available, a regression model finds the estimates that best represent the data. So, in our sample, the regression results would tell us whether children whose parents read to them more have higher test

scores than other children whose parents have the same level of education and household income but read to them less. A coefficient (i.e., estimated effect) of 10 on the reading variable would mean that each additional time per week that a parent reads to a child, that child's test score is expected

to rise by 10 points. Using a p-value (see above) we can determine whether that coefficient is significant.

There are several types of methods that researchers may use. Some of the different types are illustrated below:

Method	Definition	Pros	Cons
Experiment	Subjects are randomly assigned to experimental groups (who receive a treatment or program) and control groups (who do not). If the two groups are identical in every way, then the impact of the treatment can be determined by comparing outcomes between the two groups.	This is the only way to truly determine whether a treatment causes an outcome (i.e., did reading to children more lead to increased test scores?).	It is difficult to obtain truly identical treatment and control groups.
Pre and post test	The same subjects are surveyed before and after intervention.	Easy to conduct.	Even if the group demonstrates a change over time in outcomes, without a comparison group, we cannot know whether this change is due to the intervention, or to some other factor.
Survey research	The collection of information from a sample through their responses to questions. Could be done over the phone, in person, through the mail, or on the internet.	Can cover a wide range of topics, is relatively easy and quick, can draw a large sample.	A good survey relies on asking good questions, getting a representative sample, and getting a good response rate—all of which can be hard to achieve (see above).
Qualitative research	Methods that rely on observing people's natural behavior or gathering their responses to open-ended questions. The major qualitative methods are participant observation, intensive interviewing, and focus groups.	Can shed light on phenomena that have not been well-understood and in more detail than can be captured in survey data. Gives a richer picture of what is going on in people's lives.	Can only state results for the sample involved, not a larger population. Very time-intensive. Can be difficult to test causal hypotheses. Hard to compare the results of one study to that of another.

Reese and Cox used an experimental design, as indicated on page 23 ("children were matched on their PPVT-R scores and gender and then randomly assigned to one of the three reading styles"). Rather than having experimental and control groups, this study had three different experimental groups, representing three different types of reading styles. Because children in each of the three groups were identical, any differences at the end of the study can be attributed to differences in the reading styles to which they were exposed. Table 2 provides a good description of how the three different reading styles were defined.

Reese and Cox present their main results in Tables 4 and 5 and discuss them on pages 25-26. (Don't worry if you can't make complete sense of the tables... the goal is to gain a basic understanding of what they did, and then, by reading the text and tables, what their main findings were). One of their findings is that "children with higher initial vocabulary skills gained the most from the performance-oriented style, whereas children with lower initial vocabulary skills gained the most from a describer style of reading" (p. 25).

Step 7: Conclusions—what does it all mean? At the end of the paper, the authors should summarize what they found, and tie their results in with the other literature in the area. They should discuss instances in which their findings differ from others', suggest possible reasons why, and offer interpretations of their findings. For example, if the study found that reading to children daily is associated with improved test scores, what can we take from this in terms of policy and practice? What still remains to be learned? Importantly, the Conclusion section should also discuss the limitations of the study. All studies have limitations. Unless a study contains a randomized experiment, it is difficult to determine cause and effect (i.e., did the frequency of reading *really* cause children's test score to rise, or was it something else that the researchers were not able to observe—such as perhaps

children whose parents read to them more also talk to them more throughout the day). Therefore, the authors should discuss what they cannot know from their data and should not over-state their conclusions to be stronger than the evidence permits.

Reese and Cox summarize what they found on page 26 ("the main finding was that a describer style of book reading with children does appear to provide overall benefits for their receptive vocabulary and print skills in comparison with the other two reading styles"). They go on to further describe their findings ("another main finding... was that . . . children with higher initial vocabulary benefited most... from a performance-oriented style... "). The authors then discuss the importance of these findings and what next steps are needed for research.

Summary: The purpose of this brief is to be a resource for navigating through scholarly research, and to emphasize the importance to educators of reading such research. Although it may be difficult at times to sift through longer, academic research articles, there are many benefits of staying up-to-date on research related to one's work. By reading original research, readers can draw their own conclusions as to the relevance of research findings to their daily work. Research can be used to inform the design of community programs and update existing resources materials. As a result, educators can integrate major findings from scholarly research into their programmatic work and have a greater impact on the community in which they work.

Sources:

Gonick, L., & Smith, W. (1993). *A Cartoon Guide to Statistics*. New York: Harper Perennial.

Reese, E. & Cox, A. (1999). "Quality of Adult Bookreading Affects Children's Emergent Literacy." *Developmental Psychology*, 35(1): 20-28.

Skutt, R. (2001). *Understanding the Social World*. Thousand Oaks, CA: Pine Forge Press

Rachel Dunifon is an Associate Professor in the Department of Policy Analysis and Management at Cornell University

