Spring 2015

Public understanding of chemistry research in print news

Michael D. Hands
Purdue University

Graduate School Form 30
(Updated 11/20/2014)

PURDUE UNIVERSITY
GRADUATE SCHOOL
Thesis/Dissertation Acceptance

This is to certify that the thesis/dissertation prepared

By Michael D. Hands

Entitled PUBLIC UNDERSTANDING OF CHEMISTRY RESEARCH IN PRINT NEWS

For the degree of Doctor of Philosophy

Is approved by the final examining committee:

George M. Bodner

Adam Wasserman

Lyudmila Slipchenko

Gabriela Weaver co-chair

To the best of my knowledge and as understood by the student in the Thesis/Dissertation Agreement, Publication Delay, and Certification/Disclaimer (Graduate School Form 32), this thesis/dissertation adheres to the provisions of Purdue University's "Policy on Integrity in Research" and the use of copyrighted material.

Approved by Major Professor(s): George M. Bodner

Gabriela Weaver

Approved by: R. E. Wild 04/09/2015

Head of the Department Graduate Program Date

PUBLIC UNDERSTANDING OF CHEMISTRY RESEARCH IN PRINT NEWS

A Dissertation

Submitted to the Faculty

of

Purdue University

by

Michael D. Hands Jr.

In Partial Fulfillment of the

Requirements for the Degree

of

Doctor of Philosophy

May 2015

Purdue University

West Lafayette, Indiana

For Faith

ACKNOWLEDGEMENTS

I would first like to thank Faith Weeks for all of her support and encouragement. She is the love of my life and I wouldn't have been able to complete this Ph.D. without her. Her advice, emotional support, and patience listening to me talking about my research were invaluable.

I would like to thank my advisor Dr. Gabriela Weaver for her willingness to allow me to pursue this research and for her guidance. I would also like to thank her for the opportunities that she provided me.

I would also like to thank all of my participants, as this research would not be possible without their involvement.

In addition, I would like to thank Dr. Lyudmila Slipchenko for her support, both professional and financial, as I worked on my Ph.D. while working with her on my M.S.

Dr. George Bodner was very supportive in guiding me when I had questions about the field of chemistry education and with professional development.

Dr. Jake Jensen was incredibly helpful in the planning of my research. As this research was a new area for my advisor and me, his expertise was vital in ensuring that my dissertation work was the best that it could be.

I am very grateful for the financial support and opportunities to learn about chemical information that I received from Jeremy Garritano and David Zwicky. Working with them has expanded my knowledge of chemical education issues into chemical information and they each have been great to work with.

I would like to thank all of the members of the Discipline Based Education Research Graduate Student group. Interacting with other graduate students in related fields has been incredibly valuable and I have learned so much from our interactions. I hope that the group continues on and thrives.

I would also like to thank Dr. Nicole Cook for her friendship, support, and input along the way to finishing my dissertation. As Weaver group members left and others joined, Nicole was always there to exchange ideas and provide advice along the way. She is also a great friend and I enjoyed the fun times she spent with Faith and I.

Finally, I would also like to thank the other Weaver group members, past and present. In particular, Dr. Matt Pilarz, Dr. Gabriela Szteinberg, Dr. Kermin Martinez-Hernandez, and Dr. Kellie Green were especially helpful in welcoming me into the group and mentoring me. I also need to thank Tony Chase for his assistance with inter-rater reliability and for being generally being a sounding board for my ideas.

TABLE OF CONTENTS

Page

LIST OF TABLES

Tables Page

LIST OF FIGURES

ABSTRACT

Hands, Michael D. Ph.D., Purdue University, May 2015. Public Understanding of Chemistry Research in Print News. Major Professor: Gabriela Weaver.

Despite numerous calls for improving scientific literacy, many American adults show a lack of understanding of experiments, scientific study, and scientific inquiry. News media is one important avenue for science learning, but previous research investigating health and/or environmental science news has shown that it is inconsistent in the presentation of scientific research limitations, potentially impacting reader understanding.

In the first phase of this dissertation, seventeen news articles reporting on a single chemistry research article, along with associated press releases and research articles, were analyzed using move analysis to determine the structure of each type of text. It was found that the overall structure of each text genre was similar, with the main difference being that research articles start by presenting background information, while the others lead with highlighting overall research outcomes. Analysis of the steps revealed that, as seen for health and environmental science news articles, descriptions of the study limitations and methods were generally omitted in the news articles.

Using these findings, a pilot study was conducted where study limitations were added to a chemistry research news article and the effect of its presence on staff members employed at a large Midwestern university (n=12) and science faculty employed at the same institution (n=6) was explored. Interviews with the participants revealed that including limitations enhanced readers' ability to identify conclusions and evaluate claims, but decreased their trust in the information.

In the final part of this study, the trends seen in the previous phase were explored to determine their generalizability. Members of the public (n=232) and science faculty (n=191) read a randomly assigned news article either presenting or omitting the study limitations and research methods. Participants reading articles presenting limitations were able to evaluate the reasonableness of claims based on the article better than those who read the article omitting limitations when accounting for their views on the tentativeness of science (ToS). Presenting limitations was important in identifying unreasonable claims for both public and science faculty, while ToS views predicted ability to identify reasonable claims for the public. Including limitations also decreased readers' trust in the conclusions of the research. However, it did not impact their ability to determine the conclusions of the research and including methods did not have any effect on the measured outcomes.

CHAPTER 1. INTRODUCTION

1.1 Introduction

Over the past few decades, there has been a rise in concern about scientific literacy, both in academic and general public circles. While the term “scientific literacy” is ill-defined, it can be used to refer not just to what one learns in a classroom, but how one can use their science knowledge in other settings. According to the *National Science Education Standards,*

> Scientific literacy entails being able to read with understanding articles about science in the popular press and to engage in social conversation about the validity of the conclusions. Scientific literacy implies that a person can identify scientific issues underlying national and local decisions and express positions that are scientifically and technologically informed. A literate citizen should be able to evaluate the quality of scientific information on the basis of its source and the methods used to generate it. Scientific literacy also implies the capacity to pose and evaluate arguments based on evidence and to apply conclusions from such arguments appropriately (National Research Council, 1996).

Publications from other national organizations concerned with science education, *Science for All Americans: Project 2061* (American Association for the Advancement of Science, 1989), *Benchmarks for Science Literacy* (American Association for the Advancement of

Science, 1993), and *Science and Engineering Indicators 2014* (National Science Board, 2014), are in agreement with the NRC of the importance of scientific literacy and that citizens must be able to use their science knowledge to make informed decisions. The criteria for being scientifically literate indicate that citizens must not only have science content knowledge, but fairly sophisticated scientific reasoning skills. However, even with this emphasis on scientific literacy, it is clear that most Americans are not scientifically literate (Miller, 1986, 2004; National Science Board, 2014).

Calls to improve science literacy have generally focused on formal education settings (Alberts, 2009; Committee on Science, Engineering, and Public Policy, 2007; L. M. Lederman & Malcom, 2009; National Academies of Science, 2011). Research has been done investigated methods of improving scientific literacy in classrooms, but comparably little has been done to address the issue in other learning environments (Falk & Needham, 2013). However, American adults only spend approximately 5% of their lifetime in classrooms, with an even smaller percentage of that time devoted to science learning (Falk & Dierking, 2010). Of the approximately 30% of US adults age 25 and over that have at least a bachelor's degree (U.S. Census Bureau, 2013), about one-third have degrees in science and engineering (National Science Board, 2014). Therefore, much less than half of the US population has formal STEM (Science, Technology, Engineering, and Mathematics) education beyond high school, either by majoring in a STEM discipline or from other undergraduate course requirements. Recently, there has been greater recognition of the potential contributions to science learning from informal settings (Bell, Lewenstein, Shouse, & Feder, 2009; Stocklmayer, Rennie, & Gilbert, 2010), suggesting the need for more research in this area.

Large surveys have found that American adults obtain most of their science information in informal learning settings, such as visiting museums, zoos, and aquariums or using various forms of media (Falk, Storksdieck, & Dierking, 2007; National Science Board, 2014). Fifty percent of all American adults stated that they had visited a zoo or aquarium during the previous year, while 26% claim to have visited a science/technology museum and 27% claim to have visited a natural history museum (National Science Board, 2014). For those adults who did not have a minor in their household, attendance dropped to 44% for zoos/aquariums, 24% for natural history museums, and 25% for science/technology museums (National Science Board, 2014). This small drop in attendance among adults without children suggests that they are using these informal learning opportunities at least partially for their own benefit. When asked what their primary source of science information was, the vast majority of American adults cited some form of media, with 26% indicating online or print newspapers and magazines (National Science Board, 2014), making it the second most popular source, behind television.

However, there is little research on how adults judge the credibility of science news or how their level of scientific literacy affects their understanding. The vast majority of internet users report checking science information they found online in at least one other source (Horrigan, 2006), but that does not provide information on how they are interpreting the science. It is likely that an individual's level of scientific literacy influences their understanding of science news because it provides a framework for the individual to interpret the information presented. There is also the potential that reading science news could impact a person's scientific literacy because it is a learning experience (Britt, Richter, & Rouet, 2014). If that is the case, then science news could be an important avenue for promoting scientific literacy.

In this dissertation, I will describe the overall purpose and rationale for the study, and then discuss relevant background literature. The overall study is divided into three related that will be presented separately:

1. The Structure of Chemistry Research Reports, Press Releases, and News Reports
2. Chemistry Research in the News: The Effect of Described Study Limitations on Public Understanding
3. Chemistry Research in the News: The Effect of Described Methods and Study Limitations on Public Understanding

Research questions, methods, results and conclusions for each phase will be described in the associated chapter.

1.1.1 Purpose and Rationale

The overall purpose of this study is to examine how aspects of science news reports affect two specific aspects of scientific literacy, namely reading with understanding about science and evaluating arguments based on evidence. In this study, reading with understanding will hereafter be referred to as content understanding. Promoting scientific literacy has generally been viewed as an issue for formal education research. However, the majority of American adults are finished with their formal education, leaving only informal educational avenues to promote scientific literacy. For this population, informal education experiences could be an important avenue for science learning. Research in formal education environments can benefit future generations, while research in informal education could benefit the current adults. Therefore, it would be beneficial to present science information, such as science news, in a way that promotes scientific literacy. Achieving this goal requires research into how different aspects of news reports impact readers.

1.2 Literature Background

1.2.1 General Public Scientific Literacy

Measures of scientific literacy have traditionally focused exclusively on science content knowledge (Brossard, 2006; Miller, 1986, 1998, 2004). Over the past 25 years, public responses to basic science knowledge questions have remained fairly consistent, ranging from a low of 39% correctly identifying *The universe began with a huge explosion* as a true statement to a high of 84% correctly identifying that *The center of the Earth is very hot* is a true statement in the latest survey (National Science Board, 2014). However, there is evidence that some science content knowledge questions, particularly about evolution and the big bang, are actually measuring religious belief rather than science knowledge (Roos, 2012). Science content knowledge is only one aspect of scientific literacy, with scientific reasoning skills being at least as important and perhaps a better measure.

More recently, literacy measures have asked questions to assess understanding of probability, experiments, scientific inquiry, and scientific study (National Science Board, 2014). Responses to these questions have also remained fairly constant over the past 20 years, with levels of 65%, 34%, 33%, and 20% respectively for understanding of probability, experiments, scientific inquiry, and scientific study in the most current survey (National Science Board, 2014). Having less than a third of the adult population show an understanding of the scientific process is far below what I believe is desirable in order to have a scientifically literate society.

1.2.2 Informal Science Education

While the vast majority of research on science learning has historically focused on formal education settings, an increased interest in how people learn science outside of the classroom has led researchers to study the public understanding of science. The field of informal science education is extremely diverse, spanning settings from museums, zoos, aquariums, and science centers to the home, community projects, workshops, and hobbyist organizations, with varying research goals in the cognitive, affective, interpersonal, or behavioral realms (Brody, Bangert, & Dillon, 2007). Studies tend to either evaluate experiences or seek to understand how people learn in informal settings (Falk & Needham, 2011; Rennie & Williams, 2000, 2006; Stutchbury, 1999). When evaluating informal education programs, one area of focus has been to assess the impact of institutions on their community. Telephone surveys of Los Angeles residents found that most reported visiting a new science center within a ten-year period and that they believed it had influenced their science understanding (Falk & Needham, 2011). While attendance demonstrates some impact, it is not possible to interpret the influence on science understanding because many other factors could have been involved and it was self-reported, rather than an objective measure. Surveys of “Science in the Pub” events in Sydney, Australia indicated that 72% of attendees felt that they learned something new, but long term effects and other objective measures were not examined (Stutchbury, 1999). A study of attendees of one public science lecture in a series about human genetics in Western Australia found that participants showed increased positive attitudes towards and interest in genetics (Rennie & Williams, 2000). Results from this study were consistent with results from attendees of museums and science centers (Rennie & Williams, 2006). However, these two studies focused on individuals who were self-selected to attend an informal science education opportunity.

In addition to research evaluating the impact of informal education events, other studies have examined aspects of how people learn in these settings. Observation studies of both family and student groups visiting museums and science centers have found that people spend a relatively short amount of time at each exhibit, but do appear to have learning agendas that guide their interactions (Dierking & Falk, 1994; Rennie & McClafferty, 1995). Observations have been correlated to learning (Falk, 1983), but directly measuring learning by surveying or interviewing participants may provide a clearer understanding of what was learned. It has also been found that an important factor for learning in informal environments is the background and goals of the individual (Falk & Adelman, 2003; Falk & Storksdieck, 2010; Rennie, 1995). This may be due to the fact that people participate in these settings by choice (Falk, 2001; Falk & Dierking, 2000), so personal factors are very important. Therefore, the effectiveness of the informal learning materials interact with individual attributes to generate true impact.

The previous studies reviewed were done in the context of informal learning environments where the individual went to a physical location whose main purpose is to teach about science. However, other learning opportunities can occur in media environments. From 2007 – 2010, science news comprised 5 – 12% of all news stories in traditional media (National Science Board, 2014). Yet, much less than 0.05% of science research papers were discussed in the news media (Suleski & Ibaraki, 2009). Still, science news has the potential to reach large numbers of people, making it an important informal science education avenue, despite the vast majority of scientific research not being widely disseminated.

On television, varying views of nature of science are depicted, from science being a collection of facts to questioning scientific discoveries (Dhingra, 1999). Work in this area that focused on young children has found they were able to learn science from television in the short-term and those watching educational programs on a regular basis seem to have a small advantage in school readiness (Dhingra, 2006). High school students viewed nature of science differently based on the type of science program watched (Dhingra, 2003), with documentaries and magazine-format shows promoting science as a collection of facts and news segments promoting science as uncertain. However, the news programs promoted a sense of uncertainty stemming from ethical considerations and the social consequences of science, rather than scientific uncertainty, so it is not clear if the students really had more sophisticated views of nature of science. Other attempts to assess learning science from television have found that less than half of adults recalled any science content from news stories they had watched that were presented over a six-week period (Miller, Augenbraun, Schulhof, & Kimmel, 2006). The results were attributed to typical viewing habits where the viewer is not necessarily focused on the news story. Given that limitation, it is not clear how reasonable it is to expect adults to learn science content from television news, but learning in other forms of science media or promoting other aspects of scientific literacy may be possible.

1.2.3 Presentation of Science in the News

As discussed previously, a substantial portion of the population use print media as their primary source of science information. Therefore, it is important to understand how science is presented in these contexts. It has been suggested that most science news is presented from a perspective that highlights the benefits of scientific research on human life, which can de-emphasize other aspects of science, particularly scientific uncertainty (Maier, Rothmund, Retzbach, Otto, & Besley, 2014). Brechman, Lee, and Cappella (2009) analyzed press releases and subsequent news reports about genetics. They found that both sources presented discoveries in a simplified, deterministic way and reported different content a substantial portion of the time (Brechman, Lee, & Cappella, 2009). Another study tracked one cancer genetics story from the primary research report to newspaper reports and found that the meaning of the findings changed, in addition to the way the findings were reported (Kua, Reder, & Grossel, 2004).

These results were confirmed with analysis of multiple cancer genetics stories, with experts judging the press release as more representative of the original research article than news reports (Brechman, Lee, & Cappella, 2011). These results are consistent with other research indicating that science information in the news is simplified, science content can be different, conclusions can change, limitations and caveats are removed, and the research is depicted as more certain than in the original research article (Nelkin, 1995; Pellechia, 1997; Stocking, 1999; Tankard & Ryan, 1974).

Press releases (Woloshin & Schwartz, 2002) and news reports of scientific conference presentations (Woloshin & Schwartz, 2006) also omit research study limitations. These studies suggest that any confusion about the findings of science research might be due to distortions made from the research article, but more research is needed to determine if this is truly the case.

1.2.4 Aspects of News Reports Affecting Reader Outcomes

Some research has been conducted to look at how manipulations of news articles affect readers, demonstrating that the language used in science news reports can affect readers' perception of the science (Budescu, Broomell, & Por, 2009; Corbett & Durfee, 2004; Jensen, 2008; Jensen et al., 2011; Yaros, 2006). The inclusion of a broader context and views of other scientists had a significant impact on readers' perception of the certainty of climate change (Corbett & Durfee, 2004). Members of the general public interpret phrases conveying probabilities differently than scientists, so altering the language used can bring reader interpretation either more in line or less in line with the scientists' intent (Budescu et al., 2009). It has also been found that readers find scientists and journalists more trustworthy when the study limitations were reported in cancer research news reports (Jensen, 2008). Including limitations also reduced reader cancer fatalism and nutritional backlash (Jensen et al., 2011), which are negative responses to cancer news associated with unhealthy habits. Along with language used, the organization of a science news report can affect reader interest and comprehension (Yaros, 2006). Yaros altered the structure of two *New York Times* articles, one about cancer research and another about nanotechnology research, and found that the modified structure changed reader interest and comprehension. These studies suggest that the addition of context and study limitations could have positive impacts upon readers.

While both the method of presentation of science in text news reports and how readers perceive that information have been investigated previously (Brechman et al., 2009, 2011; Budescu et al., 2009; Corbett & Durfee, 2004; Jensen, 2008; Jensen et al., 2011; Kua et al., 2004; Nelkin, 1995; Pellechia, 1997; Stocking, 1999; Tankard & Ryan, 1974; Woloshin & Schwartz, 2002, 2006; Yaros, 2006), most of the news reports examined have been in the context of health and/or politically controversial science. These areas are the most common type of science reported (National Science Board, 2014; Suleski & Ibaraki, 2009), but studies investigating this type of science news may not be measuring just a reader's scientific literacy. A reader may have a personal connection to news about health or refer to their political values when reading politically controversial science that informs their perception of science news. Therefore, it is important to investigate science news coverage of topics in other fields that may have less of a personal connection to the reader in order to isolate scientific literacy from other influences on readers.

1.3 Overall Research Questions

This study was guided by the following research questions:

1. How is chemistry research reported in print news?
2. How do aspects of news articles on chemistry research impact reader's understanding and perceptions of the research?

Chapter 2 will address work done to answer the first question, while chapters 3 and 4 will focus on the second question.

CHAPTER 2. STRUCTURE OF CHEMISTRY RESEARCH TEXTS

In order to explore methods of potentially improving reader understanding of science in the news, it is essential to understand how science is currently presented in various settings. Although there has been research demonstrating distortions of science in the news compared to the primary research for health and environmental news (Brechman et al., 2009, 2011; Kua et al., 2004; Nelkin, 1995; Pellechia, 1997; Stocking, 1999; Tankard & Ryan, 1974), it is not clear how well these findings apply to other science news topics. Therefore, the purpose of the first phase of this study was to investigate how non-politically controversial or health related science is presented in news reports, press releases, and research reports, as well as to examine any differences between the formats. An implicit assumption in the literature is that reader understanding would be improved if news articles were more similar to research articles. The results of this study reveal differences between the types of texts that could be added or altered in news reports to determine its impact on reader understanding.

2.1 Research Questions

The research questions for this phase are as follows:

1. What is the general structure of research reports, press releases, and science news reports that report on non-politically controversial physical science?
2. How does the structure compare among the different reporting formats?

2.2 Methods

2.2.1 Theoretical Framework

The theoretical framework that guides this phase is Language for Specific Purposes (Swales, 1984, 1990). In this perspective, texts are assumed to have an overall purpose and are grouped into genres based on that feature. For example, science research reports could be considered a genre whose purpose is to convey the results of experiments to other scientists. Texts can then be analyzed using move analysis by identifying discourse units that contribute to the overall purpose in different ways.

In this analysis, units of text are classified into rhetorical moves based on their communicative purpose. A move is then defined as a section of text that serves a specific function, but also contributes to the overall purpose of the text. Within each move, there exist a number of sub-moves that accomplish the purpose of the move, called steps (Swales, 1990). The order in which the moves and steps occur within a text defines its structure. A consensus structure for a genre can be determined by examining the frequency of presence of particular moves and their occurrence within the texts.

Studies using move analysis have typically focused on specific sections of research articles (Hopkins & Dudley-Evans, 1988; Samraj, 2002; Swales & Najjar, 1987). However, there has been some work focusing on the structure of entire biochemistry research articles (Kanoksilapatham, 2003, 2005). Kanoksilapatham divided the articles into the Introduction, Methods, Results, and Discussion sections, each composed of three, four, four, and four moves respectively. The introductions generally announced the importance of the field, prepared for the study, and introduced the study. The methods described materials and procedures. The results also stated and justified procedures, along with stating results and providing comments on those results. In the discussions, the research articles contextualized the study, consolidated results, and stated limitations. She also defined 38 steps divided among the moves (Kanoksilapatham, 2005).

Another study performed the same type of analysis on science news reports, or “Journalistic Reported Versions (JRV)” of science (Nwogu, 1991). Nwogu analyzed science news reports from *The New Scientist*, *Newsweek*, and *The Times*. Nine moves, with 23 steps, were identified in these texts. It was found that the news reports presented background information, highlighted overall research outcomes, reviewed related research, presented new research, indicated consistent observations, described data collection procedures, described experimental procedures, explained research results, and then stated research conclusions. While the moves are different and occur in different orders in the two studies, many of them can be related between the two types of texts. Both schemes include moves related to providing background information, introducing the study, reviewing related research, describing experimental procedures, explaining results, and reaching conclusions.

2.2.2 Materials

In this phase of the study, I examined three genres: science research reports, science press releases, and science news reports. In order to identify texts for analysis, a search of the LexisNexis database was conducted for news reports in the category of chemistry containing the words “new” and “research” that were published in *The New York Times*, *The Washington Post*, *Christian Science Monitor*, *USA Today*, *LA Times*, or *The Wall Street Journal* since 2000. These newspapers were chosen because they are among the most widely circulated newspapers in the United States (Alliance for Audited Media, 2013). The results of the search were further limited by only including science news reports that reported on one specific published research study. Additionally, news reports on research related to health or climate change were removed from the data set to focus specifically on other science news topics.

After identifying the news reports, associated press releases were obtained by searching the media relations websites of the institutions where the research was conducted. News reports for which no press release could be located were excluded from analysis. Finally, the original research report referenced by the news report was obtained from the journal where it was published. This resulted in 17 sets of texts, where a set includes a news report, associated press release, and original research report, for a total of 51 texts. A list of the texts analyzed is presented in Appendix A.

2.2.3 Data Analysis

As discussed previously, move analysis of research reports (Kanoksilapatham, 2005) and science popularizations (Nwogu, 1991) have been conducted, with each analysis identifying different moves. In order to compare the different genres of texts analyzed in this phase, it is important to have a coding scheme that can be used across the genres. Many of the moves identified for research reports were similar to those identified for science popularizations. Therefore, I used the results of the previous studies as a guide to develop a preliminary coding scheme for moves that could be used across genres prior to analyzing any texts by identifying overlaps in move definitions between the two studies. For example, both Nwogu and Kanoksilapatham defined a move related to describing experimental procedures, so I included "Describing data collection procedures" as a preliminary move. An additional move related to where the research was conducted, how it was funded, and other personal or social contexts, "Researcher context", was added during my analysis. After describing the moves, steps within each move were identified without referencing any prior studies using lexical clues. For example, one of the news articles contained the text

> "The intense heat of the planet immediately after it formed means that any initial water would have quickly evaporated; scientists believe the oceans emerged around 8 million years later. The puzzle is where the water, which is vital for life on Earth, came from."

which was classified as the "Presenting background information" move. The first part of the first sentence, up to the semi-colon, explains a concept as a firm statement, so it was identified as "Explaining principles and concepts", while the latter part of that sentence informs the reader about a scientific theory, so it was identified as "Knowledge in the field." Finally, the last sentence presents the general question that the research is attempting to address, so it was identified as "Introducing the problem." For definitions of the move and steps identified, see Appendix B. Refinement of the identified moves and steps was conducted with another graduate student and an undergraduate student by analyzing two sets of texts separately and comparing the overlap between the two coders. This resulted in slight alterations of the move and step definitions to more clearly distinguish between distinct moves and steps and the elimination of an initially defined move of "Introducing the problem", as it was not consistent across the different texts. An example of the move analysis for one news report is presented in Appendix C.

It is important to note that move analysis has been developed for textual analysis, so any non-textual information was not coded, along with captions for it. The abstract, figures, tables, captions, and any supplementary information were not considered for coding in this analysis. News reports and press releases do not usually have an equivalent to an abstract or supplementary information as found in a research report, so these were not coded in order to maintain consistency.

Upon completing the move analysis, the consensus move structure of each genre was determined. Each individual text differed to some extent in structure, determining the consensus structure is needed to understand the general structure of a genre. Since the number of moves in a text varied, the moves in each text were split by occurrence into a number of groups equal to the most common number of moves within the genre. For example, a text containing 14 moves in a genre with a mode of 7 moves would have a group consisting of moves 1 and 2, another of moves 3 and 4, and so on for a total of 7 groups. In the event that the number of moves in a text was not a multiple of the most common number, the following equation was used to determine the number of moves within a group:

$$move_{\max} = (\text{group}) \times \left(\left\lfloor \frac{\text{total moves}}{\text{mode(moves in genre)}} \right\rfloor \right)$$

where $move_{max}$ is the upper cutoff of moves to be placed in a group, with the lower cutoff being the move after the upper cutoff of the previous group. For example, for a text with 9 total moves and the mode of moves in the genre is 6, the first group would contain move 1 and the second group moves 2 and 3, etc. After splitting the moves for all the texts into groups, the moves were assigned numerical values and the mode of each group within a genre was determined. The mode of the first group is considered the consensus first move for that genre. In some cases, the mode of consecutive groups was the same, so the number of consensus moves was less than the most common number of moves of texts in that genre.

2.3 Results

The consensus move structure of each the genres analyzed is presented in Table 2-1. Overall, five consensus moves were identified for news reports and research reports, while press releases had an additional move. All genres tended to present background information, describe data collection procedures, explain research outcomes, and state research conclusions at some point in the text, while only news reports and press releases tended to highlight overall research outcomes. It was found that news reports and press releases are incredibly similar in their general structure. The major difference was that press releases very often discuss the researcher context, possibly to highlight the accomplishments of their institution, while news reports do not. It is also important to note that press releases tend to explain the results of the research before discussing the methods used. In contrast to research reports, both news reports and press releases begin by highlighting overall research outcomes rather than presenting background information. Both of these types of texts open with the results, while research reports begin by framing the context of the study. In addition to presenting background information, research reports review related research to contextualize the study, while the other genres do not.

Table 2-1: Consensus Move Structure

Move	News Report	Press Release	Research Report
1	Highlighting overall research outcome	Highlighting overall research outcome	Presenting background information
2	Presenting background information	Presenting background information	Reviewing related research
3	Describing data collection procedure	Explaining research outcome	Describing data collection procedure
4	Explaining research outcome	Describing data collection procedure	Explaining research outcome
5	Stating research conclusions	Stating research conclusions	Stating research conclusions
6		Researcher context	

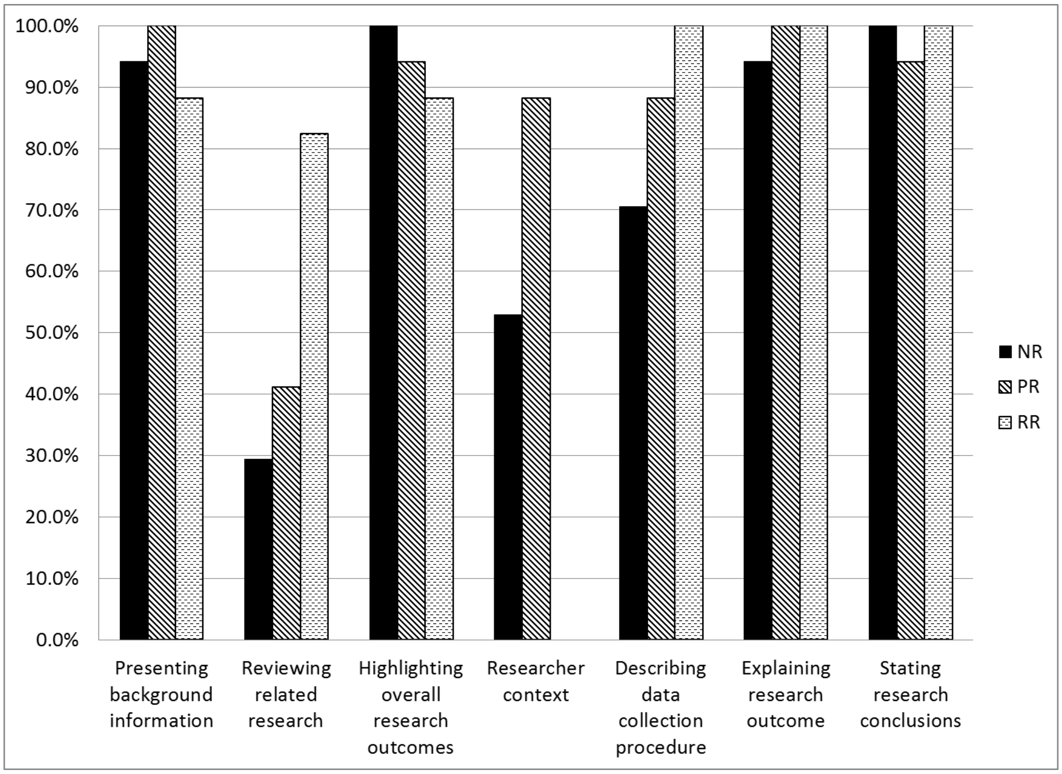


Figure 2-1 Percentage of Texts Containing Each Move
NR: News Report, PR: Press Release, RR: Research Report

While Table 2-1 shows the consensus moves for each of the genres, those moves are not the only ones present within a genre. Figure 2-1 shows the percentage of texts within a genre that contain a particular move. All of the moves are present in some news reports and press releases, while all but researcher context are present in some research reports. While the consensus moves might imply that news reports and press releases never review related research, which is inaccurate. A few texts in these genres do make mention of related research (Figure 2-1). It is important to note that while highlighting overall research outcomes is present in many research reports, it is not one of the consensus moves because its location within the text was not consistent and it did not comprise much of the research reports (Figure 2-2).

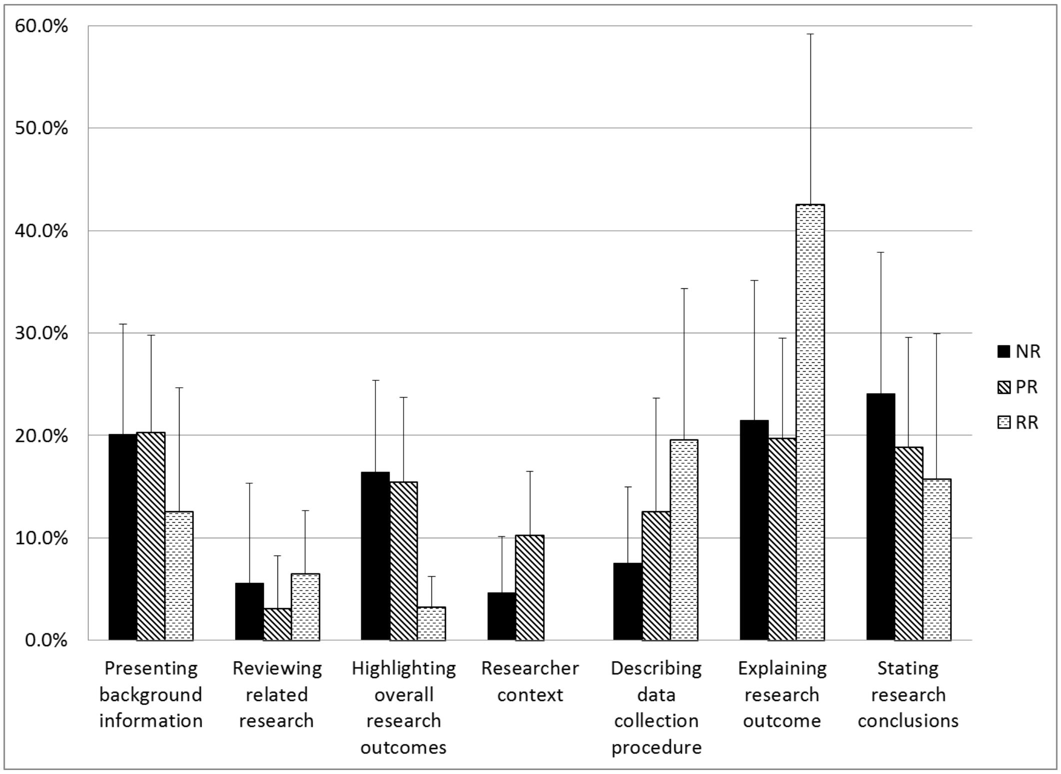


Figure 2-2 Average Percentage of Text within Each Move
NR: News Report, PR: Press Release, RR: Research Report; Error bars represent one standard deviation

Figure 2-2 shows the average amount of text within each move for each genre. This was calculated by comparing the word count within a move to the total coded word count for each text and averaging across texts within a genre. It is evident that even though most texts in all the genres present background information, highlight overall research outcomes, and state research conclusions (Figure 2-1), the news reports and press releases devote a larger percentage of their text to those moves than the research report does (Figure 2-2). The research reports allocate more space to describing data collection procedures and explaining research outcomes.

Table 2-2 Percentage of Texts Containing Each Step within the First Four Moves.

Move	Step	NR	PR	RR
Presenting background information	Explaining principles and concepts	82%	82%	88%
	Knowledge in the field	18%	12%	12%
	Introducing the problem	59%	53%	29%
	Potential Implications	0%	24%	0%
Reviewing related research	Reference to previous research	29%	41%	82%
	Indicating limitations of previous research	6%	18%	53%
Highlighting overall research outcomes	Indicating main research result	100%	94%	88%
	Implications	53%	76%	6%
	Practical Limitations	0%	6%	0%
	Scientific Limitations	6%	0%	0%
	Referencing setting/publication	29%	47%	0%
	Summary of method	18%	6%	41%
	Anecdote	6%	0%	0%
Researcher context	Referencing setting	29%	71%	0%
	Referencing publication	18%	65%	0%
	Funding	0%	47%	0%
	Anecdote	24%	6%	0%

At the move level of analysis, some differences between the different text genres were revealed, but there does not appear to be massive changes from the general structure of the research article to how it is presented in the news. However, analysis of the steps within each move provides some additional insight. Tables 2-2 and 2-3 show the percentage of texts that contain the steps identified. Within the move of presenting background information, four steps were identified. In each of the genres, the most common way to accomplish the purpose of presenting background information is to explain principles and concepts, with this step appearing a similar rate within the genres. Occasionally, the texts make some reference to the broader field of research in this move by discussing knowledge in the field. It appears that the main difference in steps within this move is that research reports are less likely to provide an explicit explanation of why the research is being done by introducing the problem. Press releases sometimes also include potential implications related to introducing the problem.

Table 2-3 Percentage of Texts Containing Each Step within the Last Three Moves.

Move	Step	NR	PR	RR
Describing data collection procedure	Materials	35%	29%	76%
	Experimental setup	59%	71%	94%
	Explanation of experiment	12%	53%	65%
	Data collected	6%	6%	29%
	Explaining principles and concepts	6%	18%	24%
Explaining research outcome	Explaining principles and concepts	41%	47%	82%
	Stating specific outcome	94%	94%	100%
	Commenting on result	71%	65%	94%
	Reference to previous research	0%	0%	29%
Stating research conclusions	Summary of results	18%	35%	88%
	Practical Limitations	24%	12%	12%
	Scientific Limitations	6%	0%	24%
	Future work	24%	24%	24%
	Implications	88%	94%	88%
	Research context	6%	18%	0%
	Explaining principles and concepts	18%	41%	53%
	Reference to previous research	0%	12%	29%
	Speculation	35%	24%	29%
	Anecdote	24%	18%	0%
	Practical Applications	12%	6%	0%

The steps within the move of reviewing related research show more differences between the genres. As the original science information reported gets further removed the original researchers, both references to previous research and indicating the limitations of previous research appear less often. While it may not be completely necessary to read about previous research in order to understand a specific study, it may have an impact on how one views science in general. It is a norm of scientific research writing to review related research to situate the study in the broader field of researcher and to demonstrate where there is a gap in the field that the researchers will address (Robinson, Stoller, Constanza-Robinson, & Jones, 2008). This norm helps emphasize that research is done within a broader context, though that is not the primary purpose of its use. Studies have shown that students often view science as a collection of facts and not embedded in a social context (Deng, Chen, Tsai, & Chai, 2011; N. G. Lederman, 2007). This may persist into adulthood and not be challenged by the presentation of science in the news.

When highlighting the overall research outcomes, all genres often indicated the main research results. This step is the main way that this move is accomplished. In addition to stating what the main result of the research was, the news reports and press releases provided additional information. These genres made reference to the implications of the research and the setting or publication much more often than the research reports did. For both steps, the press releases were more likely to contain the step than the news report. The research reports also provided additional information, specifically summarizing the methods used in the study more often than the other genres.

While all of the genres highlighted the results, the news reports and press releases were also focused on where the research took place and what the results meant, while the research reports focused on the methods used. This decreased emphasis on methods may impact how readers interpret the information in news reports, particularly if they are unsure what the researchers actually did to obtain their results.

The move of researcher context was only present in the genres of news reports and press releases. While research reports do provide information about the employer of the authors when they are listed, that information was not analyzed. Press releases were much more likely to reference the setting, publication, and funding sources of the research than news reports. This may be because one of the goals of press releases is to highlight the accomplishments of their institution. News reports were more likely to provide an anecdote about the researchers, most likely to add some human interest to the article. For example, one of the news reports included the following text

> “Mano Misra, a professor of engineering who conducted the research with Narasimharao Kondamudi and Susanta K. Mohapatra, said it was by accident that he realized coffee beans contained a significant amount of oil. "I made a coffee one night but forgot to drink it," he said. "The next morning I saw a layer of oil floating on it."”

to presumably make the researchers more relatable to the audience.

When describing data collection procedures, the farther away one gets from the researchers’ original text the less likely each of the steps is to be present. News reports do not often provide much information about the methods involved in conducting the research being reported. Information about the procedure is sometimes provided, but an explanation of why the researchers chose to collect data in that way is rarely discussed, while research reports are much more explicit about both. Overall, the research reports were more likely to provide a detailed description of the methods, along with explanations of what information was obtained using those methods. As in the case of the summary of the methods when highlighting overall research outcomes, this de-emphasis on methods may impact reader understanding. It may be difficult to judge the quality of science information or conclusions without being provided information about how the results reported were generated. A non-expert probably would not be able to completely understand the details of every method used in scientific research, but a greater emphasis on explaining the methods may be warranted.

When discussing the results of the research, it was found that all genres are very likely to simply state the findings. However, research reports are much more likely to comment on results, provide additional background information, and reference previous research compared to the other genres. News reports and press releases tended to state results without necessarily indicating that the researchers had to interpret the data collected. The research reports usually stated a result, followed by an explanation of what the researchers believed the result meant. In addition, texts in this genre also provided information to aid the reader in interpreting results. While it is not likely that these differences in presentation would affect readers' understanding of the results, the news reports and press releases present the science in a manner more consistent with science being a series of facts rather than involving interpretation, as also seen in science textbooks (Abd-El-Khalick, Waters, & Le, 2008; Chiappetta & Fillman, 2007). As with reviewing related research, this difference may not challenge potential misconceptions adults have about the nature of science.

All of the genres generally finished by stating the research conclusions, though the emphasis of this section was different between the types of texts. News reports were more likely to focus on the applications and implications of the research as it pertained to societal impact, while the research reports focused on summarizing results and discussing implications as they pertained to field of research. Press releases were some combination of these two, with an additional emphasis on the context of the research.

A potentially important difference is how the genres treat the limitations of the research. News reports were more likely to discuss limitations related to applying the research results to society than to those related to the scientific process. Research reports had the opposite trend. While only about a quarter of the research reports explicitly discussed the scientific limitations, that may be more due to the authors' assumption that the readers are professional scientists who may be able to recognize these types of limitations readily than to there not being limitations. It is potentially problematic that news reports de-emphasize discussing scientific limitations, as many readers of those texts probably do not have extensive experience judging the quality of scientific research. Therefore, presentation of these limitations may aid readers assess the validity of the claims made in scientific research.

2.4 Conclusions

In conclusion, the overall general structure of science texts in the three genres analyzed was fairly similar, with the main differences occurring in the beginning of the texts. The results of the move analysis of research reports presented here is consistent with previous studies of different types or parts of research reports (Ayers, 2008; Hopkins & Dudley-Evans, 1988; Kanoksilapatham, 2005; Skelton, 1994). The results for news reports were also fairly similar to a previous study examining similar texts, but in different publications (Nwogu, 1991), with the only major difference being that this analysis found that news reports tend to begin by highlighting overall research outcomes rather than presenting background information.

While there was some variation in how much of the texts were devoted to each move, it is clear that research reports describe data collection and present results more than the other genres. News reports place more emphasis on stating the research conclusions, with press releases somewhere in the middle. Analysis of the steps provided further evidence of the de-emphasis of the methods and results in favor of the conclusions. This could affect reader interpretation of the claims made by the researchers and should be examined. The second phase of this study explores how one aspect of the differences identified in this analysis may impact readers.

CHAPTER 3. THE EFFECT OF DESCRIBED STUDY LIMITATIONS ON PUBLIC UNDERSTANDING

The results of the first phase informed the second, with the purpose of examining how the explicit inclusion of scientific limitations in science news reports with little health or political focus affects different aspects of readers' understanding of the article. Very few news articles make explicit mention of the scientific limitations of research studies, which may impact how readers interpret the information presented. As mentioned previously, the *National Science Education Standards* (NRC, 1996) indicates that a scientifically literate person should be able to critique the quality of science information and apply conclusions appropriately, which may be difficult if the reader is not provided with information about the limitations of the study. This phase examines how the inclusion of the limitations could impact readers' understanding of the science content, trust in the results and conclusions, perceived significance of the research, and ability to evaluate additional claims.

It is not expected that including limitations would impact reader understanding of the content of the news report, but some confusion could potentially arise related to the results or conclusions. Previous work suggests that including hedged language, such as limitations, in cancer news articles causes readers to view the journalists and researchers as more trustworthy (Jensen, 2008; Jensen et al., 2011), so it may also affect reader trust in the results or conclusions. However, there is some evidence that limitations do not affect reader judgments of credibility of the science information (Maxim & Mansier, 2014). In that study, science background and views about science were the main influences on credibility judgments. Conversely, the limitations could cause readers to devalue the research because flaws are explicitly stated, resulting in a lower perceived significance. Finally, the inclusion of limitations could help readers evaluate claims made about the research by providing additional information for the reader to use in their judgment. If true, this would suggest that more emphasis should be placed on discussing scientific limitations in news reports of research. In addition to variations in the text, the reader's scientific background could also influence each of the above outcomes, so comparisons are made between professional scientists and lay readers.

3.1 Research Questions

The research questions for this phase are as follows:

1. How does the inclusion of study limitations in a science news article affect
    a. Readers' understanding of the content of the article?
    b. Readers' trust in the results and conclusions of the article?
    c. Readers' perceptions of the significance of the research?
    d. Readers' ability to evaluate claims based on the research?
2. How does the reader's science background affect the above outcomes?

3.2 Methods

3.2.1 Participants

The participants in this study were drawn from two groups. The first were administrative, professional, clerical, and service staff (n=12) employed in non-STEM departments at a large Midwestern public university. This population was recruited by placing an advertisement in the digital internal university newsletter. The staff participants (10F, 2M) ranged from 22 to 64 years old. The second group consisted of tenured science faculty (n=6) employed at the same large Midwestern public university. This population was recruited by direct email request. Of the science faculty, two were from the chemistry department, two from biology, one from physics, and one from earth, atmospheric, and planetary sciences. The faculty (1F, 5M) ranged from age 35 to over 65. Participants were randomly assigned a four-digit number as a de-identifier.

3.2.2 Design

All participants (N=18) were randomly assigned to read one of two news reports reporting on the same published chemistry research report. One of the articles explicitly discussed the limitations of the research study, while the other did not. Participants completed an online survey, then read one of the two news reports and participated in a short, semi-structured interview, during which they were asked questions related to the article they read.

3.2.3 Articles

The articles used in this study were chosen by searching the LexisNexis database for alternate versions of the news articles selected in phase I that reported on the same research study, but differed in the inclusion of the study limitations. This was done to compare articles as they occur naturally, rather than creating an artificial version of a science news report. The first article used, “Chemical analysis of a comet's ice gives a clue to source of water on Earth” from the October 11, 2011 issue of the Washington Post, contains no explicit mention of study limitations (NL). The second article, “Where did Earth's water come from?” in the October 16, 2011 issue of the Christian Science Monitor, describes the same research study and explicitly mentions study limitations (DL). Otherwise, the science content in each article is similar. The articles can be found in Appendix D. The publication, author, and date were removed from the texts to limit any bias the readers may have related to the newspaper.

3.2.4 Measures

Prior to reading one of the news reports, participants completed a survey consisting of three parts. The first part contained questions related to the participants’ science education, news reading habits, and science information seeking habits (see Appendix E). The second part was a science knowledge test (National Science Board, 2010) containing both content and reasoning questions (see Appendix F). The last part was the Views of Science Test (VOST) (Hillis, 1975) which measures an individual’s views about the tentativeness of science (see Appendix G).

After reading the news report, semi-structured interviews were conducted, in which some of the questions related to the content of the article, participants' opinion of how much they trusted the research results and conclusions, their opinion of the significance of the research, and their evaluation of claims made about the research study (see Appendix H). Additional questions mirroring the survey questions, with some adapted from the Views of Nature of Science questionnaire version C (N. G. Lederman, Abd-El-Khalick, Bell, & Schwartz, 2002), were asked to compare the survey results to open-ended interview responses. A correct response for each of the content questions was generated and responses were scored by the author using a rubric modified from work evaluating student understanding of chemistry concepts (Abraham, Grzybowski, Renner, & Marek, 1992), presented in Table 3-1. The perceived significance and trust questions were open coded to analyze the participants' reasons for their responses.

Table 3-1 Content Question Rubric

Score	Criteria
6	Contains all aspects of the correct response
5	Missing few aspects of the correct response
4	Missing many aspects of the correct response
3	Part of the response is correct, but it contains incorrect or unrelated statements
2	Response entirely incorrect or unrelated to the question
1	Don't know or no response

For evaluation of claims, participants were given statements that drew some further conclusion from what was presented in the article. Participants were then asked to indicate how reasonable or unreasonable they felt each statement was and why they felt that way. The statements, along with the ideal response and reason, are presented in Table 3-2. Subsequent references to these statements will use the labels in this table.

Table 3-2 Evaluation of Claims Statements

Label	Statement	Ideal Response	Reason
S1	Other comets from the same region of space as Hartley 2 will have a similar heavy water to light water ratio	Unsure	It is unknown how representative the measurements were of other comets
S2	Comets provided a larger portion of the initial water to Earth than was previously believed.	Disagree	This statement makes much broader claims than the authors of the original study
S3	The scientists must now reconsider what they thought they knew about water on comets.	Agree	The research findings did not fit into the current model

Table 3-3 Evaluation of Claims Rubric

Score	Criteria
6	Reasonable conclusion and logical reasoning based on evidence
5	Reasonable conclusion and logical reasoning but no evidence
4	Not reasonable conclusion but logical reasoning
3	Reasonable conclusion but unclear reasoning
2	Reasonable conclusion but unreasonable or missing reasoning
1	Not reasonable conclusion and unreasonable or missing reasoning

Participant responses were scored using a six-point rubric that assessed the reasonableness of the participants' answer and their reasoning (Table 3-3). Refinement of the rubrics for the content questions and the evaluation of claims statements were done with another graduate student by scoring responses from a subset of participants separately and comparing ratings. This resulted in clarifying what was considered a correct answer for the content questions and evidence of logical reasoning for the evaluation of claims.

3.3 Results

Overall, participants scored very well on the science knowledge test (Table 3-4), with the staff scores being consistent with those observed for members of the general population with at least a bachelor's degree (National Science Board, 2010). Although participants were randomly assigned to either the described limitations (DL) or no limitations (NL) group, staff members in the DL group appear to be more knowledgeable about science than staff members in the NL group. Scores for the VOST indicate that members of the DL staff group have similar views about the tentativeness of science as the faculty participants, while those in NL staff group view science as less tentative.

Table 3-4 Average Scores for the Science Knowledge Test and Views of Science Test. Averages for participants in the Described Limitations (DL) or No Limitations (NL) groups and overall. A higher VOST score indicates viewing science as more tentative.

		Content	Reasoning	VOST
Staff (n=12)	EL	85%	93%	143.0
	NL	69%	83%	129.7
	Overall	77%	88%	136.3
Faculty (n=6)	EL	96%	90%	147.3
	NL	100%	100%	140.7
	Overall	98%	95%	144.0

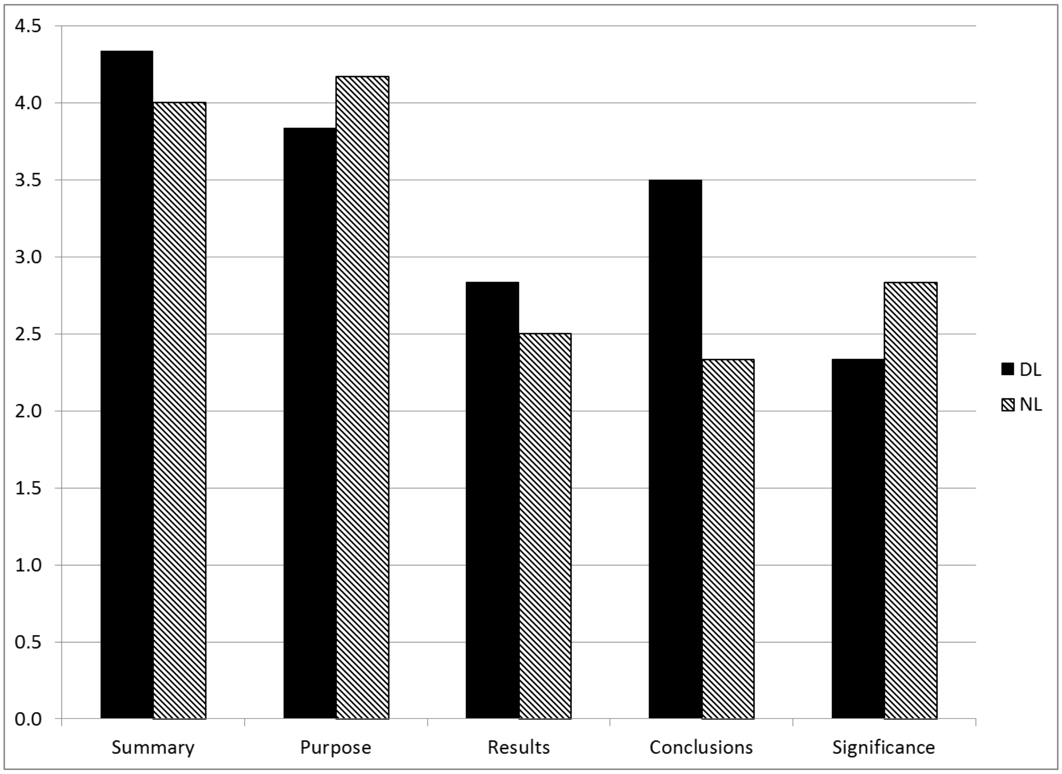


Figure 3-1 Average Score for Content Questions for Staff Participants

After reading one of the news articles, participants were asked a series of questions related to their understanding of the content of the research. The average scores for the staff participants are presented in Figure 3-1. For each of the content areas, other than conclusions, participants scored similarly regardless of the version of the article read. Those that read the article including described limitations provided more accurate explanations of the conclusions of the research than those that read the other version. Half of the participants in the NL group responded with an incorrect claim that was broader than what the researchers stated. One of the staff participants provided a typical response of this type by stating

"Ah, the conclusions was that, um, that comets were a source of water for the Earth."

None of the participants in the DL group made this type of claim. This suggests that one possible effect of the inclusion of limitations may be to help participants focus on the correct conclusions of the research, rather than causing confusion.

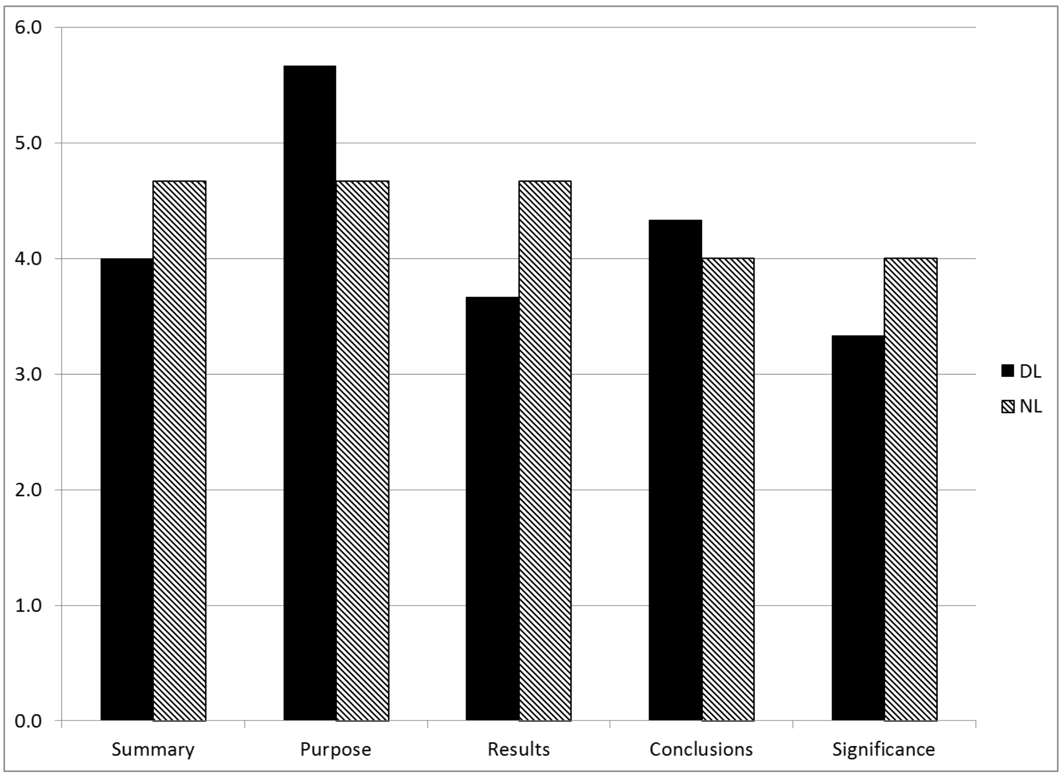


Figure 3-2 Average Score for Content Questions for Faculty Participants

In contrast to the staff participants, faculty responses did not show substantial differences between the two types of articles (Figure 3-2). They also scored higher overall for each of the questions. The disappearance of the difference in identifying the conclusions may be due to the greater science background of the faculty compared to the staff. It would be reasonable to assume that they read more science research than the staff and so have more practice identifying conclusions.

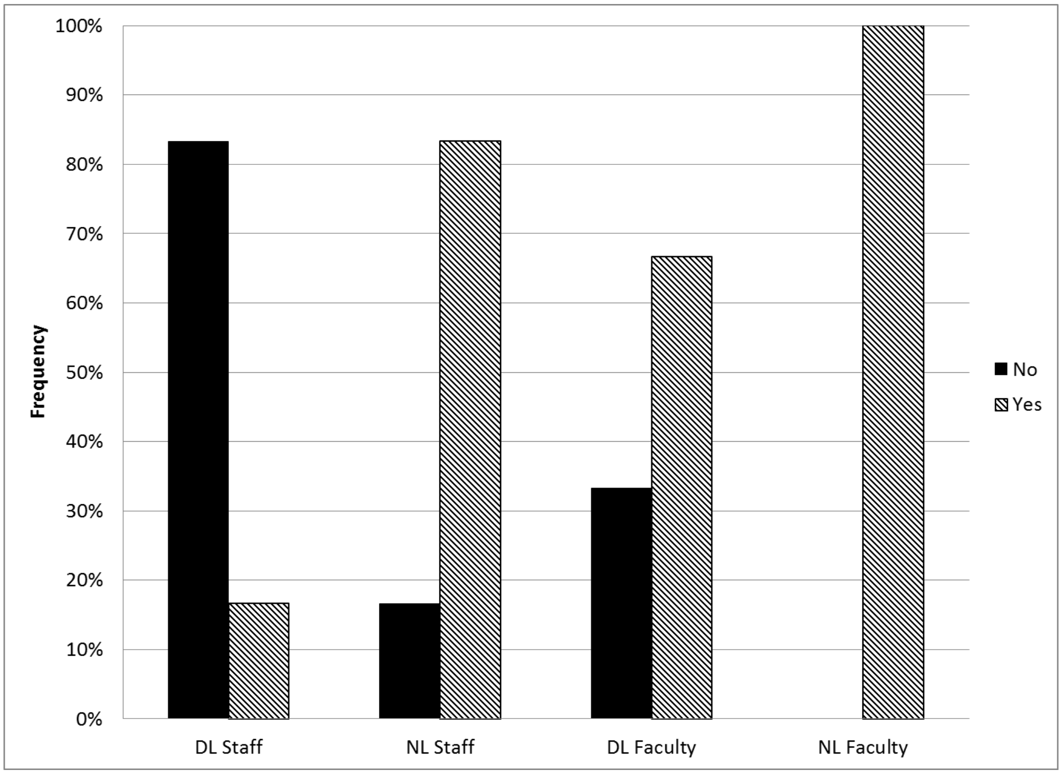


Figure 3-3 Participants' Perceived Significance of the Research
Responses were to the question "In your opinion, do you feel that the research findings were important?" DL: Described Limitations, NL: No Limitations

When participants were asked if they felt that the research findings were important, large differences were observed for the staff based on the type of article read, but not for the faculty (Figure 3-3). The staff in the DL group generally did not feel that the research findings were important, while those in the NL group did. The faculty generally felt that the findings were important. While this major difference may seem to be caused by differences between the two articles that were read, an analysis of the reasons participants' stated reasons for their responses indicates that this is likely not the case. All of the participants who felt the findings were important indicated that they did so because learning new information has some intrinsic value. For instance, one of the staff participants stated

> "Um, cuz it's good to know where it came from, how it started."

and one of the faculty participants stated

> "Um, I think it is important to understand what's going on in our solar system and, um, the universe and trying to figure out where, where things have come from and where they've came from is important."

For the participants who did not feel the research was important, most indicated that it was not personally relevant. For instance, another staff participant stated

> "I don't think they tied it in, for like, a regular person to appreciate the research… It didn't really relate to me, I don't think."

Only a single staff participant specified a reason related directly to the research,

> "I mean, I don't think it's, they have enough facts to prove it or to say that this really is true."

The vast majority of responses indicated that the reason for evaluating significance was related to personal values rather than to the information provided in the articles. This suggests that the inclusion of limitations had no effect on how readers perceive the significance of the research.

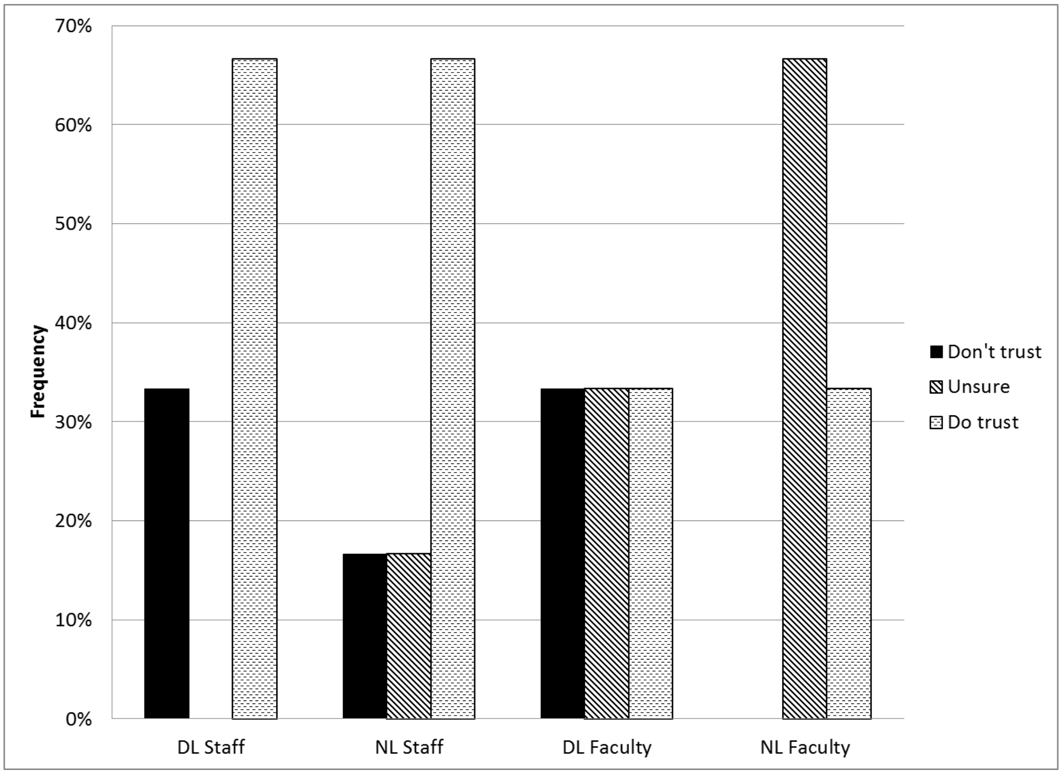


Figure 3-4 Participants' Trust of the Research Findings
Responses were to the question "How much do you trust the research findings from the article?" DL: Described Limitations, NL: No Limitations

When asked how much they trusted the findings of the research, staff participants generally trusted the findings regardless of article read, while the faculty participants were unsure (Figure 3-4). Overall, there does not appear to be much of a pattern within the staff or faculty groups. As with perceived significance, many of the reasons provided for the level of trust were not directly related to the content of the research. The most common reason for trusting the findings was that the participants deferred to the experts.

For instance, a staff participant stated that their reason was

> "…because I'm sure they're experts in their field, you know, and very knowledgeable about this, a lot more than me"

and a faculty participant stated

> "Well, it's not my area, so I have to trust."

Those that were unsure of how much to trust the findings wanted more information before making a judgment. This is evidenced by another faculty participant stating that

> "I would have to, often with, with science writing, when I see something that really interests me, I go and I find the paper. Now this is way outside my field. I don't, I don't know what someone who thinks about planet formation or evolution would say about the significance of the D to H ratio alone, so I would have to look, I would have to look into it and see."

Finally, the most common reason for not trusting the findings was that not enough research was done. One of the staff participants stated

> "And then there's only that study or one other study that's been, there's not many examples and experiments that have been taking place, so you need multiple replications to really prove something."

Based on the indicated levels of trust and the stated reasons, it is not evident that the inclusion of limitations had any effect on the responses. It appears that participants relied more on their own views of research to make their judgments than on the content of the article. However, trust in the conclusions may be more influenced by the information in the article.

When asked how much they trusted the conclusions of the research, staff participants in the NL group all trusted them, while those in the DL group were less sure (Figure 3-5).

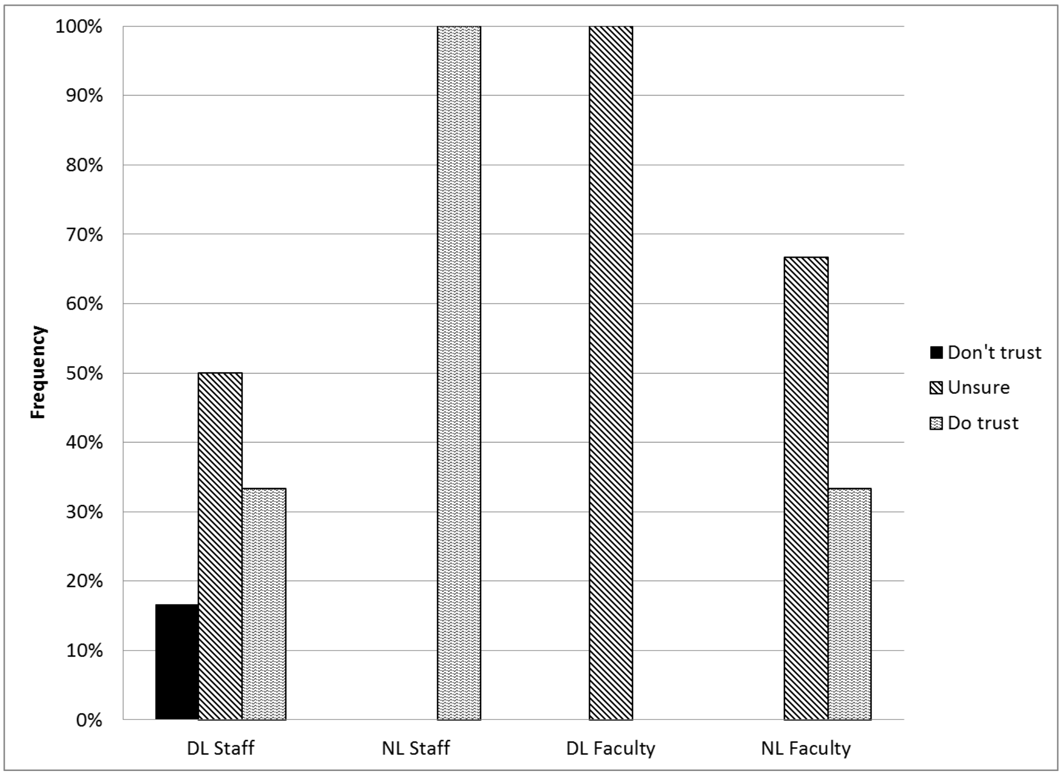


Figure 3-5 Participants' Trust of the Research Conclusions
Responses were to the question "How much do you trust the research conclusions from the article?" DL: Described Limitations, NL: No Limitations

Overall, the faculty participants were unsure how much to trust the conclusions. When determining levels of trust in the conclusions, staff participants in the NL group trusted expert opinions. For instance, a staff participant stated

> "…I think that, given the fact that, ah, that this was a published article and I assumed i – it was refereed appropriately that I would trust the conclusions."

However, staff participants in the DL group felt that the conclusions were unclear and were, therefore, unsure how much to trust them. For instance, another staff participant stated

> "I don't think they really had a good conclusion. It was pretty much another theory. Like, oh, it might have came, more water might have came, from asteroids versus comets, but we don't know how the impact would have happened for the water to get here, so it seemed like they were still questioning what they were researching.."

In addition to those that were unsure, a staff participant in the DL group didn't trust the research conclusions because she felt not enough research was done. She stated

> "Just a little. Just cuz I don't think that there were enough observations to make it believable."

It appears that staff participants who read the article with described limitations were more skeptical of the conclusions than those who read the article without limitations. This difference may be due to the effect of the inclusion of limitations on participants' perceptions of the conclusions and on their recognition of the study limitations. Some participants were unsure how much to trust the conclusions because they seemed to have difficulty dealing with the tentativeness of the researchers' conclusions. The article without described limitations made the results of the study seem less tentative, which may be a reason that participants in that group were not unsure.

While the sample size is small, some interesting trends are apparent. For the staff participants, the type of article read seems to impact how well they were able to evaluate the claims (Figure 3-6). Participants that read the DL article scored better than those that read the NL article for each statement. For S1, every participant in the DL group provided a reasonable answer and almost all provided logical reasoning, while some in the NL group provided unreasonable answers. For S2, most of the participants in the DL group provided reasonable answers with unclear reasoning, while almost all of the participants in the NL group provided unreasonable answers. Only a single participant in the NL group provided a reasonable answer to this statement. For S3, all of the DL participants provided reasonable answers with logical reasoning, while many NL participants were unable to provide clear reasoning. These differences may indicate that not including described limitations in news reports may affect what readers view as reasonable claims to make about the research.

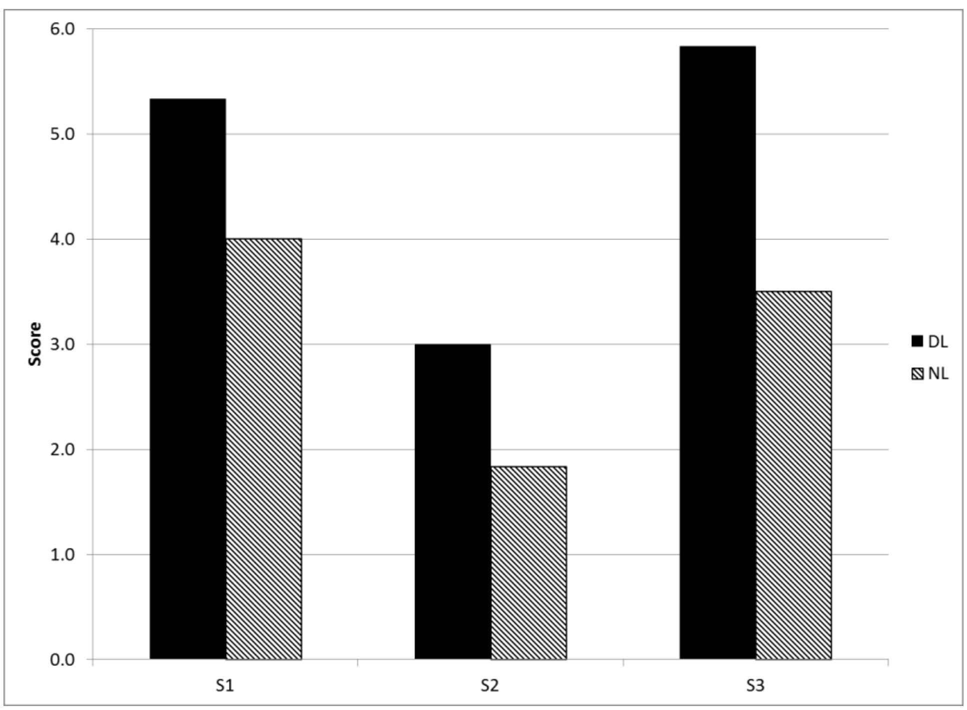


Figure 3-6 Average Score for Evaluation of Claims for Staff Participants.
S1 – Statement 1, S2 – Statement 2, S3 – Statement 3
DL: Described Limitations, NL: No Limitations

For the faculty participants, the type of article read seemed to have no impact on how well they were able to evaluate claims (Figure 3-7). Participants generally provided reasonable answers with logical reasoning to all of the questions. The disappearance of the trend observed with the staff participants may be due to the faculty's significantly greater science background. This may have allowed them to correctly identify appropriate claims based on the data without having to rely on any explicit mention of the limitations of the study.

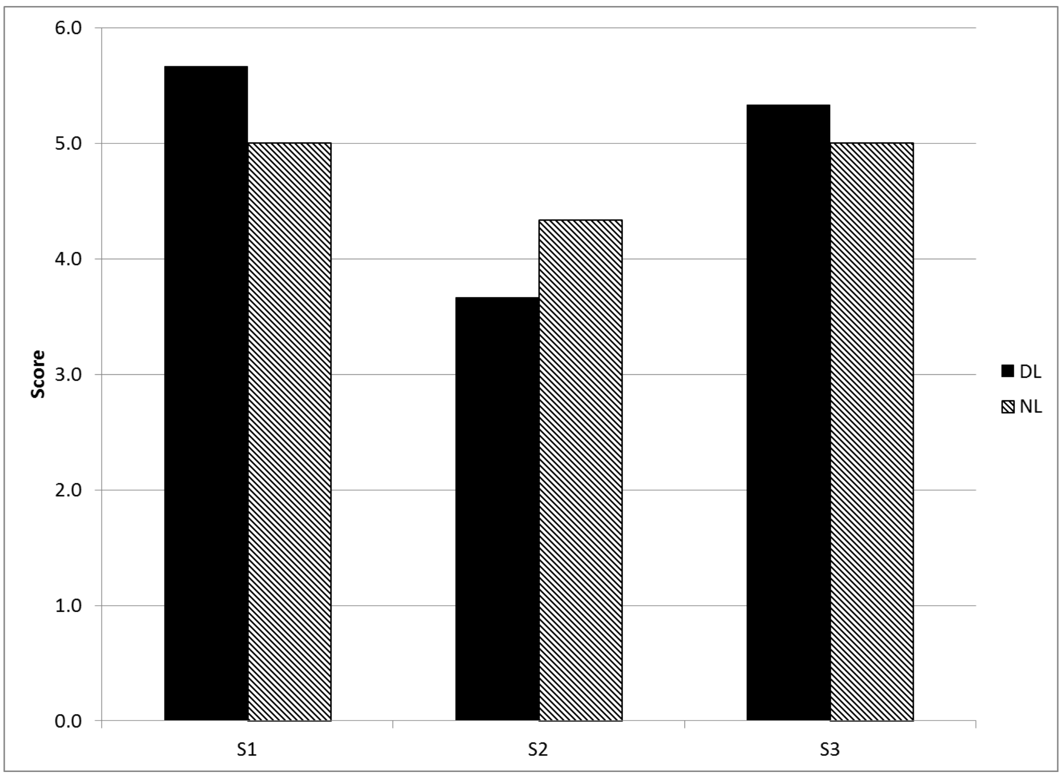


Figure 3-7 Average Score for Evaluation of Claims for Faculty Participants.
S1 – Statement 1, S2 – Statement 2, S3 – Statement 3
DL: Described Limitations, NL: No Limitations

3.4 Conclusions

The inclusion of limitations in the science news article may influence readers' ability to identify the conclusions, their trust in the conclusions, and ability to evaluate claims made about the study. It did not affect the perceived significance or trust in the findings. The effect on the trust in the conclusions was the opposite of what was observed for trust in journalists and researchers from previous studies (Jensen, 2008; Jensen et al., 2011). This suggests that trusting information may be separate from trusting the source of that information.

Although these differences were evident for the staff participants, they were not present for the faculty participants. This suggests that the inclusion of limitations may be more influential for readers with less extensive science backgrounds, while not affecting those with more extensive science backgrounds. Therefore, it may be worthwhile to include more information about the limitations of research studies in science news articles.

While this phase did indicated some interesting trends related to the inclusion of limitations, the sample size was small so it is not known how reliable or generalizable the results are. In addition, all of the participants had attained a least a bachelor's degree, so it is not clear how individuals without a college degree would be affected. Finally, the choice of articles may have also influenced the results. Only one pair of articles was used, so it is not known how much the results depended on the particular texts chosen. Also, the articles were of significantly different length (NL: 330 vs. DL: 859 words) because they were chosen to be naturally occurring, rather than artificially editing a single text. The additional length provided more background information, rather than more detail of the study, but may have had an influence on the results. The third phase of this overall study is designed to address these concerns and expand upon the results of this phase.

CHAPTER 4. THE EFFECT OF DESCRIBED METHODS AND STUDY LIMITATIONS ON PUBLIC UNDERSTANDING

The results of the previous chapter suggest that there is some relationship between presence of limitations and trust and evaluating claims, but the limitations of sample size, use of a single article topic, and the demographic composition of the participants make drawing general conclusions difficult. Therefore, these interesting trends needed to be confirmed with a larger sample size, additional articles, and a more varied population. In addition to examining the effect of the inclusion of limitations, it is also of interest to investigate the effect of including a more detailed description of the methods used in the reported research. The results from phase I (Ch. 2) indicated that the methods are de-emphasized in science news articles and some participants in phase II (Ch. 3) independently expressed a desire for more information about how the research they read was conducted. This is consistent with information requested by university students when reading science news briefs (Korpan, Bisanz, Bisanz, & Henderson, 1997). Other participants indicated they wanted more information about the research in general, which may have included a discussion of the methods. Therefore, it is also worth exploring how the inclusion of methods impacts the readers.

The previous results suggest that the inclusion of limitations did not affect readers' understanding of most of the content, with understanding of the conclusions being the only possible exception. It also appears that perceived significance and judgments of trust in the findings were mainly due to personal factors, rather than the manipulation of the articles. Therefore, this phase examined the outcomes that may be related to inclusion of limitations, specifically understanding of the conclusions, trust in the conclusions, and interpretation of limitations.

4.1 Research Questions

The research questions for this study are as follows:

1. How does the inclusion of study limitations in a science news article affect
    a. Readers' understanding of the conclusions of the article?
    b. Readers' trust in the conclusions of the article?
    c. Readers' ability to evaluate claims based on the research?
2. How does the inclusion of an explanation of the methods in a science news article affect
    a. Readers' understanding of the conclusions of the article?
    b. Readers' trust in the conclusions of the article?
    c. Readers' ability to evaluate claims based on the research?
3. How does the inclusion of study limitations interact with an explanation of the methods to affect the above outcomes?
4. How does the reader's science background affect the above outcomes?

4.2 Methods

4.2.1 Participants

In order to answer the above research questions, participants representing the general American adult population were needed, as well as science faculty at research universities to serve as a comparison group. A panel of 250 participants was purchased from Qualtrics Panels to recruit general public participants. Participants were chosen to be representative of the United States adult population, with sampling based on age and education. They were compensated by Qualtrics for participation. A total of 232 public participants provided complete responses that were included for analysis.

In order to recruit science faculty, five research universities from each of the four geographic regions of the United States (Northeast, Midwest, South, and West) were randomly selected, for a total of 20 universities. At each institution, all of the faculty members in natural science (biology, chemistry, physics, astronomy, and earth science) departments were contacted using their publicly available email addresses to participate in this study. The faculty members were only emailed a single time and were not provided compensation. A total of 191 faculty participants, with a discipline distribution equivalent to the emailed sample, provided complete responses that were included for analysis.

4.2.2 Design

This study used an online survey in a 2 (described limitations vs. no limitations) x 2 (described methods vs. no methods) x 3 (article topic) design. Using multiple articles allowed for the examination of the stability of the results across different texts. All participants (n=232) were randomly assigned to read one science news article reporting on a specific chemistry research study. Participants provided background information, then read one of the news articles, and finally answered survey questions based on the article.

4.2.3 Materials

The articles used in this study were chosen from the list of articles selected for analysis in phase I, as described in Chapter 2. Three articles were chosen, "Chemical analysis of a comet's ice gives a clue to source of water on Earth" from the October 11, 2011 issue of the Washington Post (Comet Water), "Molecular Action May Help Keep Birds on Course" from the May 5, 2008 issue of the Washington Post (Bird Compass), and "In Space, Clues to the Seeds of Life" from the January 30, 2001 issue of the Washington Post (Space Membranes). The first article was selected because a version describing limitations was available. The other two articles were the only others that were sampled in phase I that discussed the scientific limitations. Using articles that already mention limitations allows for fewer confounding variables because it limits the amount of additions necessary to the texts. Short amounts of text was either added, in the case of the first article, or deleted to create alternate versions of the articles regarding limitations. In order to create alternate versions regarding methods, text from a related press release and/or research article was adapted and added to the news article (see Appendix I). All of the edited versions were reviewed by the Senior Writer/Editor of a large Midwestern public university's marketing and media news service for style.

4.2.4 Measures

Prior to being asked to read one of the study's news articles, participants completed a survey consisting of three parts. The first part contained questions related to the participants' age, sex, education, news reading habits, and science information seeking habits. The second part was the science knowledge test (National Science Board, 2010) containing both content and reasoning questions used in the previous phase. The last part was a subscale from the Nature of Scientific Knowledge Scale (Rubba, 1977) measuring an individual's views about the tentativeness of science (ToS). This measure was used in place of the VOST (Hillis, 1975) used in phase II due to its shorter length.

After reading the news article, participants were asked to identify the conclusions of the research, rate their trust in the conclusions, and evaluate claims based on the research they had read. The conclusions questions provided the participant with a list of possible conclusions to choose from, with each article containing two correct statements (Appendix J). The order of the options was randomized by the survey software. The list of options for conclusions for each article was reviewed by the corresponding author of the research studies used. Each researcher confirmed that the statements listed as correct were actual conclusions from their work and that the incorrect statements were not legitimate conclusions. Participant responses to this question were scored using the rubric in Table 4-1.

Table 4-1 Understanding of Conclusions Rubric

Score	Criteria
5	Chose both correct conclusions and no incorrect conclusions
4	Chose one correct conclusion and no incorrect conclusions
3	Chose both correct conclusions and at least one, but not all, incorrect conclusions
2	Chose one correct conclusion and at least one, but not all, incorrect conclusions
1	Chose only incorrect conclusions
0	Chose either all or none of the conclusions

While participants in phase II were questioned about their trust in the conclusions of the research they read, it was asked as an open-ended question. A validated measure of trust in the conclusions was needed for this phase, so the trust in conclusions questions were adapted from a believability index for newspapers (Meyer, 1988). This index was developed for the purpose of judging news articles as a whole, but should also apply to assessing parts of an article, in this case, the conclusions. In both instances, the reader is tasked with evaluating information they were presented. This measure contains four 6 – point Likert-type scales, so participant responses to each scale were averaged to generate an overall trust score. After each trust question, participants were asked to indicate the reason that they chose their response (Appendix J). Possible options, based on the reasons participants provided in phase II, were given, as well as the option to type in an alternative reason.

The evaluation of claims questions asked participants to indicate how much they agree with statements making claims about the research that they read (Appendix J). All of the statements were verified as reasonable/unreasonable by the corresponding author of the original research article that was reported in the news. Participants were given statements that drew some further conclusion from what was presented in the article. They were then asked to indicate how much they agreed with each statement on a 6 – point Likert-type scale. Three of the statements were unreasonable, so they should be rated low. These statements made claims well beyond the conclusions presented by the scientists involved in the research. The other two statements were reasonable, so they should be rated high. These statements included a rewording of a conclusion from the article and a general statement that "Scientists must now reconsider what they thought they knew about _____", where the blank was research topic specific. Ratings for the unreasonable claims were averaged to provide a measure of ability to identify unreasonable claims. A score for identifying reasonable claims was similarly generated. Finally, ratings for the unreasonable claims were reversed and averaged with the ratings for the reasonable claims to generate an overall indicator of ability to evaluate claims. After participants rated each claim, they provided reasons for their rating chosen from a pre-generated list with the option to type in their own reason. This list was created to reflect the reasons offered by participants in phase II.

.

4.2.5 Data Analysis

In order to answer the research questions, initial three-way mixed model ANCOVAs were performed, with the outcome of interest as the dependent variable, limitations and methods as fixed factors, article topic as a random factor, and science literacy score and tentativeness of science score as covariates. Then, the best fit model was selected using the backwards elimination procedure (Devore, 2008) of sequentially removing the variable contributing the least explanatory power to the model until only variables with $p < 0.05$ were left. Tukey post-hoc tests were performed for the selected models. Additionally, the frequency of reasons chosen when evaluating claims for the public and faculty separated by limitations condition were analyzed using Fisher's exact test to provide some insight into the differences observed in the ANCOVAs.

4.3 Results

In order to answer the research questions regarding participants' understanding of the conclusions, the analysis was conducted separately for the public and faculty participants. The best fit model for the public's understanding of the conclusions of the news articles included only significant main effects for article topic, $F(2,228) = 10.06$, $p < 0.0001$, and science literacy score, $F(1,229) = 6.14$, $p = 0.014$. There were no significant main effects for methods or limitations and there were no significant interaction effects. A Tukey post-hoc test for article topic (Table 4-2) showed that participants had a better understanding of the conclusions of the Comet Water article than they did of Bird Compass, with a moderate effect size ($d = 0.45$), or Space Membranes, with a large effect size ($d = 0.72$). The parameter estimate for science literacy was 0.05, indicating that the higher a participant's science literacy, the better their understanding of the article conclusions.

Table 4-2 Means for Article Topic for Public Understanding of Conclusions. Cells with different letters are significantly different

	Mean
Comet Water	2.58[a]
Bird Compass	2.06[b*]
Space Membranes	1.74[b**]

*p = 0.0157, **p < 0.0001

The lack of a significant interaction effect between article topic and science literacy score indicates that participants at all science literacy levels understood the conclusions of the Comet Water article better than the other articles. This result signifies that there was some difference in how readers understood each of the articles, but the lack of other significant effects indicates that neither the presence of limitations nor methods affected reader understanding. The fact that there are no main or interaction effects for limitations or methods shows that the public were seemingly unaffected by the presence or absence of these pieces of information in the news article. It is not completely surprising that neither including study limitations nor methods impacted participants' understanding of the conclusions, as neither specifically refers to the research conclusions. However, it is clear from the mean scores (Table 4-2) that public participants do not have a good understanding of the conclusions, as they were sometimes able to identify at least one appropriate conclusion, but also misidentified multiple inappropriate conclusions as being from the article. Therefore, it may be important to explore other methods of improving reader understanding of conclusions in science news articles.

Science faculty should provide the ideal case for understanding conclusions from science news articles, as they routinely identify conclusions of scientific research as part of their work. Therefore, it is of interest to compare their results to those of the public. The best fit model for the faculty's understanding of conclusions included only a significant main effect for article topic, $F(2,188) = 37.61$, $p < 0.0001$. There were no significant main effects for limitations, methods, or science literacy score and there were no significant interaction effects. A Tukey post-hoc test for article topic (Table 4-3) showed that the faculty participants, just like the public participants, had a better understanding of the conclusions of the Comet Water article than they did of Bird Compass, with a large effect size ($d = 1.06$), or Space Membranes, with a large effect size ($d = 1.50$).

Table 4-3 Means for Article Topic for Faculty Understanding of Conclusions. Cells with different letters are significantly different

	Mean
Comet Water	4.22^{a}
Bird Compass	3.16^{b}
Space Membranes	2.71^{c}

$p < 0.0001$ for all comparisons except Bird Compass and Space Membranes ($p = 0.038$)

As with the analysis of the data from public participants, the lack of significant effects for limitations and methods suggests that their inclusion did not affect reader understanding. The absence of an effect for science literacy score is more likely due to the science faculty all scoring very high on this measure because of their extensive science backgrounds, so it is not a meaningful variable to distinguish between any differences in understanding conclusions. As expected, the science faculty demonstrated a greater understanding of the conclusions from the news articles (Table 4-3) than the public, as they were able to identify appropriate conclusions, though sometimes also misidentified inappropriate ones as appropriate. Based on these results it is clear that there are some differences between the articles related to reader understanding, but they were not due to the manipulations made to test the inclusions of limitations or methods. However, analysis of the other outcomes of interest will demonstrate effects of limitations.

Table 4-4 Means for Limitations for Public Trust in the Conclusions. Cells with different letters are significantly different

	Mean
Omitted Limitations	3.72[a]
Included Limitations	3.51[b]

$p = 0.043$

In order to answer the research questions regarding participants' trust in the conclusions, the analysis was conducted separately for the public and faculty participants as before. The best fit model for the public's trust in the conclusions of the news articles included significant main effects for limitations, $F(1,226) = 5.90$, $p = 0.016$, and article topic, $F(2,226) = 26.37$, $p < 0.0001$, along with a marginally significant interaction effect between science literacy and limitations, $F(1,226) = 3.39$, $p = 0.067$. There were no significant main effects for science literacy or methods and there were no other significant interaction effects. A Tukey post-hoc test for limitations (Table 4-4) showed that participants trusted the conclusions less when the limitations were presented, with a small effect size ($d = 0.27$). While the interaction between science literacy and limitations was marginally significant, it does provide some additional interesting insight

(Figure 4-1). Participants with low science literacy trusted the conclusions from the news articles less when the limitations were presented, as seen in the ANCOVA analysis. However, this difference in trust disappears as the reader science literacy increases, to the point where the trend flips slightly at the highest science literacy scores. A Tukey post-hoc test for article topic (Table 4-5) showed that participants trusted the conclusions of the Space Membranes article less than Bird Compass, with a large effect size (d = 1.03), or Comet Water, with a large effect size (d = 1.00).

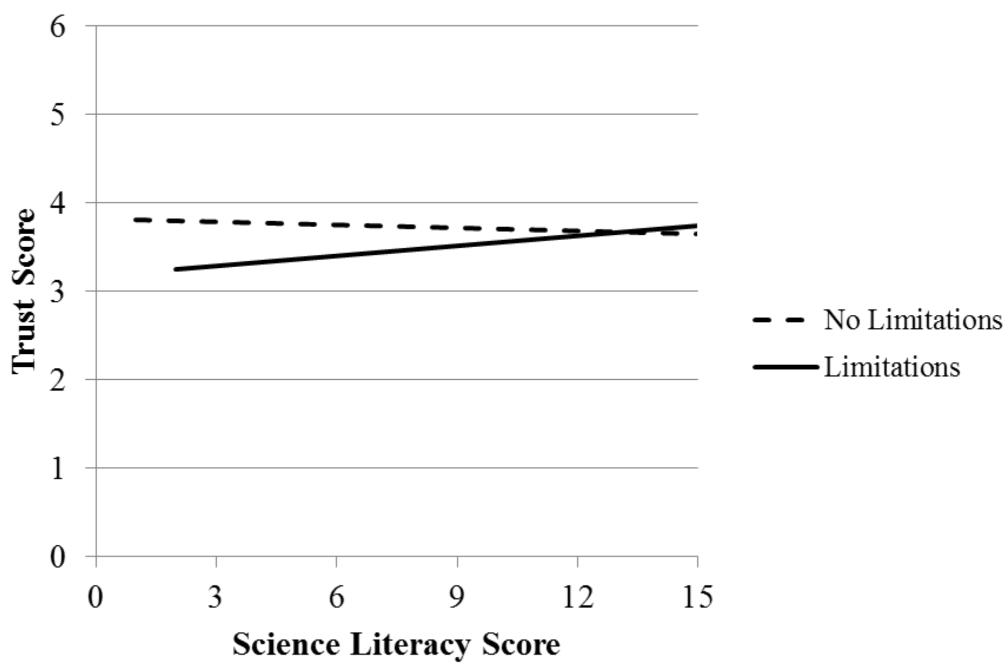


Figure 4-1 Trust by Science Literacy Score for Public Participants

Table 4-5 Means for Article Topic for Public Trust in Conclusions. Cells with different letters are significantly different

	Mean
Comet Water	3.86[a]
Bird Compass	3.89[a]
Space Membranes	3.10[b]

$p < 0.0001$

While article topic once again showed a significant main effect, the lack of a significant interaction effect between it and limitations or a significant triple interaction between it, science literacy, and limitations indicates that the effect of limitations on public participants' trust in the conclusions was stable across the difference texts. For all three articles, the addition of limitations decreased trust for participants with low science literacy and had no effect on participants with high science literacy. It appears that including limitations is viewed negatively by the public, particularly by those with low science literacy, as evidenced by decreased levels of trust. However, including methods had no effect on participants' trust in the conclusions, possibly indicating that it was not an important factor in their determination of trust. In addition, participants may have trusted the Space Membranes article less than the others because it touched on a somewhat controversial topic – the origins of life. A number of participants indicated, through open-ended responses, that they did not trust the conclusion due to it conflicting with their religious beliefs, where this was much less of an issue with the other new articles used in this study.

In contrast to the public, the best fit model for the faculty's trust in the conclusions of the news articles included only a significant main effect for article topic,

$F(2,188) = 33.52$, $p < 0.0001$. There were no significant main effects for limitations, methods, or science literacy and no significant interactions. A Tukey post-hoc test for article topic (Table 4-6) showed that, just like the public, faculty participants trusted the conclusions of the Space Membranes article less than Bird Compass, with a large effect size ($d = 1.11$), or Comet Water, with a large effect size ($d = 1.37$).

Table 4-6 Means for Article Topic for Faculty Trust in Conclusions. Cells with different letters are significantly different

	Mean
Comet Water	4.38[a]
Bird Compass	4.21[a]
Space Membranes	3.47[b]

$p < 0.0001$

Unlike the public, the faculty's trust in the conclusions was not affected by the inclusion of study limitations. It may have been expected that including limitations would increase their trust, as acknowledging limitations is a part of scientific practice, but that was not observed here. As before, the presence of a description of methods had no effect. The faculty generally trusted the conclusions more than the public (Tables 4-5 and 4-6), but also did not trust the conclusions of the Space Membranes article as much as the other articles. However, their open ended responses indicate that this is not due to religious beliefs, as it may have been for the public, but rather is due to the research conclusions being somewhat tentative.

Table 4-7 Frequency of Reasons Provided for Trust in the Conclusions. Participants could choose more than one reason

		Omitted Limitations	Included Limitations
Trust the researchers	Public	138[ab]	63[ab]
	Faculty	46[ab]	72[ab]
Don't trust the researchers	Public	22	26
	Faculty	1	1
Not enough information	Public	107	102
	Faculty	110	106
Based on the results presented	Public	168[ab]	114[ab]
	Faculty	186[b]	192[b]
The study was conducted well	Public	91[b]	70[b]
	Faculty	53[b]	68[b]
The conclusions are unclear	Public	77[a]	106[a]
	Faculty	23	20
It is unclear how the experiment was done	Public	68	52
	Faculty	58[a]	35[a]
The researchers could have made a mistake	Public	50[ab]	99[ab]
	Faculty	29[b]	26[b]
There was not enough research	Public	62[a]	93[a]
	Faculty	60	59
Other	Public	20	32
	Faculty	107	93

[a] Significant difference between omitted vs. included limitations within a group (Public or Faculty), $p < 0.05$
[b] Significant difference in distribution of omitted and included limitations between Public and Faculty, $p < 0.05$

In order to provide more insight into the differences in participants' trust in the conclusions, the reasons they indicated for their responses were analyzed. When comparing the frequency of reasons provided by the public, significant differences in the frequency based on the presence/absence of limitations for the reasons "trust the researchers", "based on the results", "the conclusions are unclear", and "there was not enough research" were found (Table 4-7). Public participants reading an article without limitations were significantly more likely to indicate that they trusted the researchers and that they based their trust in the conclusions on the results presented than those reading an article including limitations. They also were less likely to indicate that the conclusions were unclear, the researchers could have made a mistake, and that there was not enough research when the limitations were absent. It is interesting that public participants seem to trust researchers and the results less, leading to their decrease in trust in the conclusions, when study limitations were presented in the news article. This contrasts with the science faculty, who were more likely to indicate that they trusted the researchers when the limitations were presented and showed no significant difference in the likelihood of indicating that they based their trust of the conclusions on the results presented (Table 4-7). Comparisons between the public and faculty showed significant differences in the frequency of responses for trust in the researchers, based on the results, the study was conducted well, and the researchers could have made a mistake. In the first three cases, the public were more likely to choose that response if the limitations were absent, while the faculty showed the opposite trend. For the last case, the public were more likely to indicate the researchers could have made a mistake when the limitations were presented, while the faculty showed a slight opposite trend. This pattern of the public choosing more positive reasons to explain their level of trust in the conclusions when the limitations were absent, while the opposite was true of the faculty, may indicate that each group evaluates the presence of study limitations differently.

In order to answer the research questions pertaining to participants' evaluation of claims, the analysis was conducted separately for each group as before. The best fit model for the public's overall evaluation of claims included only significant main effects for limitations, $F(1,229) = 11.12$, $p = 0.001$, and ToS score, $F(1,229) = 5.94$, $p = 0.016$. There were no significant main effects for methods or article topic and there were no significant interaction effects. A Tukey post-hoc test for limitations (Table 4-8) showed that participants reading an article describing limitations were better able to evaluate claims, with a small effect size ($d = 0.31$). The parameter estimate for ToS was 0.13, indicating that the more tentative a participant viewed science, the better their ability to evaluate the claims presented.

Table 4-8 Means for Limitations for Overall Evaluation of Claims. Significance calculated between omitted vs. included limitations

	Omitted Limitations	Included Limitations
Public	3.39^{a}	3.57^{a}
Faculty	3.81^{b}	3.97^{b}

^{a} $p = 0.001$, ^{b} $p = 0.038$

The lack of a significant interaction effect between ToS and limitations indicates that presenting limitations had a positive impact on participants' ability to evaluate claims at all levels of ToS views. This result indicates that describing limitations in a news article increased the ability to evaluate claims, even for those participants with a more sophisticated view of the tentativeness of science. The fact that there is no main or interaction effect for article topic shows that this result is stable across different texts. This indicates that the results may be general and not specific to a given article. It may have been expected that individuals with more naïve views of science may benefit from describing study limitations, but these results show that it also benefits those with more sophisticated views. Therefore, including a description of study limitations in science news articles may increase a readers' ability to evaluate claims from the article. However, there was no significant main or interaction effect for methods, indicating that the brief description of methods added to the news articles had no effect on the public participants' ability to evaluate claims.

In contrast to the public, the best fit model for the faculty's overall evaluation of claims included only significant main effects for limitations, $F(1,187) = 4.35$, $p = 0.038$, and article topic, $F(2,187) = 19.29$, $p < 0.0001$. There were no significant main effects for methods or ToS and there were no significant interaction effects. A Tukey post-hoc test for limitations (Table 4-8) showed that participants reading an article describing limitations were better able to evaluate claims, with a small effect size ($d = 0.19$). A Tukey post-hoc test for article topic (Table 4-9) revealed that the faculty participants evaluated the claims made about the Space Membranes article less well than they did for the other articles.

Table 4-9 Means for Article Topic for Faculty Overall Evaluation of Claims. Cells with different letters are significantly different

	Mean
Comet Water	4.10^{a}
Bird Compass	4.03^{a}
Space Membranes	3.54^{b}

$p < 0.0001$

While the significant main effect for article topic showed that there was a difference in the evaluation of claims among the texts, the lack of a significant interaction effect between article topic and limitations indicates that the effect of interest was stable. The overall evaluation of claims was lower for the Space Membrane article, but no interaction effect between article topic and limitations indicates that the effect of including limitations was similar in all three texts. This result indicates that describing limitations in a news article increased the ability to evaluate claims, just as it did for the public, though to a lower extent. This decrease in effect size may be due to the more extensive science background of the faculty causing a ceiling effect. The absence of a significant main or interaction effect for ToS is most likely a result of the faculty having similar views, as the vast majority of them scored very highly on that scale. It was unexpected that the faculty benefitted from the presence of study limitations in the news articles almost as much as the public did, as their greater familiarity with evaluating claims made by other scientists was expected to compensate for the absence of study limitations. However, most of the faculty participants were presented with articles outside of their field of expertise, so their familiarity may not have completely transferred to other science contexts. Finally, similar to the public, there was no significant main or interaction effect for methods, indicating that the brief description of methods added to the news articles had no effect on either group's ability to evaluate claims.

The previously discussed results focused on participants' overall evaluation of claims, however examining their evaluation of unreasonable and reasonable claims separately also yielded interesting results. The best fit model for the public's evaluation of unreasonable claims included only a significant main effect for limitations, $F(1,230) = 10.75$, $p = 0.001$. There were no significant main effects for ToS, methods, or article topic and no significant interaction effects. A Tukey post-hoc test for limitations (Table

4-10) showed that participants reading an article describing limitations were better able to evaluate unreasonable claims, with a small effect size (d = 0.30), as evidenced by the decrease in average agreement.

Table 4-10 Means for Limitations for Evaluation of Unreasonable Claims. Significance calculated between omitted vs. included limitations. A decrease in value indicates more disagreement with the statement and, therefore, a more accurate response.

	Omitted Limitations	Included Limitations
Public	4.23^{a}	3.89^{a}
Faculty	3.69^{b}	3.43^{b}

^{a} $p = 0.001$, ^{b} $p = 0.05$

Once again, the main and interaction effects for article topic were not significant, indicating that the results were stable across the different texts. In contrast to overall ability to evaluate claims, participants' views of ToS had no effect on evaluating unreasonable claims. Only the presence of limitations predicted ability to identify unreasonable claims. This result indicates that participants' views of the tentativeness of science did not help or hinder their ability to evaluate unreasonable claims, suggesting that including a description of limitations is more important in aiding an individuals' ability to identify unreasonable claims than views of the tentativeness of science. As with the overall evaluation of claims, inclusion of research methods had no effect on the public's ability to evaluate unreasonable claims.

As a comparison, the best-fit model for the faculty's evaluation of unreasonable claims included only significant main effects for limitations, $F(1,187) = 3.80$, $p = 0.05$, and article topic, $F(2,187) = 13.47$, $p < 0.0001$. There were no significant main effects for ToS or methods and no significant interaction effects. A Tukey post-hoc test for limitations (Table 4-10) showed that participants reading an article describing limitations were better able to evaluate unreasonable claims, with a small effect size (d = 0.18), as evidenced by the decrease in average agreement. A Tukey post-hoc test for article topic (Table 4-11) once again revealed that faculty participants evaluated the unreasonable claims about the Space Membranes article less well than those of the other articles.

Table 4-11 Means for Article Topic for Faculty Evaluation of Unreasonable Claims. Cells with different letters are significantly different. A decrease in value indicates more disagreement with the statement and, therefore, a more accurate response.

	Mean
Comet Water	3.29[a]
Bird Compass	3.35[a]
Space Membranes	4.03[b]

$p < 0.0001$

As with the results for overall evaluation of claims, the lack of a significant interaction between article topic and limitations indicated that the effect of including limitations was stable across the different texts. Similar to the public's evaluation of unreasonable claims, participants' views of ToS and inclusion of methods had no effect, while the presence of limitations predicted ability to identify unreasonable claims. These results indicate that including limitations in the news article aided both the public and the faculty in evaluating unreasonable claims.

The best fit model for the public's evaluation of reasonable claims included just a significant main effect for ToS, $F(1,230) = 9.33$, $p = 0.003$. There were no significant main effects for limitations (Table 4-12), methods, or article topic, and no significant interaction effects. The parameter estimate for ToS was 0.33, indicating that the more tentative a participant viewed science, the better they were at identifying reasonable claims. The main and interaction effects for article topic were not significant, indicating that the results were stable across the different texts. In addition, the main and interaction effects of methods were also not significant. In contrast to overall ability to evaluate claims and ability to evaluate unreasonable claims, the presence of limitations had no effect on evaluating reasonable claims. Only participant views of ToS predicted ability to identify reasonable claims. This result indicates that presenting limitations had no effect on ability to evaluate reasonable claims, suggesting that participants are able to identify reasonable claims using their views of the tentativeness of science and limitations are not needed.

Table 4-12 Means for Limitations for Evaluation of Reasonable Claims. No significance differences were found between omitted vs. included limitations

	Omitted Limitations	Included Limitations
Public	4.29	4.24
Faculty	4.55	4.59

Unlike the results for the public's evaluation of reasonable claims, there was no best fit model for the faculty's evaluation of reasonable claims. None of the variables measured in this study predicted their ability to evaluate reasonable claims. This result suggests that the science faculty were generally adept at evaluating reasonable claims, as their average scores were fairly high (Table 4-12), and that the manipulations to the news articles did not contribute to their evaluations. The faculty are likely to be more familiar with the scientific process than the general public, which may have contributed to their high ability to evaluate reasonable claims.

Table 4-13 Frequency of Reasons Provided for Evaluation of Unreasonable Claims. Participants could choose more than one reason

		Omitted Limitations	Included Limitations
Trust the experts	Public	80[a,b]	49[a,b]
	Faculty	12[b]	20[b]
Based on information in the article	Public	166[b]	146[b]
	Faculty	138[a,b]	174[a,b]
Results may not be representative	Public	8[a,b]	20[a,b]
	Faculty	36[b]	28[b]
Results are representative	Public	62	51
	Faculty	21[a]	9[a]
Not enough information	Public	40	50
	Faculty	66	66
Results are insufficient	Public	22	33
	Faculty	39	47
Results are sufficient	Public	47	40
	Faculty	11	13
Results are unclear	Public	27[a]	53[a]
	Faculty	11	13
Other	Public	13	17
	Faculty	58	63

[a] Significant difference between omitted vs. included limitations within a group (Public or Faculty), $p < 0.05$
[b] Significant difference in distribution of omitted and included limitations between Public and Faculty, $p < 0.05$

In order to provide more insight into participants' evaluation of claims, the reasons indicated for their responses were analyzed. For the evaluation of unreasonable claims (Table 4-13), the public showed significant differences in the frequency of reasons chosen based on the presence/absence of limitations for the reasons “trust the experts”, “the results may not be representative”, and “the results are unclear”. Public participants reading an article without limitations were more likely to indicate that they based their evaluation of the claims on their trust in the experts than those reading an article including limitations. They also were less likely to indicate that the results may not be representative and that the results were unclear when the limitations were absent. It is interesting that while the public evaluated the unreasonable claims better when limitations were included in the news article, they were less likely to indicate that they trusted the experts and more likely to indicate the results may not be representative and were unclear. This may indicate that including limitations affected the public’s trust in the researchers and their confidence in identifying the results. It is likely that knowledge of the results of the research reported in the news article is necessary to properly evaluate the claims, so it is interesting that public participants were able to better evaluate claims with limitations present while simultaneously feeling that the results were unclear. However, there were a few public participants who indicated that the results may not be representative when the limitations were present, which would be a valid reason for disagreeing with some of the unreasonable claims. It appears that these participants used the limitations to correctly identify a flaw in some of the unreasonable claims, which may partially explain why including limitations led to better evaluation of unreasonable claims.

For the faculty’s evaluation of unreasonable claims (Table 4-13), there were significant differences in the frequency of reasons chosen based on the presence/absence of limitations for the reasons “based on information in the article” and “the results are representative”. They were more likely to indicate that they based their evaluation on the information in the article and less likely to indicate the results were representative when the limitations were present. This result may imply that the faculty felt more comfortable relying on the information in the news article when the limitations were present. It is also interesting that the public used the presence of limitations as an indicator that the results may not be representative, while the faculty used the absence of limitations as an indicator that the results were representative. In addition to the differences within participant groups, the public and faculty were significantly different in their frequency of indicating they used the reasons of trust the experts, based on the information in the article, and the results may not be representative. Overall, the public used their trust of the experts as a reason much more often than the faculty. However, they were less likely to indicate that they used their trust in the experts to evaluate the unreasonable claims when limitations were present, while the faculty had a slight trend in the opposite direction. A similar trend was observed for the reason based on the information in the article. The opposite trend was observed for using the reason that the results may not be representative, possibly indicating a difference in how the faculty and public view limitations.

For the evaluation of reasonable claims (Table 4-14), the public again showed significant differences in the frequency of reasons chosen based on the presence/absence of limitations for the reason “trust the experts”. They more frequently indicated that they used their trust in the experts as a reason when evaluating reasonable claims when the limitations were absent, just as when evaluating unreasonable claims. In contrast, the faculty showed no significant differences based on the presence/absence of limitations. However, when comparing between participant groups, there once again was a significant difference in the frequency of use of the reason of trusting the experts. Overall, the public used their trust of the experts as a reason much more often than the faculty, while being less likely to indicate that they used their trust in the experts to evaluate the reasonable claims when limitations were present and the faculty having slight trend in the opposite direction.

Table 4-14 Frequency of Reasons Provided for Evaluation of Reasonable Claims. Participants could choose more than one reason

		Omitted Limitations	Included Limitations
Trust the experts	Public	56[a,b]	34[a,b]
	Faculty	11[b]	17[b]
Based on information in the article	Public	116	116
	Faculty	131	144
Results may not be representative	Public	3	8
	Faculty	3	9
Results are representative	Public	35	29
	Faculty	16	6
Not enough information	Public	24	24
	Faculty	17	16
Results are insufficient	Public	7	12
	Faculty	4	7
Results are sufficient	Public	30	30
	Faculty	28	25
Results are unclear	Public	29	36
	Faculty	4	5
Other	Public	8	10
	Faculty	39	39

[a] Significant difference between omitted vs. included limitations within a group (Public or Faculty), $p < 0.05$

[b] Significant difference in distribution of omitted and included limitations between Public and Faculty, $p < 0.05$

4.4 Discussion

This phase provided evidence that the results from phase II are generalizable to a larger American adult population. Both the public and science faculty were unaffected by the presence of limitations or methods when demonstrating their understanding of the conclusions of the research in the science news articles presented, though there were differences in levels of understanding between the different articles. It is encouraging that including study limitations did not decrease reader understanding of the conclusions, since their inclusion may be beneficial in evaluating unreasonable claims.

The public trusted the conclusions of the research more when the study limitations were absent, with this effect being more pronounced for participants with lower science literacy and the difference essentially disappearing for participants with high science literacy. This contrasts with the faculty, who showed no difference in levels of trust based on the presence/absence of study limitations. Once again, introduction of a discussion of methods had no effect on either group, but there was a difference between the different articles. It is interesting that the public trusted the conclusions less when the limitations were presented. The reasons provided by the participants imply that this difference may be due to a decrease in their trust in the researchers and/or results, though this should be explored further.

Finally, describing limitations in science news articles may impact readers' ability to evaluate claims based on the research. Only including limitations had a significant effect when evaluating unreasonable claims, while only ToS views had a significant effect when evaluating reasonable claims. These results imply that while a more sophisticated view of ToS may be needed to identify reasonable claims, a description of limitations may be necessary to identify unreasonable claims. While it may have been expected that including limitations would aid the public in evaluating claims, it was unexpected that it also aided the science faculty. The more sophisticated views of ToS that the faculty had decreased the effect that including the limitations had, but did not eliminate it.

While both the public and faculty showed similar results from the ANCOVA analysis, there were some differences in the reasons provided to evaluate the claims. The public tended to provide reasons that imply that they viewed the inclusion of limitations more negatively than the faculty. This possible difference in the interpretation of the limitations should be investigated further to determine any effects on reader perceptions of the news articles. Including limitations could affect readers' trust in the researchers and/or results, beliefs about the quality of the research, or willingness to apply the results in their daily lives. Correctly evaluating claims indicates an understanding of the results reported in the news article, but does not provide information on what the reader might think about the research.

Science news articles are inconsistent with their presentation of research limitations, which this phase suggests may have an impact on how readers' trust in the conclusions and their evaluation of claims about the research. Omitting study limitations may inhibit a readers' ability to identify unreasonable claims about the research, though may increase their trust in the conclusions. This suggests that there is a trade-off between empowering a person with the ability to evaluate claims about the research and having them trust the research conclusions presented. Therefore, careful thought may be required when deciding whether or not to include a description of research limitations in science news articles. Research is needed to identify any other factors that contribute to the ability to evaluate claims and that influence a person's trust in the conclusions. In addition, the articles chosen for this study were politically and emotionally neutral, so it is not clear if these results would also apply to other science topics.

CHAPTER 5. CONCLUSIONS

5.1 Summary and Discussion of the Results

As stated in Chapter 1, this dissertation aimed to address the following two broad research questions:

1. How is chemistry research reported upon in print news?

2. How do aspects of news reports of chemistry research impact reader's understanding and perceptions of the research?

The first phase of my study focused on the first research question, specifically studying the general structure of news reports reporting on chemistry research and comparing that structure to press releases and research reports describing the same material. It was found that the overall structure of each type of text was fairly similar, with differences occurring near the beginning of the texts. Research reports tended to start by presenting background information, while news reports and press releases first highlighted overall research outcomes. These results are consistent with previous work studying medical research articles and online popularizations (Csongor, 2013), but display differences from other studies of popularized science (Nwogu, 1991; Stejskalová, 2010). Nwogu and Stejskalová both found that popular science articles follow a more similar structure to research reports by first presenting background information.

However, their research focused on science popularizations in publications, such as *New Scientist*, *National Geographic*, and *Science Daily*, whose readers are likely to have a more extensive science background than the average adult. The differences between the results reported here and their work may be due to the different types of science news outlets that were studied and it may be of interest to further explore the landscape of science popularizations.

While the general structure of the three types of texts studied was similar, some key differences were evident when examining the steps within each rhetorical move. Research reports were more likely to discuss the methods involved in generating the results of the research and the scientific limitations of previous and/or current work. News reports rarely mentioned methods and focused on limitations pertaining to practical considerations in applying the results to people's lives. These results are consistent with previous studies focusing on health or environmental science news reporting (Brechman et al., 2009, 2011; Kua et al., 2004; Nelkin, 1995; Pellechia, 1997; Stocking, 1999; Tankard & Ryan, 1974; Woloshin & Schwartz, 2002, 2006) in that limitations are often excluded from science news. However, this dissertation has demonstrated that this trend holds true for news about chemistry research. In addition, the previous research mentioned above has focused on the claims made in news reports and/or press releases compared to research reports, while this work has demonstrated differences in how methods are treated in each of the text genres analyzed as well. It is relatively easy to imagine how alterations to the claims made about some scientific research may affect readers' understanding, but it is less clear what impact, if any, including methods would have.

The second and third phases of this dissertation addressed the second main research question, specifically focusing on how the inclusion of scientific limitations and/or methods affected readers' understanding and perception of the research. The second phase served as a pilot study and found that non-academic university staff showed similar abilities to summarize a chemistry news article and understand the purpose, results, and significance of the research either with or without limitations presented. However, it appeared that the staff participants were better able to comprehend the conclusions of the article when limitations were presented. Science faculty showed no differences based on presence/absence of limitations in understanding any of the aspects of the article. This result is not surprising, as the faculty's more extensive science backgrounds and experience reading reports about scientific research presumably aided in their ability to understand the content. It is interesting that including limitations seemed to increase the staff participants' understanding of the conclusions, though this difference may be due to participants reading the article with limitations being more scientifically literate.

In addition to slight differences in comprehending the article, staff participants trusted the conclusions of the article more with limitations present, but there was no difference in their trust of the findings. While most of the staff indicated that they trusted the findings, regardless of article version, including limitations caused them to be unsure how much to trust the conclusions. The science faculty generally were unsure of how much the trust the findings and conclusions, regardless of whether limitations were presented. These results are contrary to previous work examining reader's trust in the researchers and/or journalists when limitations were included in cancer news (Jensen, 2008; Jensen et al., 2011). It is possible that this difference is due to adults having different criteria when making trust judgments about information versus judgments about the scientists or journalists.

The staff participants' explanations of their level of trust suggest that they trusted the findings because they deferred to the experts. Although some participants continued to defer to the experts in trusting the conclusions, some of those presented with limitations felt the conclusions were unclear and therefore were unsure of how much to trust them. These results imply that the staff participants felt less sure of the conclusions when limitations were included, even though they were better able to describe the conclusions. It may be of interest to explore a possible connection between a reader's confidence in their ability to comprehend the information in a science news article and the presence of limitations, as well as investigating what criteria adults use to judge how much to trust science news.

Additionally, the results of phase II indicate that the staff were better able to evaluate claims made about the research when limitations were included in the news article, while the science faculty showed no clear differences. As discussed in Chapter 1, a scientifically literate person should be able to evaluate claims made about science research (National Research Council, 1996) and it is encouraging that including limitations seems to increase readers' ability to do that. The science faculty may have been much less affected because their more extensive science background and presumed greater experience reading about scientific research may have allowed them to intuit potential limitations, even if they were not explicitly mentioned.

The final phase of this dissertation expanded upon the work done in phase II to a larger, general public population and included an investigation of the impact of describing research methods in the news articles. Unlike in phase II, the presence of limitations and/or methods had no impact on the public's understanding of the conclusions, supporting the notion that the differences previously seen were due to differences in scientific literacy rather than to the manipulation of the articles. As before, science faculty showed no difference in their ability to understand the conclusions. It is encouraging that including limitations and/or methods did not negatively impact reader understanding, as they may provide benefits in other areas.

As was observed in phase II, the public trusted the conclusions of the news articles less when limitations were present and the science faculty showed no differences. However, it is evident that this decrease in trust is more pronounced for adults with low scientific literacy and non-existent for those with very high scientific literacy. This implies that adults with low scientific literacy perceive limitations differently those with high scientific literacy and that there is a continuum. Unlike in the previous phase, participant reasons for their lower levels of trust suggests that they distrusted the researchers and/or results more when limitations were present, rather than being unsure of the conclusions. The inclusion of methods once again had no effect.

Finally, similar to the pilot, the public were better able to evaluate claims when limitations were present. However, somewhat surprisingly, this trend also was observed for the science faculty, though to a smaller extent. More specifically, all participants were better able to evaluate unreasonable claims after reading a science news article containing limitations. It appears that including limitations enables readers to identify unreasonable claims more readily, though it is not clear exactly how. The science faculty may have also benefitted from the inclusion of limitations because the content of the news articles was often far outside of their area of expertise, which may have caused them to be more similar to members of the public with very high scientific literacy. The presence of limitations did not affect readers' ability to evaluate reasonable claims, which was governed by their views of the tentativeness of science. Those who view science as more tentative were better at evaluating reasonable claims. This may be due to these participants having a more sophisticated view of the nature of science and, therefore, a better understanding of what a reasonable scientific claim might be. As before, including methods had no effect on evaluating claims.

This dissertation has shown that print news about chemistry research, like that of health and environmental science, tend to omit a description of research methods and scientific limitations, compared to the original research article. It has also demonstrated that the omission of methods has no effect on reader understanding of research conclusions, trust in the conclusions, or evaluation of claims. However, the lack of a discussion of the limitations appears to increase reader trust in the conclusions, but decrease their ability to evaluate unreasonable claims, with a greater effect on trust for adults with low scientific literacy.

5.2 Limitations

While the work presented here has discussed some interesting findings, there are several limitations. The evaluation of the structure of chemistry research texts included only a moderately sized sample of articles. This was mainly due to the specific criteria used to include texts for analysis and the fact that chemistry is not reported on as much as other scientific fields in the news. It is possible that including more texts for analysis may change the general move structure, though it is unlikely to substantially change the results presented above. In addition, the move analysis performed provides information on the general structure of each type of text, along with differences on how each text accomplishes the goals of each move. A linguistic analysis of the language used in each of the texts may provide a more detailed description of the differences between the texts.

Although the results of phase III are somewhat generalizable because a sample representative of the general public was obtained, a larger number of participants would further strengthen the conclusions drawn here. More participants would have allowed for the exploration of additional demographic factors that may also impact reader understanding of science news. It would also have allowed for the use of more than three difference news articles, which would have again strengthened the generalizability of the results to all science news articles. The fact that the Space Membranes article differed from the other two may indicate that some articles might not follow the trends observed in this work. However, these results do not show any evidence that the content of the article mattered when assessing the effect of including limitations and methods on readers.

In addition, the sample of public participants may not be entirely representative of the general population because individuals had to opt in to be contacted for this study, though it is likely fairly close. Therefore, certain segments of the population may have been excluded from the recruitment sample and it is possible that their results may have been different.

5.3 Implications

This dissertation has demonstrated that the omission of limitations from news about chemistry research has an impact on the readers' ability to evaluate claims and their level of trust in the conclusions. Previous research has also shown that including limitations affects reader perceptions and trust in the researchers and/or journalists (Jensen, 2008; Jensen et al., 2011), indicating that it may be worthwhile to include them in science news articles. It is interesting that adding language about study limitations decreases trust in the conclusions, but increases ability to evaluate claims, as demonstrated in this work, and trust in the source, as shown by Jensen. This implies that including limitations causes readers to be more skeptical of the research, but not of the researchers, which is arguably a positive outcome. An aspect of being scientifically literate (National Research Council, 1996) is evaluating the quality of scientific research and being more skeptical may be an indication that the readers are being more critical in their evaluations. Ideally, this would allow them to be better able to distinguish between "bad science" and "good science" (Goldacre, 2010) when making decisions in their lives.

There is also no reason why the effects of including limitations observed here should not be applicable to other modes of communication about scientific research. Presenting limitations in other forms of news reporting, such as television, radio, or online publications, should have a similar impact on their audience. Given that including limitations did not greatly increase the length of the news articles used in this dissertation, it seems worthwhile to present them when possible. This is not to imply that science journalists are at fault for omitting limitations, as researchers may also need to be explicit about the limitations of their work and the importance of including it when presenting their work. It may be valuable to explore ways of ensuring that people are provided with study limitations to help inform their decision making.

5.4 Future Directions

The results of this dissertation present opportunities for future work, including exploring reporting on scientific research with a more direct societal impact, other modes of science communication, and how readers perceive scientific limitations. The most common science news reporting is on health or environmental research and findings that could inform political or societal decisions (National Science Board, 2014; Suleski & Ibaraki, 2009). The chemistry news articles chosen for this study specifically avoided these types of research because personal values are more likely to play a role in the reader's interpretation of the science. This effect was partially seen for the Space Membrane article, as some participants cited religious beliefs as their reason for their level of trust in the conclusions and evaluation of claims. While limitations clearly impacted readers in this study, it is not clear how much that effect would transfer to science news in which people may have a more personal stake. It could be that personal beliefs overwhelm the effect of including limitations, with people predisposed to be dismissive of the results using limitations as a sign that the research is of low quality and people predisposed to accept the research either ignoring limitations or using them as a sign that the work was high quality. Knowing this would have an important impact on determining the best ways to communicate science that are likely to connect to a person's values.

This work focused solely on chemistry news articles that were published in general newspapers, either in print or online, and used a sample representative of the general population. However, science is communicated through a number of other avenues, including television, radio, magazines, and online, and in outlets that are more targeted (e.g. Scientific American, Discovery Magazine, New Scientist, etc.), whose audiences are not always representative of the general population. Because there was some indication that a reader's scientific background has an impact on how they are affected by including limitations, it is likely that there may be differences in the effect of including limitations depending on the particular mode of communication. The impact of presenting limitations may be somewhat different in a television newscast compared to a radio broadcast or a story in Scientific American. Being able to determine any differences could help disseminators of science research determine best practices for their intended audience.

Along with focusing on different outlets, it may be important to know how different people regard scientific limitations. People with low scientific literacy trusted the conclusions of research less when presented with limitations and this effect steadily disappeared as scientific literacy increased. However, it is not completely clear why this is the case. This finding implies that there is some difference in how adults interpret limitations based on their scientific literacy. Exploring what these differences are could not only help communicators decide how to best present information, it could also inform educational practices. All of the adults who participated in this study went to school and presumably learned about science at some level where they could have learned about the tentativeness of science. Determining how people view limitations differently and exploring the source of the differences could give educators an indication of what areas of science education may need more emphasis in primary and secondary school. It may also provide insight into how to best continue a person's science education in informal settings throughout their lifetime, as it is possible that prolonged exposure to discussion of scientific limitations in the popular press could have a positive impact on a person's views of nature of science.

REFERENCES

REFERENCES

Abd-El-Khalick, F., Waters, M., & Le, A.-P. (2008). Representations of nature of science in high school chemistry textbooks over the past four decades. *Journal of Research in Science Teaching*, *45*(7), 835–855. doi:10.1002/tea.20226

Abraham, M. R., Grzybowski, E. B., Renner, J. W., & Marek, E. A. (1992). Understandings and misunderstandings of eighth graders of five chemistry concepts found in textbooks. *Journal of Research in Science Teaching*, *29*(2), 105–120. doi:10.1002/tea.3660290203

Alberts, B. (2009). Redefining Science Education. *Science*, *323*, 437–438.

Alliance for Audited Media. (2013). *Total Circulation for US Newspapers*. Retrieved from http://abcas3.auditedmedia.com/ecirc/newstitlesearchus.asp

American Association for the Advancement of Science. (1989). *Science for All Americans: Project 2061*. New York: Oxford University Press.

American Association for the Advancement of Science. (1993). *Benchmarks for science literacy*. New York: Oxford University Press.

Ayers, G. (2008). The evolutionary nature of genre: An investigation of the short texts accompanying research articles in the scientific journal Nature. *English for Specific Purposes*, *27*(1), 22–41. doi:10.1016/j.esp.2007.06.002

Bell, P., Lewenstein, B., Shouse, A. W., & Feder, M. A. (2009). *Learning Science in Informal Environments: People, Places, and Pursuits*. Washington, D.C.: Committee on Learning Science in Informal. Environments, National Research Council. Retrieved from http://www.nap.edu/catalog.php?record_id=12190

Brechman, J. M., Lee, C., & Cappella, J. N. (2009). Lost in Translation? *Science Communication*, *30*(4), 453 –474.

Brechman, J. M., Lee, C., & Cappella, J. N. (2011). Distorting Genetic Research About Cancer: From Bench Science to Press Release to Published News. *Journal of Communication*, *61*(3), 496–513. doi:10.1111/j.1460-2466.2011.01550.x

Britt, M. A., Richter, T., & Rouet, J.-F. (2014). Scientific Literacy: The Role of Goal-Directed Reading and Evaluation in Understanding Scientific Information. *Educational Psychologist*, *49*(2), 104–122. doi:10.1080/00461520.2014.916217

Brody, M., Bangert, A., & Dillon, J. (2007). Assessing Learning in Informal Science Contexts. *Commissioned Paper). Washington, DC: National Research Council. Accessed December*, *5*, 2011.

Brossard, D. (2006). Do They Know What They Read? Building a Scientific Literacy Measurement Instrument Based on Science Media Coverage. *Science Communication*, *28*(1), 47–63. doi:10.1177/1075547006291345

Budescu, D. V., Broomell, S., & Por, H.-H. (2009). Improving Communication of Uncertainty in the Reports of the Intergovernmental Panel on Climate Change. *Psychological Science*, *20*(3), 299–308. doi:10.1111/j.1467-9280.2009.02284.x

Chiappetta, E. L., & Fillman, D. A. (2007). Analysis of Five High School Biology Textbooks Used in the United States for Inclusion of the Nature of Science. *International Journal of Science Education*, *29*(15), 1847–1868. doi:10.1080/09500690601159407

Committee on Science, Engineering, and Public Policy. (2007). *Rising Above the Gathering Storm: Energizing and Employing America for a Brighter Economic Future*. Washington, D.C.: National Academy Press.

Corbett, J. B., & Durfee, J. L. (2004). Testing Public (Un)Certainty of Science: Media Representations of Global Warming. *Science Communication*, *26*(2), 129–151.

Csongor, A. (2013). *Rhetorical Moves and Hedging in Medical Research Articles and their Online Popularizations*. University of Pécs, Pécs, Hungary.

Deng, F., Chen, D.-T., Tsai, C.-C., & Chai, C. S. (2011). Students' Views of the Nature of Science: A Critical Review of Research. *Science Education*, *95*(6), 961–999.

Devore, J. L. (2008). *Probability and Statistics for Engineering and the Sciences* (7th ed.). Belmont, CA: Duxbury Press/Thomson Brooks/Cole.

Dhingra, K. (1999). *An ethnographic study of the construction of science on television* (Ed.D.). Retrieved from http://search.proquest.com.ezproxy.lib.purdue.edu/dissertations/docview/304517365/abstract/139971AE84D584E45EA/1?accountid=13360

Dhingra, K. (2003). Thinking about television science: How students understand the nature of science from different program genres. *Journal of Research in Science Teaching*, *40*(2), 234–256. doi:10.1002/tea.10074

Dhingra, K. (2006). Science on Television: Storytelling, Learning and Citizenship. *Studies in Science Education*, *42*(1), 89–123.

Dierking, L. D., & Falk, J. H. (1994). Family behavior and learning in informal science settings: A review of the research. *Science Education*, *78*(1), 57–72. doi:10.1002/sce.3730780104

Falk, J. H. (1983). Time and behavior as predictors of learning. *Science Education*, *67*(2), 267–276. doi:10.1002/sce.3730670214

Falk, J. H. (2001). *Free-Choice Science Education: How We Learn Science outside of School. Ways of Knowing in Science and Mathematics Series.* Williston, VT: Teachers College Press.

Falk, J. H., & Adelman, L. M. (2003). Investigating the impact of prior knowledge and interest on aquarium visitor learning. *Journal of Research in Science Teaching*, *40*(2), 163–176. doi:10.1002/tea.10070

Falk, J. H., & Dierking, L. D. (2000). *Learning from museums: visitor experiences and the making of meaning*. Lanham, MD: Rowman & Littlefield.

Falk, J. H., & Dierking, L. D. (2010). The 95 Percent Solution. *American Scientist*, *98*(6), 486.

Falk, J. H., & Needham, M. D. (2011). Measuring the impact of a science center on its community. *Journal of Research in Science Teaching*, *48*(1), 1–12. doi:10.1002/tea.20394

Falk, J. H., & Needham, M. D. (2013). Factors Contributing to Adult Knowledge of Science and Technology. *Journal of Research in Science Teaching*, *50*(4), 431–452. doi:10.1002/tea.21080

Falk, J. H., & Storksdieck, M. (2010). Science Learning in a Leisure Setting. *Journal of Research in Science Teaching*, *47*(2), 194–212. doi:10.1002/tea.20319

Falk, J. H., Storksdieck, M., & Dierking, L. D. (2007). Investigating public science interest and understanding: evidence for the importance of free-choice learning. *Public Understanding of Science*, *16*(4), 455 –469.

Goldacre, B. (2010). *Bad Science: Quacks, Hacks, and Big Pharma Flacks* (Reprint edition.). New York: Faber & Faber.

Hillis, S. R. (1975). The development of an instrument to determine student views of the tentativeness of science. *Research and Curriculum Development in Science Education: Science Teacher Behavior and Student Affective and Cognitive Learning*, *3*, 32–39.

Hopkins, A., & Dudley-Evans, T. (1988). A genre-based investigation of the discussion sections in articles and dissertations. *English for Specific Purposes*, *7*(2), 113–121.

Horrigan, J. (2006). *The Internet as a resource for news and information about science*. Pew Internet and American Life Project. Retrieved from http://www.pewinternet.org/~/media//Files/Reports/2006/PIP_Exploratorium_ Science.pdf

Jensen, J. D. (2008). Scientific Uncertainty in News Coverage of Cancer Research: Effects of Hedging on Scientists' and Journalists' Credibility. *Human Communication Research*, *34*(3), 347–369.

Jensen, J. D., Carcioppolo, N., King, A. J., Bernat, J. K., Davis, L., Yale, R., & Smith, J. (2011). Including Limitations in News Coverage of Cancer Research: Effects of News Hedging on Fatalism, Medical Skepticism, Patient Trust, and Backlash. *Journal of Health Communication*, *16*(5), 486–503.

Kanoksilapatham, B. (2003). *A corpus-based investigation of scientific research articles: Linking move analysis with multidimensional analysis* (Ph.D.). Georgetown University, Washington, D.C.

Kanoksilapatham, B. (2005). Rhetorical structure of biochemistry research articles. *English for Specific Purposes*, *24*(3), 269–292.

Korpan, C. A., Bisanz, G. L., Bisanz, J., & Henderson, J. M. (1997). Assessing literacy in science: Evaluation of scientific news briefs. *Science Education*, *81*(5), 515–532. doi:10.1002/(SICI)1098-237X(199709)81:5<515::AID-SCE2>3.0.CO;2-D

Kua, E., Reder, M., & Grossel, M. (2004). Science in the News: A Study of Reporting Genomics. *Public Understanding of Science*, *13*(3), 309–322.

Lederman, L. M., & Malcom, S. M. (2009). The Next Campaign. *Science*, *323*, 1265–1266.

Lederman, N. G. (2007). Nature of Science: Past, Present, and Future. In S. K. Abell & N. G. Lederman (Eds.), *Handbook of Research on Science Education* (pp. 831–880). Mahwah, N.J.: Lawrence Erlbaum Associates.

Lederman, N. G., Abd-El-Khalick, F., Bell, R. L., & Schwartz, R. S. (2002). Views of nature of science questionnaire: Toward valid and meaningful assessment of learners' conceptions of nature of science. *Journal of Research in Science Teaching*, *39*(6), 497–521.

Maier, M., Rothmund, T., Retzbach, A., Otto, L., & Besley, J. C. (2014). Informal Learning Through Science Media Usage. *Educational Psychologist*, *49*(2), 86–103. doi:10.1080/00461520.2014.916215

Maxim, L., & Mansier, P. (2014). How is Scientific Credibility Affected by Communicating Uncertainty? The Case of Endocrine Disrupter Effects on Male Fertility. *Human and Ecological Risk Assessment*, *20*(1), 201–223. doi:10.1080/10807039.2012.719387

Meyer, P. (1988). Defining and Measuring Credibility of Newspapers: Developing an Index. *Journalism & Mass Communication Quarterly*, *65*(3), 567–574. doi:10.1177/107769908806500301

Miller, J. D. (1986). Reaching the attentive and interested publics for science. In S. Friedman, S. Dunwoody, & C. Rogers (Eds.), *Scientists and journalists: Reporting science as news* (pp. 55–69). New York: Free Press.

Miller, J. D. (1998). The measurement of civic scientific literacy. *Public Understanding of Science*, *7*(3), 203–223.

Miller, J. D. (2004). Public Understanding of, and Attitudes toward, Scientific Research: What We Know and What We Need to Know. *Public Understanding of Science*, *13*(3), 273–294.

Miller, J. D., Augenbraun, E., Schulhof, J., & Kimmel, L. G. (2006). Adult Science Learning from Local Television Newscasts. *Science Communication*, *28*(2), 216–242.

National Academies of Science. (2011). *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas*. Washington, D.C.: National Research Council. Retrieved from http://www.nap.edu/catalog.php?record_id=13165

National Research Council. (1996). *National Science Education Standards*. Washington, D.C.: National Academies Press.

National Science Board. (2010). *Science and Engineering Indicators 2010*. Arlington, VA: National Science Foundation.

National Science Board. (2014). *Science and Engineering Indicators 2014*. Arlington, VA: National Science Foundation.

Nelkin, D. (1995). *Selling science: How the press covers science and technology*. New York: W.H. Freeman.

Nwogu, K. N. (1991). Structure of science popularizations: A genre-analysis approach to the schema of popularized medical texts. *English for Specific Purposes*, *10*(2), 111–123.

Pellechia, M. (1997). Trends in science coverage: A content analysis of three US newspapers. *Public Understanding of Science*, *6*(1), 49–68. doi:10.1088/0963-6625/6/1/004

Rennie, L. J. (1995). Learning in science centres: What do we know and what do we need to know?'. In *Proceedings of the 20th Annual Conference of the Western Australian Science Education Association* (pp. 69–74).

Rennie, L. J., & McClafferty, T. (1995). Using visits to interactive science and technology centers, museums, aquaria, and zoos to promote learning in science. *Journal of Science Teacher Education*, *6*(4), 175–185. doi:10.1007/BF02614639

Rennie, L. J., & Williams, G. (2000). The nature and measurement of learning from attending a public lecture on human genetics. *Critical Studies in Education*, *41*, 17–34.

Rennie, L. J., & Williams, G. F. (2006). Adults' Learning about Science in Free-Choice Settings. *International Journal of Science Education*, *28*(8), 871–893.

Robinson, M. S., Stoller, F., Constanza-Robinson, M., & Jones, J. K. (2008). *Write Like a Chemist: A Guide and Resource*. Oxford; New York: Oxford University Press.

Roos, J. M. (2012). Measuring science or religion? A measurement analysis of the National Science Foundation sponsored science literacy scale 2006–2010. *Public Understanding of Science*. doi:10.1177/0963662512464318

Rubba, P. A. (1977). *The Development, Field Testing and Validation of an Instrument to Assess Secondary School Students' Understanding of the Nature of Scientific Knowledge* (Ed.D.). Indiana University, Bloomington, IN.

Samraj, B. (2002). Introductions in research articles: Variations across disciplines. *English for Specific Purposes*, *21*(1), 1–17.

Skelton, J. (1994). Analysis of the structure of original research papers: an aid to writing original papers for publication. *The British Journal of General Practice*, *44*(387), 455.

Stejskalová, T. (2010). *Lexical signals in the generic structure of popular scientific reports*. University of Pardubice, Pardubice, Czech Republic. Retrieved from http://dspace.upce.cz/handle/10195/36053

Stocking, S. H. (1999). How journalists deal with scientific uncertainty. In S. Friedman, S. Dunwoody, & C. Rogers (Eds.), *Communicating uncertainty: Media coverage of new and controversial science* (pp. 23–42). Mahwah, NJ: Lawrence Erlbaum.

Stocklmayer, S. M., Rennie, L. J., & Gilbert, J. K. (2010). The roles of the formal and informal sectors in the provision of effective science education. *Studies in Science Education*, *46*(1), 1–44.

Stutchbury, R. (1999). Science in the pub. In *Proceedings of the International Conference on Learning Science in Informal Contexts* (pp. 54–65). Canberra.

Suleski, J., & Ibaraki, M. (2009). Scientists are talking, but mostly to each other: a quantitative analysis of research represented in mass media. *Public Understanding of Science*, *19*(1), 115–125. doi:10.1177/0963662508096776

Swales, J. (1984). Research into the structure of Introductions to journal articles and its application to the teaching of academic writing. In *Common ground : shared interests in ESP and communication studies* (1st ed., pp. 77–86). New York: Pergamon Press.

Swales, J. (1990). *Genre Analysis: English in Academic and Research Settings*. Cambridge University Press.

Swales, J., & Najjar, H. (1987). The Writing of Research Article Introductions. *Written Communication*, *4*(2), 175 –191.

Tankard, J. W., & Ryan, M. (1974). News Source Perceptions of Accuracy of Science Coverage. *Journalism & Mass Communication Quarterly*, *51*(2), 219–225. doi:10.1177/107769907405100204

U.S. Census Bureau. (2013). *Educational Attainment in the United States: 2013*. Retrieved from http://www.census.gov/hhes/socdemo/education/

Woloshin, S., & Schwartz, L. M. (2002). Press Releases: Translating Research Into News. *JAMA: The Journal of the American Medical Association*, *287*(21), 2856–2858.

Woloshin, S., & Schwartz, L. M. (2006). Media reporting on research presented at scientific meetings: more caution needed. *Medical Journal of Australia*, *184*(11), 576–580.

Yaros, R. A. (2006). Is It the Medium or the Message? Structuring Complex News to Enhance Engagement and Situational Understanding by Nonexperts. *Communication Research*, *33*(4), 285 –309.

APPENDICES

Appendix A Texts Selected for Phase I

Table A-1 Texts Used in Move Analysis

Title	Newspaper	Date	Press Release Institution	Research Report Journal
The chemistry of space grows more complex	Christian Science Monitor	2-Aug-07	University of Arizona	Nature
Making plastic from plants?	Los Angeles Times	17-Feb-12	Utrecht University	Science
Researchers Create Nanostructures, and Whip Up a Recipe, Too	The New York Times	6-Sep-10	Northwestern University	Angewandte Chemie
What Exalts Stradivarius? Not Varnish, Study Says	The New York Times	4-Dec-09	Musee de la Musique	Angewandte Chemie
Diesel, Made Simply From Coffee Grounds (Ah, the Exhaust Aroma)	The New York Times	16-Dec-08	University of Nevada, Reno	Journal of Agricultural and Food Chemistry
What's the Peppery Note In Those Shirazes?	The New York Times	4-Jun-08	Australian Wine Research Institute	Journal of Agricultural and Food Chemistry
Sewage's Toxic Smell, Smothered by Coffee	The New York Times	21-Feb-12	City University of New York	Journal of Hazardous Materials
By Happy Accident, Chemists Produce a New Blue	The New York Times	24-Nov-09	Oregon State University	Journal of the American Chemical Society
Special Adhesive Helps Oysters Stick Together	The New York Times	31-Aug-10	Purdue University	Journal of the American Chemical Society
Researchers Develop A Type of Rubber That Can Repair Itself	The New York Times	26-Feb-08	Centre National de la Recherche Scientifique	Nature

Table A-2 Texts Used in Move Analysis Continued

Title	Newspaper	Date	Press Release Institution	Research Report Journal
Theory and Experiment Meet, and a New Form of Boron Is Found	The New York Times	3-Feb-09	Stony Brook University	Nature
Fingerprint Test Shows Not Only Who, but What	The New York Times	8-Aug-08	Purdue University	Science
The eyes have it	The Washington Post	2-Feb-10	Centre National de la Recherche Scientifique	Analytical Chemistry
Molecular Action May Help Keep Birds on Course	The Washington Post	5-May-08	University of Oxford	Nature
Chemical analysis of a comet's ice gives a clue to source of water on Earth	The Washington Post	11-Oct-11	Max Planck Institute	Nature
In Space, Clues to the Seeds of Life	The Washington Post	30-Jan-01	Ames Research Center	Proceedings of the National Academy of Sciences
Scientists Strengthen Spider Silk by Mixing In Metal	The Washington Post	27-Apr-09	Max Planck Institute	Science

Appendix B Move and Step Definitions

Moves

Presenting Background Information
Any text that provides some information that the reader must or should know in order to fully understand the current work.

Ex. Definition of terms, explanation of concepts, reference to previous work

Reviewing Related Research
Any text that references previous related research and/or discusses the limitations of previous work

Highlighting Overall Research Outcomes
Text near the beginning of the article that summarizes important findings or conclusions

Researcher Context
Text that describes who the researchers are, where the research was conducted, how it was funded, etc.

Describing Data Collection Procedure
Text that explains how the research was done

Explaining Research Outcome
Text that states or comments upon a result

Stating Research Conclusions
Text that summarizes the overall results, discusses implications, and/or talks about potential future work

Steps

Presenting Background Information
1. Explaining principles and concepts – explanation of a general concept or general background knowledge
2. Knowledge in the field – explanation of a general concept with some hedging (ie. it is thought that…, theory states…., hypothesized that…..)
3. Introducing the problem – statement providing background into the specific issue that needed to be researched
4. Potential implications – statement of importance/potential implications of achieving a specific goal (the research may contribute to, but not actually achieve the goal)

Reviewing Related Research
1. Reference to previous research – statement of what has been done previously
2. Indicating limitations of previous research – statements of limitations of previous work

Highlighting Overall Research Outcomes
1. Indicating main research result – summary statement of main result(s)
2. Implications – statement indicating the importance or implications of the main result(s)
3. Practical Limitations – explanation of the limitations of the research as related to practical applications
4. Scientific Limitations – explanation of the limitations of the research as related to scientific research
5. Referencing setting – statement of who the researchers were or where the research was conducted
6. Summary of method – summary statement of method(s) and/or materials
7. Anecdote – statements that show the human element of research

Researcher Context
1. Referencing setting - statement of who the researchers were or where the research was conducted
2. Referencing publication – statement of where the work has been published
3. Funding – statement describing the source of funding
4. Anecdote – statements that show the human element of research

Describing Data Collection Procedure
1. Materials – description of the materials used for the research
2. Experimental setup – description of the procedure of the experiment
3. Explanation of experiment – explanation of how the experiment works or why it was done
4. Data collected – description of what data were collected
5. Explaining principles and concepts – explanation of a general concept

Explaining Research Outcome
1. Explaining principles and concepts – explanation of a general concept
2. Stating specific outcome – statement of a result
3. Commenting on result – explaining/elaborating on a result
4. Reference to previous research – statement of what has been done previously

Stating Research Conclusions
1. Summary of results – summarization of all results
2. Practical Limitations – explanation of the limitations of the research as related to practical applications
3. Scientific Limitations – explanation of the limitations of the research as related to scientific research
4. Future work – proposal of future work

5. Implications – statement indicating the importance or implications of the result(s)
6. Research context – relation of the research to other researchers
7. Explaining principles and concepts – explanation of a general concept
8. Reference to previous research – statement of what has been done previously
9. Speculation – statements speculating about results
10. Anecdote – statements that show the human element of research
11. Practical applications – how the research results are being used

Appendix C Example Move Analysis

Below is the text of a news article from August 8, 2008 in *The New York Times* divided into the different identified moves.

Fingerprint Test Shows Not Only Who, but What

Highlighting Overall Research Outcome

With a new analytical technique, a fingerprint can now reveal much more than the identity of a person. It can now also identify what the person has been touching: drugs, explosives or poisons, for example.

Writing in Friday's issue of the journal Science, R. Graham Cooks, a professor of chemistry at Purdue University, and his colleagues describe how a laboratory technique, mass spectrometry, could find a wider application in crime investigations.

Presenting background information

The equipment to perform such tests is already commercially available, although prohibitively expensive for all but the largest crime laboratories. Smaller, cheaper, portable versions of such analyzers are probably only a couple of years away.

Describing Data Collection Procedure

In Dr. Cooks's method, a tiny spray of liquid that has been electrically charged, either water or water and alcohol, is sprayed on a tiny bit of the fingerprint. The droplets dissolve compounds in the fingerprints and splash them off the surface into the analyzer. The liquid is heated and evaporates, and the electrical charge is transferred to the fingerprint molecules, which are then identified by a device called a mass spectrometer. The process is repeated over the entire fingerprint, producing a two-dimensional image.

The researchers call the technique desorption electrospray ionization, or Desi, for short.

In the experiments described in the Science paper, solutions containing tiny amounts of various chemicals including cocaine and the explosive RDX were applied to the fingertips of volunteers. The volunteers touched surfaces like glass, paper and plastic. The researchers then analyzed the fingerprints.

Explaining Research Outcome

Because the spatial resolution is on the order of the width of a human hair, the Desi technique did not just detect the presence of, for instance, cocaine, but literally showed a pattern of cocaine in the shape of the fingerprint, leaving no doubt who had left the cocaine behind.

Stating Research Conclusions

"That's an advantage that this technique would have," said Bruce Goldberger, professor and director of toxicology at the University of Florida who runs a forensics laboratory that helps medical examiners and law enforcement. Dr. Goldberger was not involved in the research.

The chemical signature could also help crime investigators tease out one fingerprint out of the smudges of many overlapping prints if the person had been exposed to a specific chemical, said Demian R. Ifa, a postdoctoral researcher and the lead author of the Science paper.

Prosolia Inc., a small company in Indianapolis, has licensed the Desi technology from Purdue and is already selling such analyzers as add-ons to large laboratory mass spectrometers, which cost several hundred thousand dollars each.

Prosolia has so far sold about 70 analyzers, said Peter T. Kissinger, the company's chairman and chief executive. The most sophisticated $60,000 version that would be needed for fingerprint analysis went on sale this year.

However, fingerprints are not the main focus for Prosolia or Dr. Cooks. "This is really just an offshoot of a project that is really aimed at trying to develop a methodology ultimately to be used in surgery," Dr. Cooks said.

If a Desi analyzer can be miniaturized and automated into a surgical tool, a surgeon could, for example, quickly test body tissues for the presence of molecules associated with cancer. "That's the long-term aim of this work," Dr. Cooks said.

In unpublished research, the researchers have successfully tested the method on bladder tumors in dogs.

Prosolia is collaborating with Griffin Analytical Technologies, a subsidiary of ICx Technologies, on a Desi analyzer that works with a portable mass spectrometer. That product is probably a year or two away from the market, Dr. Kissinger said.

As it becomes cheaper and more widely available, the Desi technology has potential ethical implications, Dr. Cooks said. Instead of drug tests, a company could surreptitiously check for illegal drug use by its employees by analyzing computer keyboards after the workers have gone home, for instance.

Appendix D News Articles for Phase II

Article without described limitations

Chemical analysis of a comet's ice gives a clue to source of water on Earth

Astronomers find big clue to Earth's water

Astronomers have found the first comet with ocean like water, giving a major boost to the theory that celestial bodies were a significant source of water for a thirsty early Earth.

The intense heat of the planet immediately after it formed means that any initial water would have quickly evaporated; scientists believe the oceans emerged around 8 million years later. The puzzle is where the water, which is vital for life on Earth, came from.

Chemical analysis of water-ice from comets had suggested they could have delivered no more than 10 percent of the water in today's oceans. But research by Paul Hartogh of Germany's Max Planck Institute for Solar System Research and colleagues showed that a comet called 103P/Hartley 2 has the same chemical composition as the Earth's oceans.

The finding substantially increases the amount of water that might have originated from comets, which are largely made up of rock and ice. Previous models of the early Earth implied that most water came from asteroids.

In the case of Hartley 2, researchers using infrared instruments found that ice on the comet has a near identical "D/H" ratio to seawater. D/H measures the proportion of deuterium, or heavy hydrogen, to ordinary hydrogen.

"It was a big surprise when we saw the ratio was almost the same as what we find in the Earth's oceans," Hartogh said.

"It means it is not true anymore that a maximum of 10 percent of water could have come from comets. Now, in principle, all the water could have come from comets."

Hartogh, whose research was published online last week in the journal Nature, believes Hartley 2, whose current orbit around the sun does not extend much beyond Jupiter, started life in a different part of the solar system than other comets studied.

It probably formed in the Kuiper belt, which lies about 30 to 50 times farther from the sun than the Earth, while the others come from the Oort Cloud, some 5,000 times farther away.

Article with described limitations

Where did Earth's water come from?

Comet Hartley 2 offers new clues. The composition of comet Hartley 2 suggests that comets might have been a bigger source of Earth's water than previously thought. It's also challenging models of solar system formation.

For years, astronomers have been drafting a Kipling-like "Just So" story one might call "How the Earth Got Its Oceans." But they have had a tough time figuring out how to divvy up the credit between two potential sources - comets and asteroids.

Now, it seems, comets may have played a more significant role in drenching the third rock from the sun than previously thought.

Comet 103P/Hartley 2, which made its closest approach to the sun last October, contains water with virtually the same chemical signatures as water in the oceans, according to a study published Thursday in the journal Nature.

That signature shows up in the relative abundance of two forms of water: a typical water molecule, H_2O; and a much rarer type known as heavy water, in which one of the two hydrogen atoms has a neutron in its nucleus and the other doesn't.

But the findings raise new questions. The proportion of heavy water in the vapor spewed by Hartley 2 is much lower than theory says it should be, given where astronomers believe the comet formed. It's also lower than the proportion astronomers have measured in other comets so far.

"To me, this changes the problem," says Edwin Bergin, a University of Michigan astronomer and member of the team reporting the results.

Questions of the source for Earth's oceans are giving way to trying to figure out why comets have these differences in their water's chemistry and what that might imply for the formation and evolution of the solar system.

"That wasn't yesterday's problem," he acknowledges with a chuckle.

The team, led by Paul Hartogh with the Max Planck Institute for Solar System Research in Katlenburg-Lindau, Germany, used the European Space Agency's Herschel Space Observatory to analyze Hartley 2's halo, or coma, when the comet passed within 11 million miles of Earth shortly before its closest approach to the sun last year.

Much of the significance of Hartley 2 is where it comes from - a broad swath of frigid objects orbiting the sun beyond Neptune, called the Kuiper Belt.

Until now, scientists have only been able to measure the chemical signatures of comets from the Oort Cloud - a halo of comets much farther away from the sun at a distance of more than 5,000 astronomical units. (The Earth is 1 AU from the sun.)

These comets are thought to have formed just beyond the outer edge of today's asteroid belt between Mars and Jupiter, where sunlight is too feeble to thaw water ice. As Jupiter and other gas giant planets grew, their gravity flung these planetesimals deeper into space to form the Oort Cloud, researchers say.

From there, the comets might have slammed into Earth, delivering water and other volatiles, such as nitrogen.

But measurements of water in six Oort Cloud comets such as Hyakutake in 1996, or former Oort Cloud comets, such as Halley's Comet in 1986, showed twice the concentration of heavy water to normal water as did Earth's seawater.

Comets' stock plunged as cosmic tankers. Researchers concluded that comets could have contributed no more than about 10 percent of the oceans' water.

That would leave the bulk of the delivery to asteroids, which also contain water, notes Daruisz Lis, an astronomer at the California Institute of Technology and another member of the team reporting the results.

With Hartley 2, the pendulum may be swinging back in the comets' direction. But therein lie additional puzzles, Dr. Lis continues.

The reason: Hartley 2's composition doesn't appear to fit its birthplace.

Based on scientist's current understanding of cosmic chemistry, the deeper the chill, the higher the relative abundance of heavy water compared with H_2O in ices. The warmer it is, the lower the abundance.

That would imply that the comets that formed closest to the sun - the Oort Cloud comets - should have lower abundances of heavy water in their ices than comets that formed farther out in the Kuiper Belt.

Yet in Hartley 2, a Kuiper Belt native, astronomers have a comet with less heavy water in its ices than is present in comets that formed closer to the sun.

Hartley 2 is only one example, Lis acknowledges. It's unclear how representative the comet is of its relatives in the Kuiper Belt, called the Jupiter Family.

If it is representative, it could mean one of two things.

Scientists may need to revamp their models of heavy-water distribution in the disk of dust and gas surrounding the young sun. Or Hartley 2 could be a sign that objects early in the solar system's evolution were moving toward and away from the sun with some regularity as the giant planets in particular migrated from their birthplaces to their current orbits.

As for cometary collisions that would have delivered the water? Lis says they would have to have been very gentle.

"If you smash a comet into proto-Earth at very high velocity, there's a pretty good chance that the debris will be ejected back into space. Nothing would stay," he says. This would have to be true for asteroids as well as comets delivering water, he says.

An early Earth watered by cometary kisses?

Appendix E Demographics Survey Questions

Thank you for participating in this study. Please provide the following information about yourself.

Age

Under 21	22 - 34	35 - 44	45 - 54	55 - 64	65 and Over
○	○	○	○	○	○

Sex

○ Male

○ Female

Number of College or University Science Courses Completed

0	1 - 2	3 - 5	6 - 9	10 or More
○	○	○	○	○

Do you use your knowledge of science as part of your job?

Yes	No
○	○

On average, how many times per week do you read the news, either from a print newspaper or from the website of a news organization?

0	1 - 2	3 - 4	5 - 7
○	○	○	○

On average, how many times per week do you read <u>science</u> news, either from a print newspaper or from the website of a news organization?

0	1 - 2	3 - 4	5 - 7
○	○	○	○

On average, how often do you look for science information from any source (e.g. internet, book, magazine, museum, etc.)?

Never	Less than Once a Month	Once a Month	2-3 Times a Month	Once a Week	2-3 Times a Week	Daily
○	○	○	○	○	○	○

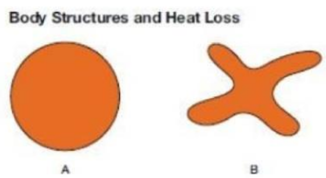
Appendix F Science Literacy Survey Questions

Now, we are going to do some detailed questions on science and technology. Please do not use any external resources to answer any of the following questions. Scientists and educators are interested in how familiar adults are with the things being taught in today's schools. Many of these questions are likely to concern things that weren't taught or emphasized when you were in school. Some of the questions involve pictures or graphs.

What property of water is most important for living organisms?

- ○ It is odorless
- ○ It does not conduct electricity
- ○ It is tasteless
- ○ It is liquid at most temperatures on Earth

The two objects shown below have the same mass, but object B loses heat more quickly than object A.



Referring to the above diagram, which combination of bodily features would be BEST suited to a small animal that lives in a cold climate and needs to minimize heat loss?

- ○ Long ears and a long body.
- ○ Small ears and a short tail.
- ○ A long nose and a long tail.
- ○ A short nose and large ears.
- ○ A long tail and a short nose.

Which of the following is a key factor that enables an airplane to lift?

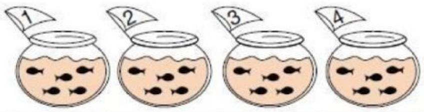
- ○ Air pressure beneath the wing is greater than that above the wing.
- ○ Pressure within the airplane is greater that that of the outside.
- ○ Engine power is greater than that of friction.
- ○ The plane's wing is lighter than air.

Lightning and thunder happen at the same time, but you see the lightning before you hear the thunder. Explain why this is so.

A solution of hydrochloric acid (HCl) in water will turn blue litmus paper red. A solution of the base sodium hydroxide (NaOH) in water will turn red litmus paper blue. If the acid and base solutions are mixed in the right proportion, the resulting solution will cause neither red nor blue litmus paper to change color. Explain why this is so.

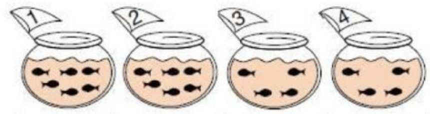
A student wants to find out if temperature affects the behavior of goldfish. He has 4 fish bowls and 20 goldfish. Which of the experiments shown below should he do?

○ A



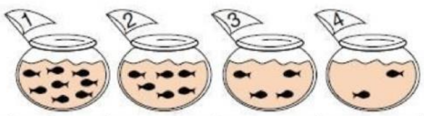
Number of fish	5 fish	5 fish	5 fish	5 fish
Temperature	15°C	20°C	25°C	30°C

○ B



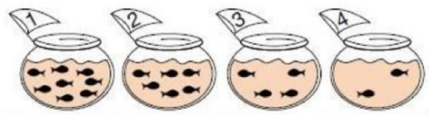
Number of fish	6 fish	6 fish	4 fish	4 fish
Temperature	20°C	20°C	30°C	30°C

○ C



Number of fish	8 fish	6 fish	4 fish	2 fish
Temperature	25°C	25°C	25°C	25°C

○ D



Number of fish	8 fish	6 fish	4 fish	2 fish
Temperature	15°C	20°C	25°C	30°C

A farmer thinks that the vegetables on her farm are not getting enough water. Her son suggests that they use water from the nearby ocean to water the vegetables. Is this a good idea?

- ○ Yes, because there is plenty of ocean water.
- ○ Yes, because ocean water has many natural fertilizers.
- ○ No, because ocean water is too salty for plants grown on land.
- ○ No, because ocean water is much more polluted than rainwater.

Which one of the following is NOT an example of erosion?

- ○ The wind in the desert blows sand against a rock.
- ○ A glacier picks up boulders as it moves.
- ○ A flood washes over a riverbank, and the water carries small soil particles downstream.
- ○ An icy winter causes the pavement in a road to crack.

Traits are transferred from generation to generation through the…

- ○ sperm only.
- ○ egg only.
- ○ sperm and egg.
- ○ testes.

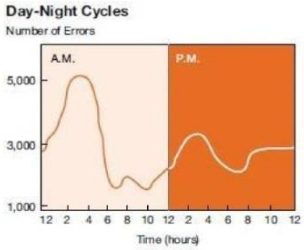
How do most fish get the oxygen they need to survive?

- ○ They take in water and break it down into hydrogen and oxygen.
- ○ Using their gills, they take in oxygen that is dissolved in water.
- ○ They get their oxygen from the food they eat.
- ○ They come to the surface every few minutes to breathe air into their lungs.

For which reason may people experience shortness of breath more quickly at the top of a mountain than along a seashore?

- ○ A slower pulse rate.
- ○ A greater gravitational force on the body.
- ○ A lower percent of oxygen in the blood.
- ○ A faster heartbeat.
- ○ A slower circulation of blood.

Day-night rhythms dramatically affect our bodies. Probably no body system is more influenced than the nervous system. The figure below illustrates the number of errors made by shift workers in different portions of the 24-hour cycle.



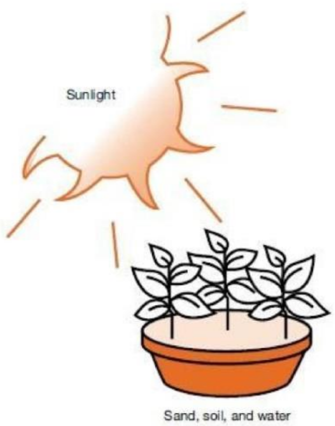
Based on the data illustrated in the figure above, during which of these time periods did the most errors occur?

- 2 A.M. to 4 A.M.
- 8 A.M. to 10 A.M.
- 12 P.M. to 2 P.M.
- 2 P.M. to 4 P.M.
- 8 P.M. to 10 P.M.

As part of a laboratory experiment, five students measured the weight of the same leaf four times. They recorded 20 slightly different weights. All of the work was done carefully and correctly. Their goal was to be as accurate as possible and reduce error in the experiment to a minimum. Which of the following is the BEST method to report the weight of the leaf?

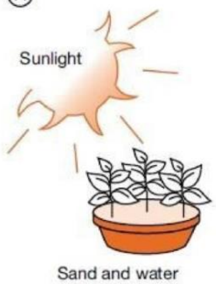
- Ask the teacher to weigh the leaf.
- Report the first measurement.
- Average all of the weights that were recorded.
- Average the highest and lowest weights recorded.
- Discard the lowest five weights.

A gardener has an idea that a plant needs sand in the soil for healthy growth. In order to test her idea she uses two pots of plants. She sets up one pot of plants as shown below.



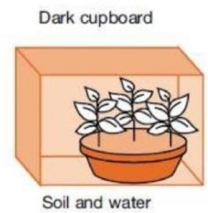
Which ONE of the following should she use for the second pot of plants?

Ⓐ

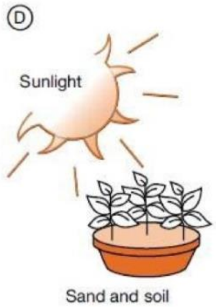


Ⓑ

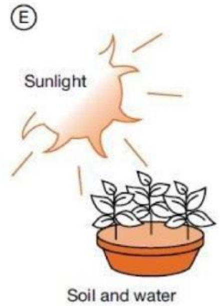
Ⓒ



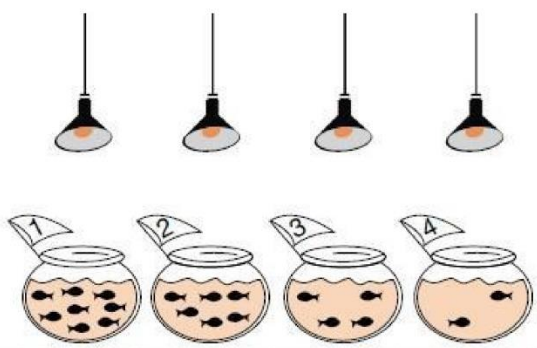
Ⓓ



Ⓔ



What might a scientist be trying to find out from the experiment represented below?



Number of fish	8 fish	6 fish	4 fish	2 fish
Temperature	25°C	25°C	25°C	25°C

- ○ If the number of fish in the fish bowl affects the behavior of the fish.
- ○ If the temperature of the fish bowl affects the behavior of the fish.
- ○ If the temperature and the amount of light affect the behavior of the fish.
- ○ If the number of fish, the temperature, and the amount of light affect the behavior of the fish.

Why did you choose that answer?

- ○ Because I already know what affects the behavior of fish.
- ○ Because that is what is allowed to change in this experiment.
- ○ Because that is what stays the same in this experiment.
- ○ Because that is what the scientist decided to include in this experiment.

Appendix G VOST Survey Questions

The following list contains items related to the views of science. Please indicate the extent to which you agree with the statements.

We all see the same moon because the moon is out there, outside ourselves, for all to see

Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree
○	○	○	○	○

Our scientific theories, especially any developed within the last ten years, are not likely to ever be changed.

Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree
○	○	○	○	○

Scientists do not agree entirely on the basic concept of the atom.

Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree
○	○	○	○	○

It is impossible to eliminate error and uncertainty from the measurement process, even with the very best equipment.

Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree
○	○	○	○	○

Atoms are thought to exist, but this has not been observed directly.

Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree
○	○	○	○	○

Science has gradually discovered that its nature, standing by its own strength, was an assumption rather than an established fact.

Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree
○	○	○	○	○

When the same experiment is performed any number of times, under exactly the same circumstances, the result is necessarily always the same.

Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree
○	○	○	○	○

In practice, the scientist follows a rigid step-by-step procedure in solving problems to ensure accurate results.

Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree
○	○	○	○	○

We cannot experience the whole of nature; consequently, we can never hope to understand it completely.

Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree
○	○	○	○	○

In sending missions to the moon, many assumptions were made because they are essential in scientific thinking.

Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree
○	○	○	○	○

Since scientific knowledge is changing all the time, scientific ideas are subject to being revised or thrown away.

Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree
○	○	○	○	○

Scientists do not know if the mass remains the same during chemical reactions, but they do know if any change occurs i must be small.

Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree
○	○	○	○	○

Scientific knowledge is constantly subject to revision.

Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree
○	○	○	○	○

There is no reason why we cannot obtain knowledge and have it change with the passage of time.

Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree
○	○	○	○	○

Scientists using the very best instruments can measure things exactly.

Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree
○	○	○	○	○

Science is the true and certain way to solve problems of nature and man.

Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree
○	○	○	○	○

When two people observe a chair, the sensations which this produces will never be quite identical to both people.

Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree
○	○	○	○	○

We can never say that any theory is final or corresponds to absolute truth because, at any moment, new facts may be discovered and compel us to abandon it.

Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree
○	○	○	○	○

In sending missions to the moon, no assumptions were made because everything had to be certain.

Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree
○	○	○	○	○

No two people ever observe the same rainbow in the same way.

Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree
○	○	○	○	○

It has been proven that there is no gain or loss of mass in any chemical reaction.

Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree
○	○	○	○	○

Scientific theories can be proven to be true.

Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree
○	○	○	○	○

Scientists will never be able to discover the exact position and exact speed of motion of every particle in the universe.

Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree
○	○	○	○	○

Our knowledge of nature can be visualized, little by little, in a number of different pictures, although no single picture enables us to visualize the whole of nature at once.

Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree
○	○	○	○	○

The scientist is content with a single exact observation.

Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree
○	○	○	○	○

In science, most evidence is arrived at or derived from some particular set of experimental data and then extended to an all-embracing theory.

Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree
○	○	○	○	○

Scientific theories are not provable in a classroom or in a well-equipped laboratory.

Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree
○	○	○	○	○

In laboratory experiments, it is impossible to record all possible observations.

Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree
○	○	○	○	○

The picture or model provides a representation, not of objective nature, but only of our knowledge of nature.

Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree
○	○	○	○	○

Observations are often very difficult to explain in terms of scientific theories.

Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree
○	○	○	○	○

We can know nothing of the world outside ourselves for certain.

Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree
○	○	○	○	○

The scientist no longer sees nature as something entirely distinct from himself or herself.

Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree
○	○	○	○	○

A measurement depends on the object being measured, the measuring instrument, and the observer.

Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree
○	○	○	○	○

With the exception of counting a small number of objects, there is always uncertainty and/or error in measurement.

Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree
○	○	○	○	○

When an experiment is repeated several times under identical conditions, several different results may be obtained.

Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree
○	○	○	○	○

The notion that scientific knowledge is certain is an illusion.

Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree
○	○	○	○	○

Appendix H Phase II Interview Protocol

Thank you for taking time to answer some questions for me today. Now that you've read that science news article, I'd like to ask you a few questions about it.

1. Could you please describe to me what the article that you just read was about?

2. According to the article, why was this research conducted?

3. According to the article, what were the results of the research?

4. According to the article, what were the scientists able to conclude from their research?

5. According to the article, why were these findings important?

6. In your opinion, do you feel that the research findings were important? Why or why not?

7. How much do you trust the research findings from this article? Why do you feel that way?

8. How much do you trust the research conclusions from this article? Why do you feel that way?

I am now going to read to you a series of statements that draw conclusions from the article that you just read. After each one, I would like you to tell me how correct you feel the statement is and why you feel that way.

1. Other comets from the same region of space as Hartley 2 will have a similar deuterium to hydrogen (D/H) ratio.

2. Comets provided a larger portion of the initial water to Earth than previously believed.

3. Scientists must now reconsider what they thought they knew about water on comets.

I would like to ask you a few more questions before we finish.

1. Would you describe yourself as having an interest in science? Why or why not?
    a. Why does that interest you?

2. Would you describe yourself as knowledgeable about science compared to the average person? Why or why not?

3. Could you please describe to me what a scientific theory is?

4. After scientists have developed a theory (e.g. atomic theory), does the theory ever change? Why or why not?
    a. If you believe that theories do change, explain why we bother to teach scientific theories.

5. What does an atom look like?
    a. How certain are scientists about the characteristics of atoms?
    b. What specific evidence do you think scientists use to determine what an atom looks like?

6. Some astronomers believe that the universe is expanding while others believe that it is shrinking; still others believe that the universe is in a static state without any expansion or shrinkage. How are these different conclusions possible if all of these scientists are looking at the same experiments and data?

Appendix I Phase III News Articles

The versions of the articles shown below contain both described limitations (**bold)** and explanation of methods (***bold/italic***). The other versions have either one or both sections removed and, if necessary, remaining sentences were altered slightly for grammatical correctness. Sentences that are in [ ] indicate that it is only present if the adjacent section is removed.

Comet Water

Chemical analysis of a comet's ice gives a clue to source of water on Earth

Astronomers find big clue to Earth's water

Astronomers have found the first comet with ocean like water, giving a major boost to the theory that celestial bodies were a significant source of water for a thirsty early Earth.

The intense heat of the planet immediately after it formed means that any initial water would have quickly evaporated; scientists believe the oceans emerged around 8 million years later. The puzzle is where the water, which is vital for life on Earth, came from.

Chemical analysis of water-ice from comets had suggested they could have delivered no more than 10 percent of the water in today's oceans. But research by Paul Hartogh of Germany's Max Planck Institute for Solar System Research and colleagues showed that a comet called 103P/Hartley 2 has the same chemical composition as the Earth's oceans.

The finding substantially increases the amount of water that might have originated from comets, which are largely made up of rock and ice. Previous models of the early Earth implied that most water came from asteroids.

In the case of Hartley 2, researchers using infrared instruments found that ice on the comet has a near identical "D/H" ratio to seawater. D/H measures the proportion of deuterium, or heavy hydrogen, to ordinary hydrogen.

In October and November, Hartley 2 passed Earth as closely as ever before since its discovery. During this encounter, the instruments on board the space observatory Herschel were aimed at the comet. With the help of exact observations of its coma - the sheath of gas and dust surrounding comets, when they come close to the Sun - the researchers hoped to determine the deuterium-to-hydrogen ratio.

"The water molecules within the coma emit a characteristic radiation in the far infrared", says Hartogh. This also holds true for the heavier cousin of water: water molecules in which one hydrogen atom has been replaced by a deuterium atom. "From this characteristic radiation we can determine the ratio of deuterium to hydrogen", he adds. However, since the heavy water is very rare, its radiation intensity is extremely weak. Nevertheless, with Herschel's HIFI instrument, the most sensitive detector for water vapor, the researchers were able to detect the molecule with an astonishingly good signal-to-noise ratio.

"It was a big surprise when we saw the ratio was almost the same as what we find in the Earth's oceans," Hartogh said.

"It means it is not true anymore that a maximum of 10 percent of water could have come from comets. Now, in principle, all the water could have come from comets."

Hartogh, whose research was published online last week in the journal Nature, believes Hartley 2, whose current orbit around the sun does not extend much beyond Jupiter, started life in a different part of the solar system than other comets studied.

It probably formed in the Kuiper belt, which lies about 30 to 50 times farther from the sun than the Earth, while the others come from the Oort Cloud, some 5,000 times farther away.

Hartley 2 is only one example, Daruisz Lis, an astronomer at the California Institute of Technology and another member of the team reporting the results, acknowledges. It's unclear how representative the comet is of its relatives in the Kuiper Belt.

As for cometary collisions that would have delivered the water? Lis says they would have to have been very gentle.

"If you smash a comet into proto-Earth at very high velocity, there's a pretty good chance that the debris will be ejected back into space. Nothing would stay," he says. This would have to be true for asteroids as well as comets delivering water, he says.

An early Earth watered by cometary kisses?

<u>Bird Compass</u>

Molecular Action May Help Keep Birds on Course

Four decades after scientists showed that migratory birds use Earth's magnetic field to orient themselves during their seasonal journeys, researchers have at last found a molecular mechanism that may explain how they do it.

If the hypothesis is true, the planet's magnetic field lines -- which arch around Earth from north to south -- may be plainly visible to birds, like the dashed line in the middle of a road.

The work, described online yesterday in the journal Nature, **was conducted in a test tube and does not prove that birds actually use the mechanism. And researchers aligned with a competing model say they are not convinced.**

But by identifying for the first time a molecule that reacts to very weak magnetic fields, the experiments prove the plausibility of a long-hypothesized method of avian navigation that has had a credibility problem because no one had ever found a molecule with the required sensitivity.

"This is a proof of principle that a chemical reaction can act as a magnetic compass," said Peter Hore of the University of Oxford, who with fellow chemist Christiane Timmel led the research.

Hore is testing similar molecules, called cryptochromes, isolated from the eyes of migratory birds. Devens Gust, a chemist at Arizona State University who worked with

Hore and Timmel, said the molecules "seem to have the right structural and chemical features to allow them to show this effect."

The seasonal comings and goings of birds have mystified people for millennia. Some early observers, noting that certain species routinely disappeared each year as others appeared, presumed that one species was somehow being transformed into the other. As late as the 18th century, an anonymous essayist who described himself simply as "a Person of Learning and Piety" concluded that many birds probably spend winters on the moon.

Recent scientific findings have seemed almost as incredible. By reversing the magnetic fields around captive birds as they prepared to migrate, scientists could induce them to take off in the wrong direction. The conclusion was that birds have a "sixth sense" that can detect magnetic energy the way eyes detect light and ears detect sound.

But how?

Two hypotheses have dominated. One centers on the discovery that birds (and other organisms, including salmon) make and store in their bodies a version of iron called magnetite, which orients itself to magnetic fields.

In birds, magnetite is often concentrated in the beak. Studies have shown that when the beaks of these birds are exposed to powerful magnetic fields -- or are numbed with an anesthetic -- the birds lose their ability to navigate properly.

But many scientists have suspected that another mechanism is also crucial -- one that can tell a bird not only which way is north but also how far it is from the equator by detecting the angle of magnetic field lines. Those lines emerge from Earth's magnetic poles perpendicular to the planet's surface, then arch overhead to meet over the equator, at which point they run parallel to the surface. If a bird could detect the angle of those lines relative to the surface, it could know, in effect, its latitude.

Scientists had theorized that a molecule with the right characteristics might change its behavior depending on the inclination of the magnetic field around it. It might react with another chemical more quickly, for example.

In the new work -- conducted in a chamber that blocks Earth's magnetic field and creates fresh ones of various strengths -- the team made a three-part molecule that, in response to light, gives up electrons at one end and passes them to the other end. There they linger for a millionth of a second or so before returning. Significantly, the precise amount of time each electron spends in its temporary home at the far end of the molecule varies with the angle of the surrounding magnetic field.

[In the model system a weak magnetic field interacts with electrons in an excited molecule, changing how long it takes to relax.]

If cryptochromes or other chemicals in a bird's eye behave as the new molecule does, they could provide the foundation of a bird's magnetic sense. Their shape would probably vary slightly, depending on how much time electrons spent at the far end, or those lingering electrons might affect the shape of another, nearby molecule in the eye. And shape determines biological function.

So depending on how far north or south a bird is from the equator, these molecules could be expected to send different signals to its brain, telling the flier whether it is veering east or west and pinpointing its latitude.

No one knows how a bird would perceive this input. Light looks like light. Sound sounds like sound. What would magnetic information "feel" or "look" like?

"It could be a bright or dark spot that would move around" in the bird's field of vision, Hore said. As in a video game, the goal might be to keep that spot centered.

But maybe not.

"I think it would be annoying to have this dot moving around," said Thorsten Ritz, a biophysicist at the University of California at Irvine, who nonetheless called the new work "breathtaking." Perhaps as a bird veered off course it would feel the way airplane passengers do in a quick descent, he suggested.

Others doubt that birds have, or need, anything more than their magnetite mouths.

"Hore is a great chemist, and this is an impressive demonstration of a weak field effect. However, I'm not sure it has any biological relevance," said Sõnke Johnsen, who studies bird navigation at Duke University.

Joe Kirschvink, an expert in magnetoreception at the California Institute of Technology, was even more dismissive, noting among other things that Hore's experiment worked only at very cold temperatures -- "a major stumbling block to the suggestion that optical effects in any organism can be used as the basis of a physiological compass," he said.

Hore and Ritz said similar molecules are expected to work at warmer temperatures.

And in the end, both camps may be right.

"Maybe there is a compass in the eye of birds," Ritz said, "and a map in their beaks."

Space Membranes

In Space, Clues to the Seeds of Life

Chemical 'Membranes' Could Revise Thinking on Origins

Scientists have for the first time shown that when simple chemicals are exposed to the harsh conditions of deep space, the molecules spontaneously arrange themselves into the hollow structures that look like the cell membranes found in all living things.

The work shows that early chemical steps considered important for the origin of life can form in space, the researchers said. It lends weight to arguments that life on Earth might have been "kick-started" billions of years ago when organic compounds such as these, born in cold interstellar clouds, landed on this planet aboard comets, meteorites and interplanetary dust.

"Scientists believe the molecules needed to make a cell's membrane, and thus for the origin of life, are all over space," said Louis Allamandola of NASA's Ames Research Center in California's Silicon Valley, who led the team. "This discovery implies that life could be everywhere in the universe."

The findings provide an intriguing new clue to one of science's biggest and most complex mysteries: How did life arise? The leading theory of the origin of life on Earth proposes that the early planet provided the rich, vast soup of chemical resources within which, somewhere, conditions emerged that favored the formation of chemical compounds and processes that led to the first living organisms. Instead, the researchers said, crucial early processes appear to take place in space long before planet formation occurs, with the implication that if the resulting compounds land in any favorable environment, they can easily trigger life.

John Hayes, a biogeochemist at the Marine Biological Laboratory at Woods Hole, Mass., who was not on the discovery team, said the work is significant in that it provides a mechanism "in the right place at the right time to deliver a lot of complicated organic material to early planetary surfaces."

But he cautioned that there are "a lot of banana peels" between there and the rise of living things, and that "a lot more study needs to be done" on the nature of these structures.

No one knows how life began on Earth, whether it was through naked genetic material drifting in a primordial sea or genetic material already encapsulated in membranes. But at some point, the researchers said, membranes became important.

"All life as we know it on Earth uses membrane structures to separate and protect the chemistry involved in the life process from the outside," said Jason Dworkin, of the SETI Institute, lead author of the team's paper published in today's issue of the Proceedings of the National Academy of Sciences. "All known biology uses membranes to capture and generate cellular energy."

Dworkin compared membranes (thin, two-layered sheets made up mostly of special fatty molecules) to a kind of housing. "Maybe these molecules were just the raw lumber lying around that allowed origin-of-life chemicals to move in and set up housekeeping or construct their own houses."

Bruce Runnegar, head of UCLA's Center for Astrobiology and not a member of the Allamandola team, said that, with the new evidence, "It's getting to the point where you can at least argue that cell membranes might have been a very early step on the pathway toward life on Earth." These hollow containers "are permeable and eventually have electrical properties, and so if you can sort of expect that they'd be available anyway, delivered to the primitive Earth from comets, then it might make sense to have them as an early step."

At Ames's Astrochemistry Laboratory, the team created an environment similar to that found in "empty" space, with temperatures close to absolute zero (minus 441 degrees Fahrenheit) in an extreme vacuum. They froze a mixture of common, familiar chemicals such as water, methanol (wood alcohol), ammonia and carbon monoxide -- the same ingredients known to make up the ice particles in the dense clouds between the stars.

The researchers then zapped these simple ices with the harsh, high-energy ultraviolet radiation that a nearby star in space would emit. ***They were then able to separate the components of the ices by size and*** when they put the resulting yellowish residue in water under a microscope at the University of California, Santa Cruz, they could see the solids spontaneously organizing themselves into the soap-bubble-like membranous structures, with "inside" and "outside" layers. ***The researchers were able to trap a dye in the membranes by letting them form in the presence of the dye, showing that the***

vesicles have interior space. Some of the compounds in the self-formed vesicles are so complex they glow, Dworkin said. That is, they are able to convert energy from the ultraviolet light to the visible range.

These structures themselves are not "life," Dworkin said: They lack the genetic information they need to evolve, as required under the accepted definition. "We're just starting to understand how these things work," he said.

Scientists have long known that ultraviolet irradiation of icy solids produces chemicals more complex than those originally present in the ice. There was speculation that some of them might have played an important role in early Earth chemistry.

In the Ames laboratory, this team has routinely made copies of the extremely cold ice particles that make up the interstellar clouds -- the birthplaces of stars and star systems, planets and smaller bodies.

Their goal had been merely to identify compounds that might be found on comets and other icy bodies, to guide planning for space missions. They were so surprised by the results that, Dworkin said, they spent months checking the experiment for error. "I was sure it was a contamination problem," he said. "But I couldn't get it not to work."

"Instead of finding a handful of molecules only slightly more complicated than the starting compounds, hundreds of new compounds are produced in every mixed ice we have studied," Ames space scientist Scott Sandford said.

The structures formed from the interstellar ices are similar to those formed from compounds found in a well-studied space rock -- the primitive Murchison meteorite that landed in Australia -- in work done earlier by chemist Dave Deamer of the University of California at Santa Cruz, a member of the Allamandola team. However, these compounds were created in the lab and not directly observed in space. Still, this suggests that interstellar ices might be the source of compounds delivered to Earth in the heavy bombardment by space rubble that occurred in its infancy. Today, more than a hundred tons of space stuff rains on Earth annually, much of it in the form of organic material (carbon-based compounds, some of which might form the building blocks of life).

"We are just now beginning to realize that we are only seeing the tip of the iceberg in terms of extraterrestrial molecular complexity," Allamandola said. "Very complex organic molecules that might be important for the origin of life could well be falling on the surfaces of newly formed planets everywhere."

Appendix J Phase III Survey Questions

The possible answers are separated by article and correct answers are in **bold**.

Please indicate which of the following were conclusions in the article that you just read.

Comet Water

Comets may have provided a larger portion of the initial water to Earth than was previously believed.

It is not clear if other comets have water similar to ocean water.

Hartley 2 provided some initial water to Earth.

Asteroids did not provide any water to Earth.

All comets have water similar to ocean water.

No more than 10% of the initial water on Earth came from comets.

The researchers were not able to conclude anything from this research.

Bird Compass

Chemical reactions could act as a magnetic compass.

Molecules exist that react differently depending on the angle of a magnetic field.

Compounds isolated from bird eyes react to weak magnetic fields.

Birds can “see” magnetic fields.

Birds have a sense that allows them to detect their latitude.

All birds navigate by detecting the orientation and angle of the Earth's magnetic field lines.

The researchers were not able to conclude anything from this research.

Space Membranes

Membranes can form in outer space.

It is not clear if interstellar ice contains membranes.

Complex compounds necessary for life are formed in outer space.

Life must exist on other planets.

Complex compounds from outer space played an important role on an early Earth.

The complex compounds needed for life on Earth originated from outer space.

The researchers were not able to conclude anything from this research.

Please answer the following questions related to the conclusions from the article that you read.

I find the conclusions: (untrustworthy – trustworthy)

Very U	U	Somewhat U	Somewhat T	T	Very T

I find the conclusions: (biased – unbiased)

Very B	B	Somewhat B	Somewhat U	U	Very U

I find the conclusions: (inaccurate – accurate)

Very I	I	Somewhat I	Somewhat A	A	Very A

I find the conclusions: (unreasonable – reasonable)

Very U	U	Somewhat U	Somewhat R	R	Very R

After each question, participants were asked to indicate why they chose their answer, with the following options given:

I trust the researchers

I don't trust the researchers

I don't have enough information

I based my response on the results of the research

The study was conducted well

The conclusion was unclear

It was unclear how the experiment was conducted

The researchers could have made a mistake

Not enough research was conducted

Other (Please specify)

For each of the articles below, the first three statements are at least somewhat unreasonable claims, while the last two are reasonable. In addition, only the subset of why options that makes sense with the participants' response will be provided as options. For example, a participant indicating they agree with the first comet water statement would be given the option "The results from the study are representative", but not "The results from the study may not be representative".

You will now be presented with a series of statements that draw conclusions from the article you read. Please indicate how much you agree with each statement and then indicate why you feel that way.

6 point Likert scale:

Strongly Disagree Disagree Somewhat Disagree Somewhat Agree Agree Strongly Agree

<u>Comet Water</u>

1. Other comets from the same region of space as Hartley 2 will have a similar deuterium to hydrogen (D/H) ratio.

   Why options:

   I trust the experts.

   I based my answer on the information in the article

   The results from the study may not be representative

The results from the study are representative

The article didn’t give me enough information

The research results were insufficient

The research results were sufficient

The results of the research were unclear

Other (Please specify)

2. Comets provided a larger portion of the initial water to Earth than previously believed.

Why options:

I trust the experts.

I based my answer on the information in the article.

The results from the study may not be representative

The results from the study are representative

The article didn’t give me enough information

The research results were insufficient

The research results were sufficient

The results of the research were unclear.

Other (Please specify)

3. Asteroids did not provide much of the initial water to Earth.

   Why options:

   I trust the experts.

   I based my answer on the information in the article.

   The results from the study may not be representative

   The results from the study are representative

   The article didn’t give me enough information

   The research results were insufficient

   The research results were sufficient

   The results of the research were unclear.

   Other (Please specify)

4. Scientists must now reconsider what they thought they knew about water on comets.

   Why options:

   I trust the experts.

   I based my answer on the information in the article.

   The results from the study may not be representative

   The results from the study are representative

The article didn’t give me enough information

The research results were insufficient

The research results were sufficient

The results of the research were unclear.

Other (Please specify)

5. More than 10% of the Earth’s water could have come from comets.

Why options:

I trust the experts.

I based my answer on the information in the article.

The results from the study may not be representative

The results from the study are representative

The article didn’t give me enough information

The research results were insufficient

The research results were sufficient

The results of the research were unclear.

Other (Please specify)

Bird Compass

1. Bird eyes contain molecules that can detect the angle of a magnetic field.

   Why options:

   I trust the experts.

   I based my answer on the information in the article.

   The results from the study may not be representative

   The results from the study are representative

   The article didn't give me enough information

   The research results were insufficient

   The research results were sufficient

   The results of the research were unclear.

   Other (Please specify)

2. Birds navigate by sensing the orientation and angle of magnetic field lines.

   Why options:

   I trust the experts.

   I based my answer on the information in the article.

   The results from the study may not be representative

   The results from the study are representative

   The article didn't give me enough information

The research results were insufficient

The research results were sufficient

The results of the research were unclear.

Other (Please specify)

3. Chemicals in birds' eyes work just like the molecule used in this study.

Why options:

I trust the experts.

I based my answer on the information in the article.

The results from the study may not be representative

The results from the study are representative

The article didn't give me enough information

The research results were insufficient

The research results were sufficient

The results of the research were unclear.

Other (Please specify)

4. Scientists must now reconsider what they thought they knew about how birds navigate.

   Why options:

   I trust the experts.

   I based my answer on the information in the article.

   The results from the study may not be representative

   The results from the study are representative

   The article didn't give me enough information

   The research results were insufficient

   The research results were sufficient

   The results of the research were unclear.

   Other (Please specify)

5. Molecules exist that react to weak magnetic fields.

   Why options:

   I trust the experts.

   I based my answer on the information in the article.

   The results from the study may not be representative

   The results from the study are representative

   The article didn't give me enough information

The research results were insufficient

The research results were sufficient

The results of the research were unclear.

Other (Please specify)

Space Membranes

1. Membranes are formed in ice particles in outer space.

Why options:

I trust the experts.

I based my answer on the information in the article.

The results from the study may not be representative

The results from the study are representative

The article didn’t give me enough information

The research results were insufficient

The research results were sufficient

The results of the research were unclear.

Other (Please specify)

2. Interstellar ices brought complex compounds needed for life to Earth.

   Why options:

   I trust the experts.

   I based my answer on the information in the article.

   The results from the study may not be representative

   The results from the study are representative

   The article didn't give me enough information

   The research results were insufficient

   The research results were sufficient

   The results of the research were unclear.

   Other (Please specify)

3. Cell membranes were an early step on the path to life on Earth.

   Why options:

   I trust the experts.

   I based my answer on the information in the article.

   The results from the study may not be representative

   The results from the study are representative

   The article didn't give me enough information

   The research results were insufficient

The research results were sufficient

The results of the research were unclear.

Other (Please specify)

4. Scientists must now reconsider what they thought they knew about the source of complex chemical compounds needed to begin life on Earth.

   Why options:

   I trust the experts.

   I based my answer on the information in the article.

   The results from the study may not be representative

   The results from the study are representative

   The article didn’t give me enough information

   The research results were insufficient

   The research results were sufficient

   The results of the research were unclear.

   Other (Please specify)

5. Simple molecules can organize into complex compounds in outer space.

   Why options:

   I trust the experts.

   I based my answer on the information in the article.

   The results from the study may not be representative

   The results from the study are representative

   The article didn’t give me enough information

   The research results were insufficient

   The research results were sufficient

   The results of the research were unclear.

   Other (Please specify)

VITA

VITA

Michael Hands Jr. was born in Providence, Rhode Island and grew up in Scituate, Rhode Island. After graduating high school, he attended Worcester Polytechnic Institute in Worcester, MA, where he received his B.S. in Biology, with a concentration in Cellular & Molecular Biology and Genetics and a minor in Biochemistry. Upon finishing his undergraduate degree, Michael attended Yale University, where he received a M.S. in Molecular Biophysics and Biochemistry under the supervision of Dr. Lynne Regan. He then spent a year teaching science informally as an educator at the Maritime Aquarium at Norwalk and for Little Scientists in Connecticut. These experiences cemented his love of teaching and desire to research informal science education.

In 2009, Michael enrolled at Purdue University to pursue a Ph.D. in Chemistry Education, where he joined Dr. Gabriela Weaver's group. While at Purdue, he also became interested in computational chemistry, so he completed a M.S. in Chemistry under the supervision of Dr. Lyudmila Slipchenko while working on his Ph.D. Michael also completed a Graduate Certificate in Applied Statistics. As a member of Dr. Weaver's group, he analyzed the presentation of chemistry research in print news and investigated how including research limitations and/or methods impacted reader understanding of the science.

PUBLICATION

The following is a publication submitted to the International Journal of Science Education Part B.

The Effect of Detailed Descriptions of Study Limitations and Methods in Chemistry News on Reader Evaluations of Claims

Michael D. Hands[a] and Gabriela Weaver[b]

[a] *Department of Chemistry, Purdue University, West Lafayette, IN*

[b] *Center for Teaching and Faculty Development, University of Massachusetts Amherst, Amherst, MA*

Abstract:

Despite numerous calls for improving scientific literacy, many American adults show a lack of understanding of experiments, scientific study, and scientific inquiry. News media is one important avenue for science learning, but is inconsistent in the presentation of scientific research limitations, potentially impacting reader understanding. In this study, members of the public (n=232) and science faculty (n=191) read a randomly assigned news article either presenting or omitting the study limitations. Participants reading articles presenting limitations were able to evaluate the reasonableness of claims based on the article better than those who read the article omitting limitations when accounting for their views on the tentativeness of science (ToS). Presenting limitations was important in identifying unreasonable claims for both public and science faculty, while ToS views predicted ability to identify reasonable claims for the public.

The past few decades have a seen a rise in concern about scientific literacy, both in academic and general public circles. While the term “scientific literacy” is somewhat ill-defined, it can be used to refer not just to what one learns in a classroom, but how one can use their science knowledge in other settings. According to the *National Science Education Standards*, “Scientific literacy entails being able to read with understanding articles about science in the popular press and... implies the capacity to pose and evaluate arguments based on evidence and to apply conclusions from such arguments appropriately” (National Research Council, 1996). Publications from other national organizations concerned with science education, *Science for All Americans: Project 2061* (American Association for the Advancement of Science, 1989), *Benchmarks for Science Literacy* (American Association for the Advancement of Science, 1993), and *Science and Engineering Indicators 2014* (National Science Board, 2014), are in general agreement with the NRC of the importance of scientific literacy and that citizens must be able to use their science knowledge to make informed decisions. Adults make numerous health, consumer, and political decisions on topics that have a scientific basis. Increasing an individual’s scientific literacy could empower him/her to make the most informed decisions for his/her specific situation. The criteria for being scientifically literate indicate that citizens must not only have science content knowledge, but fairly sophisticated scientific reasoning skills. However, even with this emphasis on scientific literacy, it is clear that most Americans are not scientifically literate (Miller, 1986, 2004; National Science Board, 2014).

Adult Scientific Literacy

Measures of scientific literacy have traditionally focused exclusively on science content knowledge (Brossard, 2006; Miller, 1986, 1998, 2004). Over the past 25 years, public responses to basic science knowledge questions have remained fairly consistent, ranging from a low of 39% correctly answering if *The universe began with a huge explosion* to a high of 84% correctly answering if *The center of the Earth is very hot* in the latest survey (National Science Board, 2014). However, there is evidence that some science content knowledge questions, particularly about evolution and the big bang, are actually measuring religious belief rather than science knowledge (Roos, 2012). Science content knowledge is only one aspect of scientific literacy, with scientific reasoning skills being at least as important and perhaps a better measure.

More recently, literacy measures have asked questions to assess understanding of probability, experiment, scientific inquiry, and scientific study (National Science Board, 2014). Responses to these questions have also remained fairly constant over the past 20 years, with levels of 65%, 34%, 33%, and 20% for understanding of probability, experiment, scientific inquiry, and scientific study respectively in the most current survey (National Science Board, 2014). Having less than a third of the adult population show an understanding of the scientific process is far below what I believe is desirable to have a scientifically literate society.

Adult Science Information Sources

Large surveys have found that American adults obtain most of their science information in informal learning settings, ranging from visits to museums, zoos, and aquariums to using various forms of media (Falk et al., 2007; National Science Board, 2014). Fifty percent of all American adults stated that they had visited a zoo or aquarium during the previous year, while 26% had visited a science/technology museum and 27% had visited a natural history museum (National Science Board, 2014). For those adults who did not have a minor in their household, attendance dropped to 44% for zoos/aquariums, 24% for natural history museums, and 25% for science/technology museums (National Science Board, 2014). This small drop in attendance among adults without children suggests that they are using these informal learning opportunities at least partially for their own benefit. When asked what their primary source of science information was, the vast majority of American adults cited some form of media, with 26% indicating online or print newspapers and magazines (National Science Board, 2014), making it the second most popular source, behind television.

However, there is little research on how adults judge the credibility of science news or how their level of scientific literacy affects their understanding. The vast majority of internet users report checking science information they found online in at least one other source (Horrigan, 2006), but that does not provide information on how they are interpreting the science. It is likely that an individual's level of scientific literacy influences their understanding of science news because it provides a framework for the individual to interpret the information presented. There is also the potential that reading science news could impact a person's scientific literacy because it is a learning experience (Britt et al., 2014). If that is the case, then science news could be an important avenue for promoting scientific literacy.

Presentation of science in the news

Since a substantial portion of the population use print media as their primary source of science information, it is important to understand how science is presented in these contexts. It has been suggested that most science news is presented from a perspective that highlights the benefits of scientific research on human life, which can de-emphasize other aspects of science, particularly scientific uncertainty (Maier et al., 2014). Brechman, Lee, and Cappella (2009) analyzed press releases and subsequent news reports about genetics. They found that both sources presented discoveries in a simplified, deterministic way and reported different content a substantial portion of the time (Brechman et al., 2009). Another study tracked one cancer genetics story from the primary research report to newspaper reports and found that the meaning of the findings changed, in addition to the way the findings were reported (Kua et al., 2004). These results were confirmed with analysis of multiple cancer genetics stories, with experts judging the press release as more representative of the original research article than news reports (Brechman et al., 2011). These results are consistent with other research indicating that science information in the news is simplified, science content can be different, conclusions can change, limitations and caveats are removed, and the research is depicted as more certain than in the original research article (Nelkin, 1995; Pellechia, 1997; Stocking, 1999; Tankard & Ryan, 1974). Press releases (Woloshin & Schwartz, 2002) and news reports of scientific conference presentations (Woloshin & Schwartz, 2006) also omit research study limitations. These studies suggest that any confusion about the findings of science research might be due to distortions made from the research article.

Aspects of news reports affecting reader outcomes

Some research has been conducted to look at how manipulations of news articles affect readers (Budescu et al., 2009; Corbett & Durfee, 2004; Jensen, 2008; Jensen et al., 2011; Yaros, 2006). Previous work has shown that the language used in science news reports can affect readers' perception of the science. The inclusion of a broader context and views of other scientists has a significant impact on readers' perception of the certainty of climate change (Corbett & Durfee, 2004). Members of the general public interpret phrases conveying probabilities differently than scientists, so altering the language used can bring reader interpretation more in line with the scientists' intent (Budescu et al., 2009). It has also been found that readers find scientists and journalists more trustworthy when the study limitations were reported in cancer research news reports (Jensen, 2008). Including limitations also reduced reader cancer fatalism and nutritional backlash (Jensen et al., 2011), which are negative responses to cancer news associated with unhealthy habits. Along with language used, the organization of a science news report can affect reader interest and comprehension (Yaros, 2006). Yaros altered the structure of two *New York Times* articles, one about cancer research and another about nanotechnology research, and found that the modified structure increased reader interest and comprehension. These studies suggest that the addition of context and study limitations could have positive impacts upon readers.

While the presentation of science in text news reports and how readers perceive that information have been investigated previously (Brechman et al., 2009, 2011; Budescu et al., 2009; Corbett & Durfee, 2004; Jensen, 2008; Jensen et al., 2011; Kua et al., 2004; Nelkin, 1995; Pellechia, 1997; Stocking, 1999; Tankard & Ryan, 1974;

Woloshin & Schwartz, 2002, 2006; Yaros, 2006), most of the news reports examined have been in the context of health and/or politically controversial science. These areas are the most common type of science reported (National Science Board, 2014; Suleski & Ibaraki, 2009), but studies investigating this type of science news may not be measuring just a reader's scientific literacy. A reader may have a personal connection to news about health or refer to their political values when reading politically controversial science that informs their perception of science news. Therefore, it is important to investigate science news coverage of topics in other fields that may have less of a personal connection to the reader in order to isolate science literacy from other influences on readers.

Research Questions

The research questions for this study are as follows:

1. How does the inclusion of study limitations in a science news article affect readers' ability to evaluate claims based on the research?
2. How does the inclusion of an explanation of the methods in a science news article affect readers' ability to evaluate claims based on the research?
3. How does the inclusion of study limitations interact with an explanation of the methods to affect the readers' ability to evaluate claims based on the research?
4. How does the reader's science background affect the readers' ability to evaluate claims based on the research?

Participants

In order to answer the above research questions, participants representing the general American adult population were needed, as well as science faculty at research universities to serve as a comparison group. A panel of 250 participants was purchased from Qualtrics Panels to recruit general public participants. Participants were chosen to be representative of the United States adult population, with sampling based on age and education. They were compensated by Qualtrics for participation. A total of 232 public participants provided complete responses that were included for analysis.

In order to recruit science faculty, five research universities from each of the four geographic regions of the United States (Northeast, Midwest, South, and West) were randomly selected, for a total of 20 universities. At each institution, all of the faculty members in natural science (biology, chemistry, physics, astronomy, and earth science) departments were contacted using their publicly available email addresses to participate in this study. The faculty members were only emailed a single time and were not provided compensation. A total of 191 faculty participants, with a discipline distribution similar to the emailed sample, provided complete responses that were included for analysis.

Design

This study used an online survey in a 2 (described limitations vs. no limitations) x 2 (described methods vs. no methods) x 3 (article topic) design. Using multiple articles allows for the examination of the stability of the results across different texts. All participants (n=232) were randomly assigned to read one science news article reporting on a specific chemistry research study. Participants provided background information, then read one of the news articles, and finally evaluated claims based on the article.

Materials

The articles used in this study were chosen by first conducting a search of the Lexis-Nexis database for news articles in the category of chemistry that were published in The New York Times, The Washington Post, Christian Science Monitor, USA Today, LA Times, or The Wall Street Journal since 2000. These newspapers were chosen because they are among the most widely circulated newspapers in the United States (Alliance for Audited Media, 2013). The results of the search were further limited by only including science news articles that reported on one specific published research study. Of this list, three articles were chosen, “Chemical analysis of a comet's ice gives a clue to source of water on Earth” from the October 11, 2011 issue of the Washington Post (Comet Water), “Molecular Action May Help Keep Birds on Course” from the May 5, 2008 issue of the Washington Post (Bird Compass), and “In Space, Clues to the Seeds of Life” from the January 30, 2001 issue of the Washington Post (Space Membranes). The first article was selected because a version describing limitations was available. The other two articles were the only others that discussed the scientific limitations. Using articles that already mention limitations allows for fewer confounding variables because it limits the amount of additions necessary to the texts. Short amounts of text was either added, in the case of the first article, or deleted to create alternate versions of the articles regarding limitations. In order to create alternate versions regarding methods, text from a related press release and/or research article was adapted and added to the news article.

All of the edited versions were reviewed by the Senior Writer/Editor of a large Midwestern public university's marketing and media news service for style.

Measures

Prior to reading one of the news articles, participants completed a survey consisting of three parts. The first part contained questions related to the participants' age, sex, education, news reading habits, and science information seeking habits. The second part was a science knowledge test (National Science Board, 2010) containing both content and reasoning questions. The last part was a subscale from the Nature of Scientific Knowledge Scale (Rubba, 1977) measuring an individual's views about the tentativeness of science (ToS).

After reading the news article, participants were asked to evaluate claims based on the research they had read. They were given statements that drew some further conclusion from what was presented in the article. Participants were then asked to indicate how much they agreed with each statement on a 6 – point Likert-type scale. Three of the statements were unreasonable, so participants should rate them low. These statements make claims well beyond the conclusions presented by the scientists involved in the research. The other two statements were reasonable, so participants should rate them high. These statements include a rewording of a conclusion from the article and a general statement that "Scientists must now reconsider what they thought they knew about _____", where the blank was research topic specific. All of the statements were verified as reasonable/unreasonable by the corresponding author of the original research article that was reported in the news. Ratings for the unreasonable claims were averaged to provide a measure of ability to identify unreasonable claims. A score for identifying reasonable claims was similarly generated. Finally, ratings for the unreasonable claims were reversed and averaged with the ratings for the reasonable claims to generate an overall indicator of ability to evaluate claims.

After participants rated each claim, they provided reasons for their rating chosen from a pre-generated list with the option to write in their own reason. This list was created to reflect the reasons offered by staff members working in non-science departments and science faculty at a large Midwestern public university in interviews as part of a pilot study focusing on the same outcomes as this study.

Data Analysis

In order to answer the research questions, initial three-way mixed model ANCOVAs were performed separately for public and faculty participants, with evaluation of claims score as the dependent variable, limitations and methods as fixed factors, article topic as a random factor, and science literacy score and tentativeness of science score as a covariates. Then, the best fit model was selected using the backwards elimination procedure of sequentially removing the variable contributing the least explanatory power to the model until only variables with $p < 0.05$ were left. Tukey post-hoc tests were performed for the selected models. Additionally, the frequency of reasons chosen for the public and faculty separated by limitations condition were analyzed using Fisher's exact test to provide some insight into the differences observed in the ANCOVAs.

Results

The best fit model for the public's overall evaluation of claims included only significant main effects for limitations, $F(1,229) = 11.12$, $p = 0.001$, and ToS score, $F(1,229) = 5.94$, $p = 0.016$. There were no significant main effects for methods or article topic and there were no significant interaction effects. A Tukey post-hoc test for limitations (Table 1) showed that participants reading an article describing limitations were better able to evaluate claims, with a small effect size ($d = 0.31$). The parameter estimate for ToS was 0.13, indicating that the more tentative a participant viewed science, the better their ability to evaluate the claims presented.

The lack of a significant interaction effect between ToS and limitations indicates that presenting limitations had a positive impact on participants' ability to evaluate claims at all levels of ToS views. This result indicates that describing limitations in a news article increased the ability to evaluate claims, even for those participants with a more sophisticated view of the tentativeness of science. The fact that there is no main or interaction effect for article topic shows that this result is stable across different texts. This indicates that the results may be general and not specific to a given article. It may have been expected that individuals with more naïve views of science may benefit from describing study limitations, but these results show that it also benefits those with more sophisticated views. Therefore, including a description of study limitations in science news articles may increase a readers' ability to evaluate claims from the article. However, there was no significant main or interaction effect for methods, indicating that the brief description of methods added to the news articles had no effect on the public participants' ability to evaluate claims.

Science faculty should provide the ideal case for evaluating claims from science news articles, as they routinely evaluate scientific claims as part of their work. Therefore, it is of interest to compare their results to those of the public. The best fit model for the faculty's overall evaluation of claims included only significant main effects for limitations, $F(1,187) = 4.35$, $p = 0.038$, and article topic, $F(2,187) = 19.29$, $p < 0.0001$. There were no significant main effects for methods or ToS and there were no significant interaction effects. A Tukey post-hoc test for limitations (Table 1) showed that participants reading an article describing limitations were better able to evaluate claims, with a small effect size ($d = 0.19$). A Tukey post-hoc test for article topic (Table 2) revealed that the faculty participants evaluated the claims made about the Space Membranes article less well than they did for the other articles.

While the significant main effect for article topic showed that there was a difference in the evaluation of claims among the texts, the lack of a significant interaction effect between article topic and limitations indicates that the effect of interest was stable. The overall evaluation of claims was lower for the Space Membrane article, but no interaction effect between article topic and limitations indicates that the effect of including limitations was similar in all three texts. This result indicates that describing limitations in a news article increased the ability to evaluate claims, just as it did for the public, though to a lower extent. This decrease in effect size may be due to the more extensive science background of the faculty causing a ceiling effect. The absence of a significant main or interaction effect for ToS is most likely a result of the faculty having similar views, as the vast majority of them scored very highly on that scale. It was unexpected that the faculty benefitted from the presence of study limitations in the news articles almost as much as the public did, as their greater familiarity with evaluating claims made by other scientists was expected to compensate for the absence of study limitations. However, most of the faculty participants were presented with articles outside of their field of expertise, so their familiarity may not have completely transferred to other science contexts. Finally, similar to the public, there was no significant main or interaction effect for methods, indicating that the brief description of methods added to the news articles had no effect on either group's ability to evaluate claims.

The previously discussed results focused on participants' overall evaluation of claims, however examining their evaluation of unreasonable and reasonable claims separately also yielded interesting results. The best fit model for the public's evaluation of unreasonable claims included only a significant main effect for limitations, $F(1,230) = 10.75$, $p = 0.001$. There were no significant main effects for ToS, methods, or article topic and no significant interaction effects. A Tukey post-hoc test for limitations (Table 3) showed that participants reading an article describing limitations were better able to evaluate unreasonable claims, with a small effect size ($d = 0.30$), as evidenced by the decrease in average agreement.

Once again, the main and interaction effects for article topic were not significant, indicating that the results were stable across the different texts. In contrast to overall ability to evaluate claims, participants' views of ToS had no effect on evaluating unreasonable claims. Only the presence of limitations predicted ability to identify unreasonable claims. This result indicates a participants' views of the tentativeness of science did not help or hinder their ability to evaluate unreasonable claims, suggesting that including a description of limitations is more important in aiding an individuals' ability to identify unreasonable claims than views of the tentativeness of science. As with the overall evaluation of claims, inclusion of research methods had no effect on the public's ability to evaluate unreasonable claims.

As a comparison, the best fit model for the faculty's evaluation of unreasonable claims included only significant main effects for limitations, $F(1,187) = 3.80$, $p = 0.05$, and article topic, $F(2,187) = 13.47$, $p < 0.0001$. There were no significant main effects for ToS or methods and no significant interaction effects. A Tukey post-hoc test for limitations (Table 3) showed that participants reading an article describing limitations were better able to evaluate unreasonable claims, with a small effect size ($d = 0.18$), as evidenced by the decrease in average agreement. A Tukey post-hoc test for article topic (Table 4) once again revealed that faculty participants evaluated the unreasonable claims about the Space Membranes article less well than those of the other articles.

As with the results for overall evaluation of claims, the lack of a significant interaction between article topic and limitations indicated that the effect of including limitations was stable across the different texts. Similar to the public's evaluation of unreasonable claims, participants' views of ToS and inclusion of methods had no effect, while the presence of limitations predicted ability to identify unreasonable claims. These results indicate that including limitations in the news article aided both the public and the faculty in evaluating unreasonable claims.

The best fit model for the public's evaluation of reasonable claims included just a significant main effect for ToS, $F(1,230) = 9.33$, $p = 0.003$. There were no significant main effects for limitations (Table 5), methods, or article topic, and no significant interaction effects. The parameter estimate for ToS was 0.33, indicating that the more tentative a participant viewed science, the better they were at identifying reasonable claims. The main and interaction effects for article topic were not significant, indicating that the results were stable across the different texts. In addition, the main and interaction effects of methods were also not significant. In contrast to overall ability to evaluate claims and ability to evaluate unreasonable claims, the presence of limitations had no effect on evaluating reasonable claims. Only participant views of ToS predicted ability to identify reasonable claims. This result indicates that presenting limitations had no effect on ability to evaluate reasonable claims, suggesting that participants are able to identify reasonable claims using their views of the tentativeness of science and limitations are not needed.

Unlike the results for the public's evaluation of reasonable claims, there was no best fit model for the faculty's evaluation of reasonable claims. None of the variables measured in this study predicted their ability to evaluate reasonable claims. This result suggests that the science faculty are generally adept at evaluating reasonable claims, as their average scores were fairly high (Table 5), and that the manipulations to the news articles did not contribute to their evaluations. The faculty are likely to be more familiar with the scientific process than the general public, which may have contributed to their high ability to evaluate reasonable claims.

In order to provide more insight into participants' evaluation of claims, the reasons indicated for their responses were analyzed. For the evaluation of unreasonable claims (Table 6), the public showed significant differences in the reasons chosen based on the presence/absence of limitations for trust the experts, the results may not be representative, and the results are unclear. Public participants reading an article without limitations were significantly more likely to indicate that they trusted the experts than those reading an article including limitations. They also were less likely to indicate that the results may not be representative and that the results were unclear when the limitations were absent. It is interesting that while the public evaluated the unreasonable claims better when limitations were included in the news article, they were less likely to indicate that they trusted the experts and more likely to indicate the results may not be representative and were unclear. This may indicate that including limitations affected the public's trust in the researchers and their confidence in identifying the results. It is likely that knowledge of the results of the research reported in the news article is necessary to properly evaluate the claims, so it is interesting that public participants were able to better evaluate claims with limitations present while simultaneously feeling that the results were unclear. However, there were a few public participants who indicated that the results may not be representative when the limitations were present, which would be a valid reason for disagreeing with some of the unreasonable claims. It appears that these participants used the limitations to correctly identify a flaw in some of the unreasonable claims, which may partially explain why including limitations led to better evaluation of unreasonable claims.

For the faculty's evaluation of unreasonable claims (Table 6), there were significant differences in the reasons chosen based on the presence/absence of limitations for based on information in the article and the results are representative. They were more likely to indicate that they based their evaluation on the information in the article and less likely to indicate the results were representative when the limitations were present. This result may imply that the faculty felt more comfortable relying on the information in the news article when the limitations were present. It is also interesting that the public used the presence of limitations as an indicator that the results may not be representative, while the faculty used the absence of limitations as an indicator that the results were representative. In addition to the differences within participant groups, the public and faculty were significantly different in their frequency of indicating they used the reasons of trust the experts, based on the information in the article, and the results may not be representative. Overall, the public used their trust of the experts as a reason much more often than the faculty. However, they were less likely to indicate that they used their trust in the experts to evaluate the unreasonable claims when limitations were present, while the faculty had a slight trend in the opposite direction. A similar trend was observed for the reason based on the information in the article. The opposite trend was observed for using the reason that the results may not be representative, possibly indicating a difference in how the faculty and public view limitations.

For the evaluation of reasonable claims (Table 7), the public again showed significant differences in the reasons chosen based on the presence/absence of limitations for trust the experts. They more frequently indicated that they used their trust in the experts as a reason when evaluating reasonable claims when the limitations were absent, just as when evaluating unreasonable claims. In contrast, the faculty showed no significant differences based on the presence/absence of limitations. However, when comparing between participant groups, there once again was a significant difference in the frequency of use of the reason of trusting the experts. Overall, the public used their trust of the experts as a reason much more often than the faculty, while being less likely to indicate that they used their trust in the experts to evaluate the reasonable claims when limitations were present and the faculty having slight trend in the opposite direction.

Discussion

This study provides evidence that describing limitations in science news articles may impact readers' ability to evaluate claims based on the research. Only including limitations had a significant effect when evaluating unreasonable claims, while only ToS views had a significant effect when evaluating reasonable claims. These results imply that while a more sophisticated view of ToS may be needed to identify reasonable claims, a description of limitations may be necessary to identify unreasonable claims. While it may have been expected that including limitations would aid the public in evaluating claims, it was unexpected that it also aided the science faculty. The more sophisticated views of ToS that the faculty had decreased the effect that including the limitations had, but did not eliminate it.

While both the public and faculty showed similar results from the ANCOVA analysis, there were some differences in the reasons provided to evaluate the claims. The public tended to provide reasons that imply that they viewed the inclusion of limitations more negatively than the faculty. This possible difference in the interpretation of the limitations should be investigated further to determine any effects on reader perceptions of the news articles. Including limitations could affect readers' trust in the researchers and/or results, beliefs about the quality of the research, or willingness to apply the results in their daily lives. Correctly evaluating claims indicates an understanding of the results reported in the news article, but does not provide information on what the reader might think about the research.

Science news articles are inconsistent with their presentation of research limitations, which this study suggests may impact how readers evaluate claims about the research. Omitting study limitations may inhibit a readers' ability to identify unreasonable claims about the research. Therefore, it may be beneficial to include a description of research limitations in science news articles to improve communication between scientists and the public. Research is needed to identify any other factors that contribute to the ability to evaluate claims. In addition, the articles chosen for this study were politically and emotionally neutral, so it is not clear if these results would also apply to other science topics.

References

Abd-El-Khalick, F., Waters, M., & Le, A.-P. (2008). Representations of nature of science in high school chemistry textbooks over the past four decades. *Journal of Research in Science Teaching*, *45*(7), 835–855. doi:10.1002/tea.20226

Abraham, M. R., Grzybowski, E. B., Renner, J. W., & Marek, E. A. (1992). Understandings and misunderstandings of eighth graders of five chemistry concepts found in textbooks. *Journal of Research in Science Teaching*, *29*(2), 105–120. doi:10.1002/tea.3660290203

Alberts, B. (2009). Redefining Science Education. *Science*, *323*, 437–438.

Alliance for Audited Media. (2013). *Total Circulation for US Newspapers*. Retrieved from http://abcas3.auditedmedia.com/ecirc/newstitlesearchus.asp

American Association for the Advancement of Science. (1989). *Science for All Americans: Project 2061*. New York: Oxford University Press.

American Association for the Advancement of Science. (1993). *Benchmarks for science literacy*. New York: Oxford University Press.

Ayers, G. (2008). The evolutionary nature of genre: An investigation of the short texts accompanying research articles in the scientific journal Nature. *English for Specific Purposes*, *27*(1), 22–41. doi:10.1016/j.esp.2007.06.002

Bell, P., Lewenstein, B., Shouse, A. W., & Feder, M. A. (2009). *Learning Science in Informal Environments: People, Places, and Pursuits*. Washington, D.C.: Committee on Learning Science in Informal. Environments, National Research Council. Retrieved from http://www.nap.edu/catalog.php?record_id=12190

Brechman, J. M., Lee, C., & Cappella, J. N. (2009). Lost in Translation? *Science Communication*, *30*(4), 453 –474.

Brechman, J. M., Lee, C., & Cappella, J. N. (2011). Distorting Genetic Research About Cancer: From Bench Science to Press Release to Published News. *Journal of Communication*, *61*(3), 496–513. doi:10.1111/j.1460-2466.2011.01550.x

Britt, M. A., Richter, T., & Rouet, J.-F. (2014). Scientific Literacy: The Role of Goal-Directed Reading and Evaluation in Understanding Scientific Information. *Educational Psychologist*, *49*(2), 104–122. doi:10.1080/00461520.2014.916217

Brody, M., Bangert, A., & Dillon, J. (2007). Assessing Learning in Informal Science Contexts. *Commissioned Paper). Washington, DC: National Research Council. Accessed December*, *5*, 2011.

Brossard, D. (2006). Do They Know What They Read? Building a Scientific Literacy Measurement Instrument Based on Science Media Coverage. *Science Communication*, *28*(1), 47–63. doi:10.1177/1075547006291345

Budescu, D. V., Broomell, S., & Por, H.-H. (2009). Improving Communication of Uncertainty in the Reports of the Intergovernmental Panel on Climate Change. *Psychological Science*, *20*(3), 299–308. doi:10.1111/j.1467-9280.2009.02284.x

Chiappetta, E. L., & Fillman, D. A. (2007). Analysis of Five High School Biology Textbooks Used in the United States for Inclusion of the Nature of Science. *International Journal of Science Education*, *29*(15), 1847–1868. doi:10.1080/09500690601159407

Committee on Science, Engineering, and Public Policy. (2007). *Rising Above the Gathering Storm: Energizing and Employing America for a Brighter Economic Future*. Washington, D.C.: National Academy Press.

Corbett, J. B., & Durfee, J. L. (2004). Testing Public (Un)Certainty of Science: Media Representations of Global Warming. *Science Communication*, *26*(2), 129–151.

Csongor, A. (2013). *Rhetorical Moves and Hedging in Medical Research Articles and their Online Popularizations*. University of Pécs, Pécs, Hungary.

Deng, F., Chen, D.-T., Tsai, C.-C., & Chai, C. S. (2011). Students' Views of the Nature of Science: A Critical Review of Research. *Science Education*, *95*(6), 961–999.

Devore, J. L. (2008). *Probability and Statistics for Engineering and the Sciences* (7th ed.). Belmont, CA: Duxbury Press/Thomson Brooks/Cole.

Dhingra, K. (1999). *An ethnographic study of the construction of science on television* (Ed.D.). Retrieved from http://search.proquest.com.ezproxy.lib.purdue.edu/dissertations/docview/304517365/abstract/139971AE84D584E45EA/1?accountid=13360

Dhingra, K. (2003). Thinking about television science: How students understand the nature of science from different program genres. *Journal of Research in Science Teaching*, *40*(2), 234–256. doi:10.1002/tea.10074

Dhingra, K. (2006). Science on Television: Storytelling, Learning and Citizenship. *Studies in Science Education*, *42*(1), 89–123.

Dierking, L. D., & Falk, J. H. (1994). Family behavior and learning in informal science settings: A review of the research. *Science Education*, *78*(1), 57–72. doi:10.1002/sce.3730780104

Falk, J. H. (1983). Time and behavior as predictors of learning. *Science Education*, *67*(2), 267–276. doi:10.1002/sce.3730670214

Falk, J. H. (2001). *Free-Choice Science Education: How We Learn Science outside of School. Ways of Knowing in Science and Mathematics Series.* Williston, VT: Teachers College Press.

Falk, J. H., & Adelman, L. M. (2003). Investigating the impact of prior knowledge and interest on aquarium visitor learning. *Journal of Research in Science Teaching*, *40*(2), 163–176. doi:10.1002/tea.10070

Falk, J. H., & Dierking, L. D. (2000). *Learning from museums: visitor experiences and the making of meaning*. Lanham, MD: Rowman & Littlefield.

Falk, J. H., & Dierking, L. D. (2010). The 95 Percent Solution. *American Scientist*, *98*(6), 486.

Falk, J. H., & Needham, M. D. (2011). Measuring the impact of a science center on its community. *Journal of Research in Science Teaching*, *48*(1), 1–12. doi:10.1002/tea.20394

Falk, J. H., & Needham, M. D. (2013). Factors Contributing to Adult Knowledge of Science and Technology. *Journal of Research in Science Teaching*, *50*(4), 431–452. doi:10.1002/tea.21080

Falk, J. H., & Storksdieck, M. (2010). Science Learning in a Leisure Setting. *Journal of Research in Science Teaching*, *47*(2), 194–212. doi:10.1002/tea.20319

Falk, J. H., Storksdieck, M., & Dierking, L. D. (2007). Investigating public science interest and understanding: evidence for the importance of free-choice learning. *Public Understanding of Science, 16*(4), 455 –469.

Goldacre, B. (2010). *Bad Science: Quacks, Hacks, and Big Pharma Flacks* (Reprint edition.). New York: Faber & Faber.

Hillis, S. R. (1975). The development of an instrument to determine student views of the tentativeness of science. *Research and Curriculum Development in Science Education: Science Teacher Behavior and Student Affective and Cognitive Learning*, *3*, 32–39.

Hopkins, A., & Dudley-Evans, T. (1988). A genre-based investigation of the discussion sections in articles and dissertations. *English for Specific Purposes*, *7*(2), 113–121.

Horrigan, J. (2006). *The Internet as a resource for news and information about science*. Pew Internet and American Life Project. Retrieved from http://www.pewinternet.org/~/media//Files/Reports/2006/PIP_Exploratorium_ Science.pdf

Jensen, J. D. (2008). Scientific Uncertainty in News Coverage of Cancer Research: Effects of Hedging on Scientists' and Journalists' Credibility. *Human Communication Research*, *34*(3), 347–369.

Jensen, J. D., Carcioppolo, N., King, A. J., Bernat, J. K., Davis, L., Yale, R., & Smith, J. (2011). Including Limitations in News Coverage of Cancer Research: Effects of News Hedging on Fatalism, Medical Skepticism, Patient Trust, and Backlash. *Journal of Health Communication*, *16*(5), 486–503.

Kanoksilapatham, B. (2003). *A corpus-based investigation of scientific research articles: Linking move analysis with multidimensional analysis* (Ph.D.). Georgetown University, Washington, D.C.

Kanoksilapatham, B. (2005). Rhetorical structure of biochemistry research articles. *English for Specific Purposes*, *24*(3), 269–292.

Korpan, C. A., Bisanz, G. L., Bisanz, J., & Henderson, J. M. (1997). Assessing literacy in science: Evaluation of scientific news briefs. *Science Education*, *81*(5), 515–532. doi:10.1002/(SICI)1098-237X(199709)81:5<515::AID-SCE2>3.0.CO;2-D

Kua, E., Reder, M., & Grossel, M. (2004). Science in the News: A Study of Reporting Genomics. *Public Understanding of Science*, *13*(3), 309–322.

Lederman, L. M., & Malcom, S. M. (2009). The Next Campaign. *Science*, *323*, 1265–1266.

Lederman, N. G. (2007). Nature of Science: Past, Present, and Future. In S. K. Abell & N. G. Lederman (Eds.), *Handbook of Research on Science Education* (pp. 831–880). Mahwah, N.J.: Lawrence Erlbaum Associates.

Lederman, N. G., Abd-El-Khalick, F., Bell, R. L., & Schwartz, R. S. (2002). Views of nature of science questionnaire: Toward valid and meaningful assessment of learners' conceptions of nature of science. *Journal of Research in Science Teaching*, *39*(6), 497–521.

Maier, M., Rothmund, T., Retzbach, A., Otto, L., & Besley, J. C. (2014). Informal Learning Through Science Media Usage. *Educational Psychologist*, *49*(2), 86–103. doi:10.1080/00461520.2014.916215

Maxim, L., & Mansier, P. (2014). How is Scientific Credibility Affected by Communicating Uncertainty? The Case of Endocrine Disrupter Effects on Male Fertility. *Human and Ecological Risk Assessment*, *20*(1), 201–223. doi:10.1080/10807039.2012.719387

Meyer, P. (1988). Defining and Measuring Credibility of Newspapers: Developing an Index. *Journalism & Mass Communication Quarterly*, *65*(3), 567–574. doi:10.1177/107769908806500301

Miller, J. D. (1986). Reaching the attentive and interested publics for science. In S. Friedman, S. Dunwoody, & C. Rogers (Eds.), *Scientists and journalists: Reporting science as news* (pp. 55–69). New York: Free Press.

Miller, J. D. (1998). The measurement of civic scientific literacy. *Public Understanding of Science*, *7*(3), 203–223.

Miller, J. D. (2004). Public Understanding of, and Attitudes toward, Scientific Research: What We Know and What We Need to Know. *Public Understanding of Science*, *13*(3), 273–294.

Miller, J. D., Augenbraun, E., Schulhof, J., & Kimmel, L. G. (2006). Adult Science Learning from Local Television Newscasts. *Science Communication*, *28*(2), 216–242.

National Academies of Science. (2011). *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas*. Washington, D.C.: National Research Council. Retrieved from http://www.nap.edu/catalog.php?record_id=13165

National Research Council. (1996). *National Science Education Standards*. Washington, D.C.: National Academies Press.

National Science Board. (2010). *Science and Engineering Indicators 2010*. Arlington, VA: National Science Foundation.

National Science Board. (2014). *Science and Engineering Indicators 2014*. Arlington, VA: National Science Foundation.

Nelkin, D. (1995). *Selling science: How the press covers science and technology*. New York: W.H. Freeman.

Nwogu, K. N. (1991). Structure of science popularizations: A genre-analysis approach to the schema of popularized medical texts. *English for Specific Purposes*, *10*(2), 111–123.

Pellechia, M. (1997). Trends in science coverage: A content analysis of three US newspapers. *Public Understanding of Science*, *6*(1), 49–68. doi:10.1088/0963-6625/6/1/004

Rennie, L. J. (1995). Learning in science centres: What do we know and what do we need to know?’. In *Proceedings of the 20th Annual Conference of the Western Australian Science Education Association* (pp. 69–74).

Rennie, L. J., & McClafferty, T. (1995). Using visits to interactive science and technology centers, museums, aquaria, and zoos to promote learning in science. *Journal of Science Teacher Education*, *6*(4), 175–185. doi:10.1007/BF02614639

Rennie, L. J., & Williams, G. (2000). The nature and measurement of learning from attending a public lecture on human genetics. *Critical Studies in Education*, *41*, 17–34.

Rennie, L. J., & Williams, G. F. (2006). Adults' Learning about Science in Free-Choice Settings. *International Journal of Science Education*, *28*(8), 871–893.

Robinson, M. S., Stoller, F., Constanza-Robinson, M., & Jones, J. K. (2008). *Write Like a Chemist: A Guide and Resource*. Oxford; New York: Oxford University Press.

Roos, J. M. (2012). Measuring science or religion? A measurement analysis of the National Science Foundation sponsored science literacy scale 2006–2010. *Public Understanding of Science*. doi:10.1177/0963662512464318

Rubba, P. A. (1977). *The Development, Field Testing and Validation of an Instrument to Assess Secondary School Students' Understanding of the Nature of Scientific Knowledge* (Ed.D.). Indiana University, Bloomington, IN.

Samraj, B. (2002). Introductions in research articles: Variations across disciplines. *English for Specific Purposes*, *21*(1), 1–17.

Skelton, J. (1994). Analysis of the structure of original research papers: an aid to writing original papers for publication. *The British Journal of General Practice*, *44*(387), 455.

Stejskalová, T. (2010). *Lexical signals in the generic structure of popular scientific reports*. University of Pardubice, Pardubice, Czech Republic. Retrieved from http://dspace.upce.cz/handle/10195/36053

Stocking, S. H. (1999). How journalists deal with scientific uncertainty. In S. Friedman, S. Dunwoody, & C. Rogers (Eds.), *Communicating uncertainty: Media coverage of new and controversial science* (pp. 23–42). Mahwah, NJ: Lawrence Erlbaum.

Stocklmayer, S. M., Rennie, L. J., & Gilbert, J. K. (2010). The roles of the formal and informal sectors in the provision of effective science education. *Studies in Science Education*, *46*(1), 1–44.

Stutchbury, R. (1999). Science in the pub. In *Proceedings of the International Conference on Learning Science in Informal Contexts* (pp. 54–65). Canberra.

Suleski, J., & Ibaraki, M. (2009). Scientists are talking, but mostly to each other: a quantitative analysis of research represented in mass media. *Public Understanding of Science*, *19*(1), 115–125. doi:10.1177/0963662508096776

Swales, J. (1984). Research into the structure of Introductions to journal articles and its application to the teaching of academic writing. In *Common ground : shared interests in ESP and communication studies* (1st ed., pp. 77–86). New York: Pergamon Press.

Swales, J. (1990). *Genre Analysis: English in Academic and Research Settings*. Cambridge University Press.

Swales, J., & Najjar, H. (1987). The Writing of Research Article Introductions. *Written Communication*, *4*(2), 175 –191.

Tankard, J. W., & Ryan, M. (1974). News Source Perceptions of Accuracy of Science Coverage. *Journalism & Mass Communication Quarterly*, *51*(2), 219–225. doi:10.1177/107769907405100204

U.S. Census Bureau. (2013). *Educational Attainment in the United States: 2013*. Retrieved from http://www.census.gov/hhes/socdemo/education/

Woloshin, S., & Schwartz, L. M. (2002). Press Releases: Translating Research Into News. *JAMA: The Journal of the American Medical Association*, *287*(21), 2856–2858.

Woloshin, S., & Schwartz, L. M. (2006). Media reporting on research presented at scientific meetings: more caution needed. *Medical Journal of Australia*, *184*(11), 576–580.

Yaros, R. A. (2006). Is It the Medium or the Message? Structuring Complex News to Enhance Engagement and Situational Understanding by Nonexperts. *Communication Research*, *33*(4), 285 –309.

Table 1: Means for Limitations for Overall Evaluation of Claims

	Omitted Limitations	Included Limitations
Public	3.39[a]	3.57[a]
Faculty	3.81[b]	3.97[b]

Significance calculated between omitted vs. included limitations
[a] $p = 0.001$, [b] $p = 0.038$

Table 2: Means for Article Topic for Faculty Overall Evaluation of Claims

	Mean
Comet Water	4.10[a]
Bird Compass	4.03[a]
Space Membranes	3.54[b]

Cells with different letters are significantly different, $p < 0.0001$

Table 3: Means for Limitations for Evaluation of Unreasonable Claims

	Omitted Limitations	Included Limitations
Public	4.23[a]	3.89[a]
Faculty	3.69[b]	3.43[b]

Significance calculated between omitted vs. included limitations. A decrease in value indicates more disagreement with the statement and, therefore, a more accurate response.
[a] $p = 0.001$, [b] $p = 0.05$

Table 4: Means for Article Topic for Faculty Evaluation of Unreasonable Claims

	Mean
Comet Water	3.29[a]
Bird Compass	3.35[a]
Space Membranes	4.03[b]

Cells with different letters are significantly different, $p < 0.0001$. A decrease in value indicates more disagreement with the statement and, therefore, a more accurate response.

Table 5: Means for Limitations for Evaluation of Reasonable Claims

	Omitted Limitations	Included Limitations
Public	4.29	4.24
Faculty	4.55	4.59

No significance differences were found between omitted vs. included limitations

Table 6: Frequency of Reasons Provided for Evaluation of Unreasonable Claims

		Omitted Limitations	Included Limitations
Trust the experts	Public	80[a,b]	49[a,b]
	Faculty	12[b]	20[b]
Based on information in the article	Public	166[b]	146[b]
	Faculty	138[a,b]	174[a,b]
Results may not be representative	Public	8[a,b]	20[a,b]
	Faculty	36[b]	28[b]
Results are representative	Public	62	51
	Faculty	21[a]	9[a]
Not enough information	Public	40	50
	Faculty	66	66
Results are insufficient	Public	22	33
	Faculty	39	47
Results are sufficient	Public	47	40
	Faculty	11	13
Results are unclear	Public	27[a]	53[a]
	Faculty	11	13
Other	Public	13	17
	Faculty	58	63

Participants could choose more than one reason

[a] Significant difference between omitted vs. included limitations within a group ($p < 0.05$)

[b] Significant difference in distribution of omitted and included limitations between groups ($p < 0.05$)

Table 7: Frequency of Reasons Provided for Evaluation of Reasonable Claims

		Omitted Limitations	Included Limitations
Trust the experts	Public	56[a,b]	34[a,b]
	Faculty	11[b]	17[b]
Based on information in the article	Public	116	116
	Faculty	131	144
Results may not be representative	Public	3	8
	Faculty	3	9
Results are representative	Public	35	29
	Faculty	16	6
Not enough information	Public	24	24
	Faculty	17	16
Results are insufficient	Public	7	12
	Faculty	4	7
Results are sufficient	Public	30	30
	Faculty	28	25
Results are unclear	Public	29	36
	Faculty	4	5
Other	Public	8	10
	Faculty	39	39

Participants could choose more than one reason
[a] Significant difference between omitted vs. included limitations within a group ($p < 0.05$)
[b] Significant difference in distribution of omitted and included limitations between groups ($p < 0.05$)