NO POINTS FOR STYLE? ANALYSIS OF THE PSYCHOLOGICAL EFFECTS OF JOURNALISTIC WRITING CONVENTIONS

A Dissertation in
Mass Communications

by
Alyssa Appelman

© 2015 Alyssa Appelman

Submitted in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

May 2015
The dissertation of Alyssa Appelman was reviewed and approved* by the following:

Michael Schmierbach  
Associate Professor  
College of Communications  
Dissertation Adviser  
Chair of Committee

S. Shyam Sundar  
Distinguished Professor  
College of Communications

James P. Dillard  
Professor  
College of the Liberal Arts  
Department of Communication Arts and Sciences

J. Ford Risley  
Professor  
College of Communications  
Interim Associate Dean for Undergraduate and Graduate Education

*Signatures are on file in the Graduate School.
ABSTRACT

Journalistic editing has remained fairly constant over time, despite ongoing changes to the media landscape. Traditional skills are still taught to journalism students and still employed in traditional newsrooms. This study sought to determine the psychological effect of journalistic writing conventions on modern audiences. Through the paradigm of dual-processing models of persuasion (i.e., the Elaboration Likelihood Model and the Heuristic-Systematic Processing Model), this study explores two possibilities: Do readers notice mistakes and then consciously decide not to trust the content? Or do mistakes serve as distractions that unconsciously inhibit understanding? In other words, are errors heuristic cues, or are they inhibitors that block in-depth processing?

A between-subjects experiment (N = 504) was conducted to observe the relationship between errors and perceptions. Participants were shown news articles with various mistakes (i.e., grammatical errors, AP style errors, and inverted pyramid structure errors) and responded to questions about their perceptions. Participants also were put in either a high-motivation or low-motivation condition to determine whether perceptions differed based on attention.

Results showed no significant effects of errors on recall, writing quality, message credibility, or informativeness. These results held regardless of which errors were considered and which article was read. Power analysis showed that the sample was large enough to detect even a small effect size. This suggests that the effects are either contingent on study design or no longer existent. Future studies could determine whether such effects are still present.
If, as this study suggests, these effects can no longer be found, then we could be observing a shift in the way readers view historic standards for writing. Based on this study, readers are more willing to overlook style and structure errors than has been found in previous research. Therefore, this study proposes that journalists and journalism educators in the 21st century need to rethink the emphasis on style in the newsroom and the classroom.
## TABLE OF CONTENTS

List of Tables ................................................................................................................. vii  
List of Figures ................................................................................................................ viii  
Acknowledgements ........................................................................................................ ix  

Chapter 1. INTRODUCTION ......................................................................................... 1

Chapter 2. LITERATURE REVIEW ........................................................................... 5  
Historical Development of Journalistic Writing Conventions ...................................... 5  
  Grammar .................................................................................................................. 6  
  Associated Press Style ........................................................................................... 7  
  Inverted Pyramid Structure ................................................................................... 9  
Cognitive Processing of Journalistic Writing Conventions .......................................... 10  
  Elaboration Likelihood Model .............................................................................. 11  
  Heuristic-Systematic Processing Model ................................................................. 12  
Journalistic Writing Conventions in the HSM Framework ......................................... 13  
  News as context .................................................................................................. 13  
  Media judgments as outcomes ........................................................................... 16  
  Errors as ability constraints ................................................................................. 16  
  Errors as heuristic cues ....................................................................................... 18  
  Motivation manipulation as analysis ................................................................... 22  
  Message factors as processing mechanisms ....................................................... 24

Chapter 3. METHODS ............................................................................................... 26  
  Design .................................................................................................................. 26  
  Participants .......................................................................................................... 26  
  Stimulus ............................................................................................................... 28  
  Manipulated Variables ....................................................................................... 30  
    Writing-convention errors ................................................................................ 30  
    Motivation ......................................................................................................... 37  
  Measured Variables ............................................................................................ 38  
    Pretest ............................................................................................................... 38  
    Processing time ................................................................................................ 38  
    Recall ................................................................................................................ 39  
    Writing quality ................................................................................................ 40  
    Message credibility ........................................................................................... 40  
    Informativeness ................................................................................................. 41  
    Source credibility ............................................................................................... 41  
    Source-related message credibility ................................................................. 42  
    Message cues ................................................................................................... 42  
    Enjoyment ......................................................................................................... 43  
    Quality .............................................................................................................. 43
Chapter 4. RESULTS
Manipulation Check: Perceived Motivation
Descriptive Statistics
Hypothesis Testing
\ H1 ................................................................. 48
\ H2 ................................................................. 49
\ H3 ................................................................. 49
\ H4 ................................................................. 50
Article differences .................................................. 52
\ H5 ................................................................. 54
\ RQ1 ................................................................. 54
Additional Analyses .................................................. 57
Motivation manipulation ............................................ 57
Demographics ....................................................... 61

Chapter 5. DISCUSSION
Interpretation of Findings
\ H1 ................................................................. 64
\ H2 ................................................................. 65
\ H3 ................................................................. 66
\ H4 ................................................................. 67
\ H5 ................................................................. 68
\ RQ1 ................................................................. 68
Demographics ....................................................... 69
Limitations and Future Research ................................... 71
Theoretical and Practical Implications ............................ 79
Errors as expectancy-violations ................................... 79
Rethinking journalism education ................................... 82
Conclusion ................................................................ 85

References ................................................................ 87

Appendix A: Stimulus News Articles – Control Condition ........................................... 95
Appendix B: Pilot Test 1 Questionnaire — Grammatical Errors ................................... 100
Appendix C: Pilot Test 1 Questionnaire — AP Style Errors ......................................... 101
Appendix D: Stimulus News Articles — Treatment Conditions ................................. 102
Appendix E: Recall Questions ................................................................................. 118
Appendix F: Discussions among MTurk Workers .................................................. 119
List of Tables

Table 1: Stimulus Article Readability Scores ..........................................................28
Table 2: Pairwise Correlations between Measured Variables .................................47
Table 3: Descriptive Statistics for Perceptions by Error Type and Motivation ............47
Table 4: Source Table for 2(Motivation) x 4(Error Type) Between-Subjects ANOVA ....51
Table 5: Descriptive Statistics for Perceptions by Error Type and Motivation for Individual Articles ..................................................................................................................53
Table 6: Standardized Regression Weights for Mediating Effects of Errors on Overall News Quality .................................................................55
Table 7: Standardized Regression Weights for Significant Mediating Effects of Errors on Overall News Quality .................................................................57
List of Figures

Figure 1: Path Analysis for Mediating Effects of Errors on Overall News Quality ..........56

Figure 2: Path Analysis for Significant Mediating Effects of Errors on Overall News Quality .................................................................................................................................57
Acknowledgements

Thank you to my wonderful adviser, Dr. Mike Schmierbach, as well as my fantastic committee — Dr. Ford Risley, Dr. Shyam Sundar, and Dr. James Dillard — for their support throughout my doctoral program. Thank you also to the American Copy Editors Society for funding this study and providing ongoing support of my research.
Chapter 1

Introduction

Journalistic editing has remained fairly constant over time, despite ongoing changes to the media landscape. Traditional skills are still taught to journalism students and still employed in traditional newsrooms. Editing textbooks from across decades have taught virtually identical rules (e.g., Brooks, Pinson, & Sissors, 2005; Hyde, 1920; McGiffert, 1972; Neal, 1933; Taylor & Scher, 1951). These foundational rules are outlined in “The Art of Editing the News” in the form of questions copy editors should ask themselves as they are reading news articles: “(1) Is the story grammatical? (2) Does it conform to style? (3) Is it consistent? (4) Is it accurate? (5) Is it complete? (6) Does it have the right lead? (7) Is it well organized? (8) Is it fair? (9) Is it objective? (10) Is it in good taste? (11) Is it clear? (12) Is it concise? (13) Is everything spelled right? (14) Am I sure about all the names? (15) Have I checked all dates and numbers? (16) Have I deleted anything that I should restore?” (McGiffert, 1972, p. 2).

This study focuses on the particular journalistic writing conventions that some sources have suggested might be obsolete. These include: grammar, Associated Press style and inverted pyramid structure. These are in contrast to the other skills mentioned above because those refer specifically to notions of accuracy. Previous research already has found that accuracy affects modern audiences (e.g., Appelman & Sundar, forthcoming; Kang & Yang, 2011; Sundar, 1999), so this study does not test that possibility. Instead, it focuses on the stylistic conventions that are being questioned in the current media environment.
South (2012) explains the modern dilemma that journalists face: “Edit a story or create a Storify? Enforce AP (Associated Press) style or craft an SEO (search engine optimization) headline? Attend SPJ (Society of Professional Journalists) or head to ONA (Online News Association)” (p. 9). He explains that these industry changes lead to questions about stylistic conventions not only within the industry, but also within academia. Which journalistic skills should we teach students? “Curriculum development can be a zero-sum game: For every skill or topic we add, we often must subtract something” (South, 2012, p. 9).

Does that mean we should teach web skills instead of writing styles? Should we eliminate lessons on grammar and Associated Press style rules in favor of social media training? Do journalism students still need to learn traditional writing conventions? Despite discussions by media practitioners (“Really, j-schools need to ditch AP style and start teaching their students SEO instead. More valuable to their careers.” — Journalist Robert Niles [South, 2012]), academic research has not fully addressed these concerns.

It could be that newer forms of media make traditional skills dated and obsolete. People have become accustomed to reading text messages and blogs and, therefore, might be less bothered by casual writing than in the past. Maybe modern news audiences do not care about journalistic writing conventions and professionalism anymore. On the other hand, it could be that journalistic writing conventions need to be maintained. Maybe traditional journalistic writing conventions help to distinguish traditional journalism outlets from user-generated content. Maybe news articles with proper style and structure are seen as more professional to readers and signal to them to think more seriously about the content. Research suggests possible answers to these questions, particularly in terms
of grammar, but we have yet to fully explore the ways other formal features of news content can affect audience perceptions. In fact, these “traditional” conventions were never based on empirical support to begin with. Journalism educators have taught AP style and inverted pyramid structure with the assumption that they help readers; however, we do not know if that is actually the case. Does that consistency of AP style help readers? Is inverted pyramid structure easier to understand than articles written in chronological order? This study seeks to address these questions and, by doing so, advance theoretical and practical understandings of the influence of journalistic writing conventions.

First, this study broadly considers how journalistic writing conventions are processed by readers. Through the framework of dual-processing theories of persuasion (Chaiken, 1980; Chaiken, Liberman, & Eagly, 1989; Petty & Cacioppo, 1986a, 1986b), this study explores two possibilities: Do readers notice mistakes and then consciously decide not to trust the content? Or do mistakes serve as distractions that unconsciously inhibit understanding? This study then looks more narrowly at the effects that these writing conventions have on reader judgments, specifically recall, writing quality, message credibility, and informativeness.

Theoretically, this study advances dual-processing theories by exploring how we process news articles and whether that is linked to journalistic writing conventions. Studies have shown that errors matter to readers’ perceptions (Appelman, 2009; Appelman & Bolls, 2011; Appelman & Schmierbach, 2014) and are noticed by readers (Vultee, 2011). In addition, Lang (1989) found that news articles written in inverted pyramid style are more difficult to process than news articles written chronologically, and
Gibson and Zillmann (1993, 1998) found that the presence of direct quotes in news articles leads to different perceptions than articles with no quotes and with paraphrases. Although these effects have been found, the psychological mechanism is unclear. Do errors cue other perceptions of the text, or do they prevent us from understanding the content? In other words, are they heuristic cues, or are they inhibitors that block in-depth processing? Sparks and Areni (2008) provided support for both possibilities. Using the Elaboration Likelihood Model, they tested whether language power acted as a peripheral cue, a biasing influence, or a distracting influence. They operationalized language power as the presence or absence of powerless speech markers (i.e., tag questions, hedges, verbal hesitations, and nonverbal hesitations). They found support for language power as both a peripheral cue and a distracting influence, and they offered the following conclusion: “In the case of language power, style does appear to be more important than substance” (Sparks & Areni, 2008, p. 57). The current study advances theory, then, by determining whether Sparks and Areni’s conclusions can be applied to the psychological processing of journalistic writing conventions.

Practically, this study informs journalists of the writing conventions they can use to effectively deliver news content, and whether historic standards for writing still matter to audiences. It will provide empirical support for the effectiveness, or lack thereof, of traditional writing styles. In doing so, this study will also inform journalism educators of the most pertinent skills to teach their students as they prepare to enter the industry.
Chapter 2

Literature Review

Historical Development of Journalistic Writing Conventions

Before exploring processes and effects, this study looks at the historical development of journalistic writing conventions. As mentioned above, journalistic writing conventions refer to the traditional rules and processes of creating news articles. In a conventional newsroom, reporters construct a news article from information gathered through interviews and background research. They then send the article to copy editors, who read the content for grammar, style, structure, and clarity. In other words, there is a reporting stage, followed by an editing stage.

A study of journalistic conventions could include any number of traditions at either stage. At the reporting stage, for example, it could examine accuracy, the journalistic method, and interviewing. As mentioned above, previous research strongly suggests that accuracy, or truth, affects modern audiences, so this study does not test that possibility (e.g., Appelman & Sundar, forthcoming; Kang & Yang, 2011; Sundar, 1999). The journalistic method refers to a technique of original information gathering, or “the pursuit of independently verifiable facts about current events through enterprise, observation and investigation” (Stephens, 1988, p. 229). Interviewing is a writing convention by which reporters ask sources for information, which the reporter then attributes to the source. This study suggests that these methods are so ingrained in our culture that they are not thought of as conventions anymore, but would, arguably, be used by anyone writing for the mass media. A blogger who writes inaccurate information without observing or interviewing, for example, would likely be discounted by most
readers. For the sake of parsimony, then, this study focuses on the journalistic writing conventions that are being questioned in the current media environment (e.g., South, 2012), which are those related to the editing stage.

For the purposes of this study, journalistic writing conventions refer to traditional rules of grammar, style, and structure. Grammar refers to formal standards of usage, spelling and punctuation; style refers to the Associated Press style; and structure refers to the inverted pyramid structure (Brooks, Pinson, & Sissors, 2005; Hyde, 1920; Neal, 1933; McGiffert, 1972; Taylor & Scher, 1951). Significant time and energy is spent on teaching these skills in journalism classrooms. Journalism departments have stand-alone editing classes that focus on these rules, and introductory reporting classes employ grammar and AP style quizzes as part of their curricula. Journalists who fill this editing role, known as copy editors, have existed in some form in newsrooms since the 19th century (Keith, 2009).

These editing-stage journalistic writing conventions are discussed in detail below, in terms of their historical development and potential effects on news readers, so as to better understand and predict how they are processed by readers.

Grammar. Grammar refers to formal standards of usage, spelling and punctuation. Because there is not really one “American” standard of grammar, this study focuses on English standards enforced by journalists in their canonic literature (e.g., Brooks, Pinson, & Wilson, 2006; O’Conner, 2003; Truss, 2006; Webster’s, 2005). “Grammar for Journalists” outlines this standard: “Whatever medium you work for, you should always strive to use ‘good English,’ which is not necessarily ‘correct English,’” as advocated by the strict grammarians. Let your chief objective be to write carefully,
The connection between journalistic grammar and quality is widely anticipated despite limited empirical research; journalism professionals have discussed it for decades. Journalism schools teach copy editing to students as part of their basic course loads; copy editors formed their own professional society, the American Copy Editors Society (ACES); and professional journalism publications analyze each other’s errors (e.g. the Columbia Journalism Review’s Language Corner). The underlying assumption of such efforts is that editing will improve the text quality (see also B. Walsh, 2008).

Professional discussions of the importance of grammar extend beyond newsrooms, as well. Columnists have suggested that speakers lose credibility when they use incorrect grammar (Grazian, 1996), recruiters overlook job candidates who misuse words (Grazian, 1996; Charney & Rayman, 1989), and grammatically incorrect emails are problematic (McGoon, 1996).

Although oft discussed in these professional settings, the relationship between grammatical errors and perceptions of news content has not been extensively studied in academic research. Exceptions include Appelman and Bolls (2011) and Vultee (2011), which found that readers notice grammatical errors in news articles and can distinguish between edited and unedited news articles. This study seeks to corroborate these grammar effects, and it extends the research by also empirically testing the potential impact of Associated Press style and inverted pyramid structure.

Associated Press style. Associated Press style, or AP style, is a traditional news media style that includes conformity regulations for spelling and abbreviations. For the
purposes of this study, it represents the journalistic convention of *style*. The first AP stylebook came out in 1953, and it is used as the standard for news media style. “It proclaimed itself the ‘most definitive and inclusive work ever undertaken by a group of newspapers,’ and yet it proved to be just the start of a work that is constantly in progress” (Curley, 2008).

Why did this become the formal style convention for journalistic writing? Tom Curley, the former president and executive officer of the Associated Press, explained the stylebook’s purpose as such: “To provide a uniform presentation of the printed word, to make a story written anywhere understandable everywhere” (Curley, 2008). Thus, AP style provides for uniformity and consistency in journalistic writing.

The push for uniformity can be attributed to several factors. It could be that the style serves the industry more than the readers. Perhaps it acts as a cue of professionalism and industry knowledge; someone applying for a journalism job without AP style knowledge, for example, could be seen as an outsider. An article directed at public relations practitioners suggests this possibility: “Too many writers violate the AP style rules, and each style error marks the writer as an amateur. When I address groups of public relations writers, I am sometimes shocked that many have no idea what the *Associated Press Stylebook* is. Editors are not impressed by writers who cannot use the correct style” (Ryan, 1995, p. 25).

On the other hand, it could be that uniformity helps readers focus on content, rather than style. Maybe the consistency helps readers to more easily read the content and, therefore, to better understand it. These potential effects are explored below.
**Inverted pyramid structure.** The *inverted pyramid structure* is a traditional news media structure in which information is presented in order of descending importance. The most important information is listed at the top, in the form of an attention-grabbing introduction, or *lede*, and the information gets increasingly less important as the article progresses. This distinguishes journalistic writing from other forms of writing, where information might be structured narratively or chronologically. For the purposes of this study, then, ledes and inverted pyramid represent the journalistic convention of *structure*.

This structure is often said to have emerged from the advent of the telegraph, the contemporary technology used to transmit communications. However, Mindich (1993) offers a compelling alternative in which he places the first inverted pyramid style story in the 1860s, before the advent of the telegraph. Mindich found that the oft-cited first inverted pyramid story, which announced Abraham Lincoln’s assassination in 1865, was actually written by a government official, rather than a journalist. The official was Edwin M. Stanton, Lincoln’s secretary of war. Mindich explains that Stanton had significant control over information and that his press releases were often published unedited in newspapers. Thus, the first inverted pyramid style news article was published because of the first inverted pyramid style press release. Mindich comments on this irony: “It is strangely fitting that the author of what might be the first inverted pyramid was also one of the most notorious censors of the nineteenth century press” (Mindich, 1993, p. 11).

Why did this become the formal structural convention for journalistic writing? Mindich says Stanton began writing in this style partly for circumstantial reasons (i.e., he was busy, and it was quicker to write this way) and partly for purposeful reasons: by
keeping the “important” information at the top, he could ensure that his message remained, even if editors had limited space. Mindich further explains this notion, and this explanation is often employed by journalism professors as a reason to write in this style: “Because it has no narrative flow, the inverted pyramid is easier to manipulate than the chronological account; it is easy to delete paragraphs and facts without disturbing the sense of the story” (Mindich, 1993, p. 20). Unlike journalism professors, however, Mindich emphasizes that this manipulation benefits authority; by putting the “most important” information first, journalists ensure that the “official” information stays, regardless of other cuts to the article.

Whether this be the case or not, it has been assumed that the inverted pyramid structure has an effect on readers. The structure allows readers to skim information and, perhaps, to learn more with less engagement. In this way, articles written in the inverted pyramid structure could lead to more knowledge gain than those that are not. Maybe the structure helps readers to gain useful knowledge despite their tendency to process superficially. This study tests that possibility.

**Cognitive Processing of Journalistic Writing Conventions**

To determine how these writing conventions affect readers, this study looks at how they are cognitively processed. Do readers notice mistakes and then consciously decide not to trust the content? Or do mistakes serve as distractions that unconsciously inhibit understanding? This study empirically tests this concern through the framework of dual-processing models of persuasion.

Two of the more common models — the Elaboration Likelihood Model and the Heuristic-Systematic Processing Model — conceptualize information processing in
similar ways. After exploring these two theories, this study explains how journalistic
writing conventions will be studied within their framework.

**Elaboration Likelihood Model.** The Elaboration Likelihood Model, developed
by Richard Petty and John Cacioppo (Petty & Cacioppo, 1986a, 1986b), is a dual-
processing model of social judgment. The model hypothesizes that people process
information by one of two routes: the central route and the peripheral route. The central-
route to attitude change is more deep, or “based on relatively extensive and effortful
information-processing activity, aimed at scrutinizing and uncovering the central merits
of the issue or advocacy” (Petty & Wegener, 1999, p. 42). The peripheral-route to attitude
change, on the other hand, is more superficial, or “based on a variety of attitude change
processes that typically require less cognitive effort” (p. 42). The model suggests an
elaboration continuum whereby motivation and ability determine which route a person
follows. According to Petty and Wegener, this continuum “implies that the type of
thought given to object-relevant information can be the same under high- and low-
elaboration conditions, but that the amount of cognitive activity varies” (Petty &
Wegener, 1999, p. 46).

Seven formal postulates were then established, two of which are of particular
relevance for this analysis (Petty & Wegener, 1999, p. 44-63). “The Elaboration
Continuum Postulate” posits: “Although people want to hold correct attitudes, the amount
and nature of issue-relevant elaboration in which they are willing or able to engage to
evaluate a message vary with individual and situational factors.” These “individual and
situational factors” most often refer to motivation, which will be manipulated in this
analysis. In “The Multiple-Roles Postulate,” they explain that variables are not inherently
processed one way or the other, but that “any one variable can have an impact on attitude change by more than one mechanism” (Petty & Wegener, 1999, p. 48). In other words, not all source cues are peripheral, and not all message cues are central; it depends on whether the variable is serving as an argument or as a cue within a particular context.

**Heuristic-Systematic Processing Model.** The Heuristic-Systematic Processing Model, developed by Shelly Chaiken and colleagues (Chaiken, 1980; Chaiken, Liberman, & Eagly, 1989) is also a dual-processing model of social judgment and builds upon the Elaboration Likelihood Model. The foundational notion of the theory is that people process information heuristically or systematically. Systematic processing, like central-route processing, refers to “a relatively analytic and comprehensive treatment of judgment-relevant information” that are “responsive to the actual content of this information” (Chen & Chaiken, 1999, p. 74). Heuristic processing, like peripheral-route processing, refers to more surface-level processing based on “activation and application of judgmental rules or ‘heuristics’ that, like other knowledge structures, are presumed to be learned and stored in memory” (Chen & Chaiken, 1999, p. 74). As is the case with the Elaboration Likelihood Model, researchers manipulate or induce a processing path and then test outcome variables to confirm whether participants used that processing path.

The Heuristic-Systematic Processing Model adds more nuance to the notion of dual-processing by arguing that the two types of processing can occur simultaneously. It outlines the possible combinations of this co-processing in three hypotheses: The additivity hypothesis (both “exert independent and judgmentally consistent effects”), the bias hypothesis (“judgmental implications of heuristic cue information may establish expectancies about subsequently encountered judgment-relevant information, which may
then bias the nature of more effortful systemic processing of this information”), and the attenuation hypothesis (“judgmental implications derived from systematic processing may contradict and thus attenuate the judgmental impact of heuristic processing”) (Chen & Chaiken, 1999, p. 75-76). Outcome measures, then, might vary depending on the combination of heuristic and systematic processing.

Writing conventions might be processed heuristically, and they might be processed systematically. As mentioned in the Introduction, Sparks and Areni’s (2008) analysis found support for language power both as a peripheral cue and as a distraction. Because it is not clear how writing conventions are processed, it is also not clear whether they are processed in only one way. Like language power, they could be processed in multiple ways. This study, therefore, uses the Heuristic-Systematic Processing Model because it allows for the possibility of errors being processed in multiple ways.

Journalistic Writing Conventions in the HSM Framework

On a basic level, this study fits into the HSM framework in that message variables (i.e., convention errors) are being manipulated and processing type (i.e., heuristic or systematic) is being inferred. This study follows this framework with a few variations. First, news will be used as the context. Second, media judgments will serve as the outcome measures. Third, and most importantly, errors will be viewed as potentially both heuristics and ability constraints, to test which explanation of their effects is more convincing. These variations are explained below.

News as context. For this study, journalistic writing serves as the context for dual-processing model analysis. These models are usually used in studies of advertising and persuasion, but they are useful for other studies of communication, including
journalism. The Heuristic-Systematic Processing Model talks about two modes of persuasion “within any given judgmental context … by which perceivers may determine their attitudes and other social judgments” (Chen & Chaiken, 1999, p. 74). Similarly, the Elaboration Likelihood Model talks about understanding “how any external or internal variable has an impact of some evaluative (e.g., good-bad) or nonevaluative (e.g., likely-unlikely) judgment,” (Petty & Wegener, 1999, p. 42). In both cases, this notion of “any given judgmental context” or “any external or internal variable” opens the theories to broader application.

In addition, journalism can be seen as a form of persuasion. Rather than persuading you to think a certain way about a product, journalism inherently persuades you to think a certain way about information. By its nature, journalism persuades people to believe that information is important, credible, and accurate. Within that context, it makes complete sense to use the framework of dual-processing models for studies of news media.

Finally, dual-processing models already look specifically at message quality and argument quality. There is no reason to believe that quality within advertisements would behave differently than quality within journalistic writing. O’Keefe (2013) specifically suggested that we need more studies of message factors: “There may be additional quality-related message characteristics—beyond outcome desirability—that enhance message persuasiveness under conditions of high elaboration. Identification of such message properties would represent an important advance in the understanding of persuasion generally and argument quality specifically” (p. 144). Mongeau and Stiff (1993) commented on this need as well. They argued that the relationship between
argument quality and message cognitions is the most understudied part of the dual-processing models: “Over a decade of testing the ELM has generated very little knowledge of the role of message variables in the persuasion process” (p. 69). Language cues represent one example of these “quality-related message characteristics” that would be useful for theoretical advancement. Therefore, this is not only an appropriate application, but also a useful one in terms of theory advancement.

Other researchers have applied dual-processing models of persuasion to non-persuasion contexts, as well, so this type of application is not new. Zuckerman and Chaiken (1998), for example, applied the Heuristic-Systematic Processing Model to the study of product warning labels. Edwards and Edwards (2013) applied the model to the study of instructor reviews, in terms of their effect on students’ perceptions. Riddle (2010) talks about the model as an explanation for cultivation effects (i.e., heuristics cue perceptions of social reality). Another study applied the Elaboration Likelihood Model to a credibility study of online informational diet and nutrition messages (Jung, 2011).

Some studies have specifically applied the models to a news context, as is done here. Kim and Paek (2008), for example, used the Heuristic-Systematic Processing Model to explain that the effect of processing on attitudes about news messages was moderated by motives. Waheed and Chung (2008) used the Heuristic-Systematic Processing Model to test information processing of political messages, in the context of coverage of candidates. Smith and Smith (2009) conducted a similar analysis, in which they used the Elaboration Likelihood Model to study the effects of race and gender cues on perceptions of political candidate coverage. Winter and Krämer (2012) applied the Elaboration Likelihood Model to a study of the factors that influence readers’ selection of blog
content. Liu (2008) talked about the Elaboration Likelihood Model as an explanation for agenda-setting theory (i.e., heuristics cue perceptions of importance). Berry, Wharf-Higgins, and Naylor (2007) use the Elaboration Likelihood Model to explain news coverage of health information. Thus, the application of dual-processing models to a news context is not only logical and justified; it also has precedent.

**Media judgments as outcomes.** Studies have found relationships between journalistic writing conventions and a variety of outcome variables. This study focuses on four judgments: recall, writing quality, message credibility, and informativeness. A number of others could be explored, as well (e.g., representativeness, liking). However, this study is particularly interested in how journalistic writing conventions impede or aid in knowledge gain. Knowledge gain is often considered to be one of the primary outcomes of journalism; these judgments represent features that would most closely relate to it.

These four judgments are discussed in terms of how processing should or should not affect them. General language research is inconclusive, so this study looks at the possibility that conventions are processed differently depending on what outcome is being tested. It argues that errors will serve as ability constraints when predicting recall, and errors will cue a *writing quality* heuristic when predicting message credibility and informativeness.

**Errors as ability constraints.** Dual-processing analyses test a variety of potential ability constraints, such as time. For this study, the ability constraint will be errors, or, more specifically, errors in journalistic writing conventions. This framework tests the possibility that errors affect readers because they serve as ability constraints, or
inhibitors. Maybe errors prevent readers from comprehending the text. Errors, then, might inhibit comprehension.

Several language complexity studies talk about language processing in a way similar to ability constraints. This line of research defines language complexity as the degree to which text is comprehensible. In other words, complex language is seen as a barrier to comprehension. Reilly and Richey (2011), for example, found that increased language complexity, operationalized as more obscure and legalistic language, on ballots leads to more roll-off from participants. They said this is because voters will not answer questions they do not understand. In this way, language harms comprehension. Wolf, Davis, Shrank, Rapp, Bass, Connor, Clayman, and Parker (2007) found similar results in a study of medical terminology. They found that patients commonly misunderstand the drug label instructions and that this is largely due to awkward or unnecessarily complicated language. They concluded that “prescription drug labels should use explicit dosing intervals, clear and simple language, within a patient-friendly label format” (p. 293). These studies support the notion of errors as inhibitors.

Recall. Recall is a measure of the information the reader remembers from the article. It can be measured by asking content-specific questions about the stimulus article. Appelman and Bolls (2011) found that grammatical errors in news articles lead to longer reading time and lower retention. Why would participants take more time to read but remember less?

A study on language complexity suggests this is based on errors as inhibitors. Van Weert, van Noort, Bol, van Dijk, Tates, and Jansen (2011) looked at the effect of complex language on recall. In the context of medical websites, they found that complex
language (i.e., medical jargon) led to significantly less recall than non-complex language (i.e., informal language). Interestingly, language complexity did not affect readers’ self-report measure of satisfaction with comprehensibility (e.g., “the website is easy to read,” “the language used on the website is easy to understand”); it only affected their actual recall scores. In other words, people do not think that complex language affects comprehension, even though it does. Language complexities, thus, seem to act as inhibitors in their effect on recall.

Journalistic writing-convention errors should act as complexities; therefore, they, too, should act as inhibitors in their effect on recall. This study, thus, predicts that errors affect recall by inhibiting processing.

**Errors as heuristic cues.** The other possibility is that errors serve, not as ability constraints, but as heuristic cues. This framework tests the possibility that errors affect readers because they trigger heuristic cues. Dual-processing analyses test a variety of potential heuristics, such as source gender or appearance. This study predicts that errors in journalistic writing conventions will trigger a writing quality heuristic. Maybe readers see errors as cues of poor writing quality; an error in style, then, might be perceived as a cue that the content cannot be trusted to be accurate.

Howard (1997), for example, used the Elaboration Likelihood Model and found that familiar idiomatic phrases have a stronger effect on attitude formation than unfamiliar literal phrases. He further showed that this was especially true for participants in low involvement and distracted conditions and for people with low need for cognition. He argued that familiar phrases are more quickly and easily understood and that this supports the “expectation of peripheral-cue effectiveness” (p. 231).
Several neuroscience studies talk about language as heuristic, rather than systematic, in its effect. Shtyrov, Pulvermüller, Näätänen, and Ilmoniemi (2003), for example, found that grammar processing takes place automatically, or, in other words, “early syntax processing in the human brain may take place outside the focus of attention” (p. 1195). Aaltonen, Hellström, Peltola, Savela, Tamminen, and Lehtola (2008) found that language violations trigger “the brain’s automatic change-detection response” (p. 56). Pulvermüller, Shtyrov, Hasting, and Carlyon (2008) similarly found that, even under significant distraction, participants can distinguish between correct and incorrect language. These studies support the notion that language can trigger responses in a more automatic manner.

This study argues that this happens because language errors cue the writing quality heuristic. Writing quality is a message perception that refers to “the degree or level of overall excellence of a news story” (Sundar, 1999, p. 381). Research shows that message errors directly affect perceptions of writing quality. For example, Appelman and Schmierbach (2014) found that grammatical errors in news articles lead to lower perceptions of article quality. Related research suggests that writing quality can, in turn, influence other judgments, such as message credibility and informativeness. Appelman and Sundar (forthcoming), for example, found that in addition to professional expertise and fairness, the variable most likely to affect message credibility perceptions was writing quality. Thus, this study looks at writing quality as the heuristic that is cued by errors and that, in turn, influences the following perceptions.

Message credibility. Credibility is a “global evaluation of the objectivity of the story” (Sundar, 1999, p. 380). The concept of credibility can be further conceptualized as
three separate concepts: source credibility, message credibility, and media credibility (Metzger, Flanagan, Eyal, Lemus, & Mccann, 2000). Because this study’s focus is on language cues, it focuses on message credibility as the relevant media credibility judgment. *Message credibility* can be defined as follows: “An individual’s judgment of the content of communication” (Appelman & Sundar, forthcoming). It can be measured by asking participants to rate how accurate, believable, and authentic they perceive the content to be (Appelman & Sundar, forthcoming).

Incorrect grammar affects perceived credibility of speakers (Grazian, 1996), job candidates (Grazian, 1996; Charney & Rayman, 1989), and email writers (McGoon, 1996). Sundar (1998) found that articles with quotes are perceived to be more credible than stories without quotes. In addition, anonymous sources have been shown to lead to lower perceived credibility than named sources (Sternadori & Thorson, 2009). Similarly, Appelman and Bolls (2011) found that grammatical errors in news articles lead to lower perceptions of credibility. Thus, overall, errors affect credibility perceptions. Why is this the case?

Several of the aforementioned dual-processing studies look at perceived credibility as an outcome measure affected by heuristic cues (e.g., Jung, 2011; Winter & Krämer, 2012). Other studies talk about credibility as an outcome measure of message-based heuristic cues. For example, message errors can cue perceptions of professionalism, which can, in turn, influence credibility perceptions. Perceived professionalism was specifically mentioned by participants as a factor that influences credibility in a set of focus groups (Appelman & Sundar, forthcoming). Several participants suggested that lack of perceived professionalism affects perceptions. For
example, when asked how they would make a news outlet credible, one participant said to “keep the way that you present your information professional.” Dochterman and Stamp (2010) conducted focus groups that yielded similar findings. Wathen and Burkell (2002) made a similar conclusion in the context of online media: “For Internet information, even a single spelling mistake can give the impression of ‘amateurism’ and lead the user to reject the site as not credible” (p. 140). In all, then, evidence suggests that errors cue heuristics, such as professionalism, which, in turn, influence credibility. This study thus predicts that errors affect message credibility by serving as heuristic cues.

Informativeness. Informativeness is the perception of the article’s ability to communicate information, or knowledge (Mutz & Reeves, 2005). Research suggests that message errors might affect perceptions of informativeness. Appelman and Schmierbach (2014), for example, predicted that grammatical errors would lead to lower perceived informativeness of news articles. The relationship was not significant, but the data did suggest a pattern that warrants further testing. If that pattern proves replicable and errors do, in fact, negatively affect informativeness, why would this be the case?

Few studies look at informativeness, but it has been discussed in ways similar to heuristic processing. Emotion, as discussed by Exemplification Theory, talks about informativeness as an outcome measure of message-based heuristic cues, such that message errors can cue emotion, which can, in turn, influence perceived informativeness. Exemplification Theory explains the use of exemplars in explaining news events, rather than relying solely on facts, or “base-rate information” (Zillmann, Gibson, Sundar, & Perkins, 1996). Zillmann (1999) explains the innate preference for exemplars: “Recipients give disproportional attention to concrete, often vividly displayed events,
especially to those that engage the recipients’ emotions, and that this attentional preference comes at the expense of attention to more abstract, comparatively pallidly presented information” (p. 70).

Sundar (2003) explains that these exemplars affect perceived learning because they evoke or convey emotions. In this way, maybe emotion mediates the relationship between journalistic writing conventions and informativeness perceptions. This study thus predicts that errors affect informativeness by serving as heuristic cues.

**Motivation manipulation as analysis.** Are journalistic writing conventions processed as ability constraints or as heuristic cues? This study is interested in whether media judgments are based on “errors as inhibitor” or “errors as heuristic cues.” As explained above, both paths are possible, and research exists that support both options. Based on previous research, this study predicts that recall will be based on errors as inhibitors, and that errors will cue a writing quality heuristic that will influence message credibility and informativeness. How will these differences be tested?

The traditional design of dual-processing model studies is experimental. Researchers manipulate or induce a processing path and then test outcome variables to confirm whether participants used that processing path. Petty and Wegener (1999) outline the main ways in which processing paths are induced: source and message variables are both manipulated, “to operationalize high-versus low-elaboration attitude change” (e.g., Heppner et al., 1995); message variables are manipulated, “to show their role as peripheral cues” (e.g., Braverman, 2008); and source variables are manipulated, to show “their role as arguments” (e.g., Di Blasio & Milani, 2008). Studies often induce or measure participants’ involvement, need for cognition, or motivation as moderators.
This study suggests the paths for this study can be analyzed through the following method. If errors matter because they are inhibitors, then motivation will not moderate the effect of errors on media judgments. If errors are inhibitors, then it should not matter how motivated a participant is in the study — it should impede processing either way. In other words, there should be a main effect of error on judgments.

On the other hand, if errors matter because they are heuristic cues, then errors will make more of a difference in low-motivation conditions, as compared with high-motivation conditions. The Heuristic-Systematic Processing Model says heuristic cues only affect participants when their motivation is low; when motivation is high, they are processing systematically and overlooking heuristic cues. Thus, if errors are heuristics, they will only affect people who have low motivation. In other words, there should be an interaction effect of error and motivation on judgments.

Thus, errors should have main effects on recall and writing quality, and they should interact with motivation in affecting message credibility and informativeness. Based on this method of measurement, and the aforementioned predictions, this study suggests the following hypotheses:

\( H1 \): There will be a main effect of errors on recall, such that articles with errors will be associated with lower recall than articles without errors.

\( H2 \): There will be a main effect of errors on perceived writing quality, such that articles with errors will be associated with lower writing quality than articles without errors.

\( H3 \): Motivation will moderate the effect of errors on perceived message credibility, such that articles with errors will only be associated with low message credibility when motivation is low.

\( H4 \): Motivation will moderate the effect of errors on perceived informativeness, such that articles with errors will only be associated with low informativeness when motivation is low.
Message factors as processing mechanisms. This study also predicts that errors will have a main effect on perceptions of overall news quality. In addition to the variables above (i.e., recall, writing quality, message credibility, informativeness), several variables could mediate this relationship: enjoyment, newsworthiness, source-related message credibility, source credibility, fairness, professionalism, familiarity, and emotion.

Maybe, for example, errors lower enjoyment, which, in turn, lowers overall quality judgments. In addition, maybe perceptions of newsworthiness mediate this relationship. Maybe articles with journalistic writing errors are seen to be less newsworthy than those without errors, which, in turn, leads to lower media perceptions. Similarly, maybe source factors, such as perceived source credibility or source quality, mediate this relationship. Maybe articles are seen as less fair when they have errors; maybe mistakes signal a lack of objectivity, which leads to negative judgments. Similarly, perceived professionalism could mediate the effects of writing conventions on overall judgments (e.g., Appelman & Sundar, forthcoming; Dochterman & Stamp, 2010; Wathen & Burkell, 2002).

Familiarity also could mediate the effects of writing conventions on judgments. Maybe readers respond positively to journalistic writing conventions purely out of routine. Maybe they expect to see content written in this way and it becomes a distraction when written differently. Dochterman and Stamp (2010) talk about familiarity in terms of familiarity with a website, rather than with writing conventions specifically, but their findings are relevant here, as well. Participants said they have more positive judgments of websites they have already visited than of new websites. Maybe people also have more positive judgments of news content that looks like other news content they have seen.
before; if they read content in traditional styles and structure, then maybe they like that content better. Finally, emotion could mediate the relationship (e.g., Sundar, 2003).

These possibilities are addressed through the final analyses:

*H5:* There will be a main effect of errors on overall news quality, such that articles with errors will be associated with lower quality than articles without errors.

*RQ1:* Are the effects of journalistic writing conventions on perceptions of overall news quality mediated by (a) enjoyment, (b) newsworthiness, (c) source-related message credibility, (d) source credibility, (e) fairness, (f) professionalism, (g) familiarity, or (h) emotion?
Chapter 3

Methods

Design

To test these predictions, this study takes the form of a 4 (writing-convention errors) x 2 (motivation) between-subjects experiment. Participants were randomly assigned to one of the error conditions (grammatical errors, \( n = 122 \); AP style errors, \( n = 126 \); inverted pyramid structure errors, \( n = 129 \); no errors, \( n = 127 \)) and one of the two motivation conditions (low, \( n = 251 \); high, \( n = 253 \)). They then read an article in their condition and answered questions to determine media judgments and infer processing.

Participants

Participants (\( N = 504 \)) were recruited from Amazon Mechanical Turk (MTurk). All participants saw the stimulus article, so all were retained for analyses.

Sample-size sufficiency. Based on an online effect size calculator (Becker, 1998), previous experiments looking at the effects of micro-editing errors have yielded relatively small effect sizes: \( r = .21 \) (Appelman & Bolls, 2011) and \( r = .26 \) (Appelman & Schmierbach, 2014). For an eight-condition experiment with a predicted effect size of \( r = .25 \) (i.e., comparable to these previous studies), G*Power (Faul, Erdfelder, Lang, & Buchner, 2007) recommends 400 participants; therefore, this sample size was acceptable. Similarly, Cohen (1992) says a sample with \( N > 464 \) should be able to detect what it deems to be a “small effect size” at \( \alpha = .01 \). Additionally, Holbert and Stephenson (2002) recommend a minimum sample of 150 for modeling analyses. In all, then, this study’s sample size is more than sufficient to allow for path analysis and detection of even small-sized effects.
MTurk as a sampling source. MTurk is an online forum, or “crowdsourcing web service” that connects researchers with participants (Paolacci, Chandler, & Ipeirotis, 2010, p. 411). According to their site, it is “a marketplace for work that requires human intelligence” (“Amazon Mechanical Turk FAQ,” n.d.). The site provides a forum by which researchers (i.e., requesters) can post studies (i.e., HITS — Human Intelligence Tasks) for participants (i.e., workers) to complete. Requesters then pay workers through Amazon, which collects 10 percent commission for the service. These tasks are “typically completed within minutes and usually pay in cents rather than dollars” (Paolacci & Chandler, 2014, p. 184). The forum markets itself as mutually beneficial: “The Mechanical Turk service gives businesses access to a diverse, on-demand, scalable workforce and gives Workers a selection of thousands of tasks to complete whenever it's convenient (“Amazon Mechanical Turk FAQ,” n.d.). The forum’s name refers to the notion of humans doing automated work: “Mechanical Turk is named after an 18th century chess playing ‘automaton’ that was in fact operated by a concealed person” (Paolacci, Chandler, & Ipeirotis, 2010, p. 411).

MTurk has become a popular online sampling source for social scientists because it provides large samples of motivated participants at low cost (Paolacci & Chandler, 2014). It is particularly relevant for studies of online news media because, by nature of the sampling technique, the participants are tech-savvy and frequent Internet users. Demographic analysis shows that MTurk samples are as representative of the U.S. population as other samples used in social science research, “with gender, race, age and education of Internet samples all matching the population more closely than college
undergraduate samples and internet samples in general” (Paolacci, Chandler, & Ipeirotis, 2010, p. 414).

Demographics were measured as possible control variables in the study. All such measures had multiple-choice responses, so as to elicit uniform, analyzable results. Participants ranged in age from 18 to 74 ($M = 34.55, SD = 11.07$). There were more men ($n = 292$) than women ($n = 211$). Most were White ($n = 411$), with relatively equal numbers of Black ($n = 34$), Hispanic/Latin(o/a) ($n = 35$) and Asian ($n = 33$). Some identified as Native-American ($n = 6$) or Other ($n = 6$). Some who marked “Other” wrote in responses, including Brazilian, Caribbean American, mixed and White/Hispanic. Participants represented a range of geographic locations in the U.S.: South ($n = 205$), Northeast ($n = 117$), Midwest ($n = 104$), and West ($n = 78$).

All “workers” who had completed the researcher’s “hits” in the past 30 days were removed from analysis, to ensure there were no pilot test participants in the final study. Requesters can set their own standard for payment, but, according to a blog for MTurk requesters, payment should be between 10 cents and 30 cents per minute of work (“Tips for academic requesters,” 2012). For this study, then, all participants were compensated $1 for their participation, and those with perfect recall scores earned an additional 50 cents. The additional money was provided through the bonus option in MTurk. This approach is discussed further in the Motivation section.

**Stimulus**

The stimulus articles for this study are science-themed news articles from The New York Times. A science-based article is ideal because it offers the possibility for learning, which is one of the core outcome measures of this study. In addition, science
articles are fairly complicated, so deep processing could conceivably affect understanding. The original stimulus articles were modified to be relatively comparable in terms of length and complexity. For the sake of parsimony, they are referred to in the following sections by their main topics: Space (Chang, 2014), Brain (Gorman, 2014), Antibiotic (Rabin, 2014), and Coral (Alvarez, 2014).

Each article was manipulated to be 12 paragraphs and 430 to 450 words. They also were manipulated to have Flesch Reading Ease scores of 30 to 50 in Microsoft Word. The Flesch Reading Ease score is a 100-point scale, calculated based on average sentence length and average number of syllables per word, where high values mean the content is easier to understand (“Test your document’s readability,” n.d.). The scores for articles in this study are relatively low, which means the articles are relatively complex. The complexity and length modifications made two articles (i.e., Space and Coral) more difficult to understand, while it kept the others (i.e., Brain and Antibiotic) consistent with the original. Original and modified readability scores are shown in Table 1. The stimuli for the control group are listed in Appendix A, with their corresponding complexity and length values.

Table 1
Stimulus Article Readability Scores

<table>
<thead>
<tr>
<th>Article</th>
<th>Version</th>
<th>Word Count</th>
<th>Flesch Reading Ease</th>
<th>Flesch-Kincaid Grade Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Space</td>
<td>Original</td>
<td>510</td>
<td>45.4</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>Modified</td>
<td>453</td>
<td>41.0</td>
<td>12</td>
</tr>
<tr>
<td>Brain</td>
<td>Original</td>
<td>449</td>
<td>40.4</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>Modified</td>
<td>448</td>
<td>40.3</td>
<td>12</td>
</tr>
<tr>
<td>Antibiotic</td>
<td>Original</td>
<td>506</td>
<td>29.3</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>Modified</td>
<td>438</td>
<td>30.0</td>
<td>12</td>
</tr>
<tr>
<td>Coral</td>
<td>Original</td>
<td>658</td>
<td>50.7</td>
<td>11.9</td>
</tr>
<tr>
<td></td>
<td>Modified</td>
<td>441</td>
<td>47.0</td>
<td>12</td>
</tr>
</tbody>
</table>

*Note:* Flesch Reading Ease scores range from 1-100, with high values indicating less complex content (Test, n.d.).

Articles listed based on topics: Space (Chang, 2014), Brain (Gorman, 2014), Antibiotic (Rabin, 2014) and Coral (Alvarez, 2014).
Manipulated Variables

Writing-convention errors. The study’s core independent variable is writing-convention errors. For this study, these errors are conceptualized as grammatical errors, style errors, and structure errors. One condition has no errors, to serve as a control group. As discussed in the literature review, grammar errors are deviations from formal standards of usage, spelling and punctuation; style errors are deviations from Associated Press style; and structure errors are deviations from inverted pyramid structure. The three writing-convention errors (i.e., grammar, style, and structure) were written into the articles by the researcher, based on results of the following two pilot tests.

Pilot Test 1. To determine which errors to include in the experiment, journalists (N = 39) were recruited and asked to rank style errors. Of that group, 29 were currently employed as journalists (range = 1 to 42 years, M = 15.54, Mdn = 10, SD = 13.95, two students), nine had been employed as journalists in the past (range = 5 to 25 years, \( M = 13.88, Mdn = 13, SD = 8.47 \)), and one was a student not working in journalism.

First, they were given a list of 25 grammatical errors (e.g., incorrect subject-verb agreement), with examples in parentheses (e.g., they is), and asked to select the five that were the most problematic. Similarly, they were given a list of 25 AP style errors (e.g., incorrect construction of p.m. and a.m.) and asked to select the five that were the most problematic. For both sets of questions, the order of options was randomized by participant, and an “other,” write-in option was given. Appendix B and Appendix C show the full lists of grammatical errors and AP style errors, respectively. Additionally, participants were each given one of the four stimulus articles with the paragraph order
randomized, and they were asked to rearrange the paragraphs to be in descending order of importance (i.e., in inverted pyramid structure).

For the grammar question, participants selected six of the errors more than the others: incorrect subject-verb agreement \( (n = 26) \), incorrect spelling of proper nouns \( (n = 24) \), confused homophones \( (n = 24) \), confusion of possessive and plural cases of nouns \( (n = 17) \), incorrect spelling of common nouns \( (n = 16) \), and incorrect pronoun-antecedent agreement \( (n = 15) \).

For the AP style question, abbreviation-related responses were, collectively, the most popular (confusion between postal code and AP state abbreviations, \( n = 13 \); confusion between spelling out proper names and abbreviations, \( n = 11 \); and incorrect abbreviations, \( n = 11 \)). This was followed by placing commas and periods outside of quotation marks \( (n = 22) \), confusion between spelling out numbers and using numerals \( (n = 17) \), incorrect capitalization of formal titles \( (n = 13) \), adding a serial comma \( (n = 12) \), and using \textit{yesterday/today/tomorrow} instead of days of the week \( (n = 11) \).

For the inverted pyramid questions, the results showed relative agreement across participants. The majority of participants identified a specific paragraph as the lede. For the Space article \( (n = 7) \), most \( (n = 5) \) said the first paragraph should be, “Instead of splashing into the Pacific Ocean in 2020 as planned … .” For the Brain article \( (n = 8) \), most \( (n = 6) \) said the first paragraph should be, “The National Institutes of Health set an ambitious $4.5 billion price tag … .” For the Antibiotic article \( (n = 17) \), most \( (n = 13) \) said the first paragraph should be, “A single infusion of an antibiotic can clear serious bacterial skin infections … .” For the Coral article \( (n = 7) \), most \( (n = 4) \) said the first paragraph should be, “With many of the reefs in South Florida faltering … .”
Based on the first pilot test, then, the treatment conditions were constructed as follows. The articles in the grammatical error condition were each given 10 grammatical errors from the list of six (i.e., incorrect subject-verb agreement, incorrect spelling of proper nouns, confused homophones, confusion of possessive and plural cases of nouns, incorrect spelling of common nouns, and incorrect pronoun-antecedent agreement), with each of the six present at least once. The articles in the AP style error condition were each given 10 AP style mistakes from the list of six (i.e., abbreviation errors, placing commas and periods outside of quotation marks, confusion between spelling out numbers and using numerals, incorrect capitalization of formal titles, adding a serial comma, and using yesterday/today/tomorrow instead of days of the week). Each of the six were present when possible (Note: there was no serial comma in the Brain or Antibiotic articles because the original articles did not contain lists.). For the grammar and AP style conditions, no mistakes were added to the headline or quotes, so as to ensure that participants attributed errors to reporters and editors, not to designers or sources. The articles in the inverted pyramid structure error condition were written such that the two paragraphs deemed most important were placed within the article, rather than at the beginning of the article (i.e., what journalists call “burying the lede”).

These manipulated articles are attached as Appendix D. The added mistakes are highlighted here for clarity; in the actual study, the mistakes were not identified.

**Pilot Test 2.** A second pilot test was conducted to confirm the manipulations of journalistic writing conventions and to identify potential confounding variables. Participants ($N = 111$) were recruited from Amazon Mechanical Turk. They were each assigned to read all four articles, but the versions they saw were randomized (i.e., one
participant might have seen the Space article in the grammar condition, the Brain article in the AP condition, the Antibiotic article in the grammar condition, and the Coral article in the control condition). For each article, they were asked to indicate their agreement with three statements regarding the manipulation: This article has no obvious grammatical errors, This article has no obvious AP style errors, and This article was written from most important to least important. These measures were written by the researcher and were measured on a five-point scale from $I = \text{strongly disagree}$ to $5 = \text{strongly agree}$.

*Enjoyment as a possible confound.* Participants also were asked to indicate their agreement with three statements regarding the potential confounding variable of enjoyment: This article was enjoyable, This article was boring (reversed), and This article was entertaining (Oliver, Weaver, & Sargent, 2000), also on a scale from $I = \text{strongly disagree}$ to $5 = \text{strongly agree}$. An enjoyment scale was created for each article as the average of three. For each story, the reliability of this scale was well above .80 (Space, $\alpha = .90$; Brain, $\alpha = .91$; Antibiotic, $\alpha = .87$; Coral, $\alpha = .87$). Thus, an enjoyment variable was created for each article: Space ($M = 3.61, SD = .89$), Brain ($M = 2.85, SD = 1.02$), Antibiotic ($M = 3.12, SD = .94$), and Coral ($M = 3.16, SD = .94$).

Analysis showed no significant relationship between condition and enjoyment for the Space article, $F(3, 107) = .02, p = .10, \eta^2_\text{p} = .00$; the Brain article, $F(3, 107) = .52, p = .67, \eta^2_\text{p} = .01$; or the Antibiotic article, $F(3, 107) = 1.32, p = .27, \eta^2_\text{p} = .04$. For the Coral article, analysis showed significant differences in enjoyment based on condition, $F(3, 107) = 3.87, p = .01, \eta^2_\text{p} = .10$. Post hoc analysis using Tukey’s HSD showed that the version with inverted pyramid errors ($M = 3.53$) was perceived as significantly more
enjoyable than the version with the grammatical errors ($M = 2.74, p = .01$). There were no such effects with the article in the AP style error condition ($M = 3.12$) or the control condition ($M = 3.24$). Because the Coral article suggested a possible confounding effect of enjoyment, enjoyment was measured in the full study as a potential confound.

**Confirming manipulation of journalistic writing conventions.** Independent samples $t$-tests were used to confirm the manipulation of journalistic writing conventions. For the Space content, the article in the grammatical error condition was perceived as having significantly worse grammar ($M = 2.81, SD = 1.47$) than the articles in the other three conditions ($M = 3.76, SD = .90$), $t(32.5) = -3.17, p < .001, d = .78$. The article in the AP style error condition was perceived as having worse AP style ($M = 3.43, SD = 1.00$) than the articles in the other three conditions ($M = 3.69, SD = .97$); however, this difference was not significant, $t(109) = -1.21, p = .23$. The article in the inverted pyramid error condition was perceived as having worse inverted pyramid structure ($M = 3.03, SD = .94$) than the articles in the other three conditions ($M = 3.32, SD = .93$); however, this difference was not significant, $t(109) = -1.40, p = .16$.

For the Brain content, the article in the grammatical error condition was perceived as having worse grammar ($M = 3.25, SD = 1.43$) than the articles in the other conditions ($M = 3.42, SD = 1.04$); however, this difference was not significant, $t(37.05) = -.59, p = .56$. The article in the AP style error condition was perceived as having worse AP style ($M = 3.14, SD = 1.08$) than the articles in the other conditions ($M = 3.30, SD = 1.06$); however, this difference was not significant, $t(109) = -.68, p = .50$. The article in the inverted pyramid error condition was perceived as having significantly
worse inverted pyramid structure \((M = 2.70, SD = .84)\) than the articles in the other conditions \((M = 3.31, SD = .77)\), \(t(109) = -3.62, p < .001, d = .76\).

For the Antibiotic content, the article in the grammatical error condition was perceived as having significantly worse grammar \((M = 3.04, SD = 1.24)\) than the articles in the other conditions \((M = 3.74, SD = .92)\), \(t(32) = -2.63, p = .013, d = .64\). The article in the AP style error condition was perceived as having the same AP style \((M = 3.59, SD = .83)\) as the articles in the other conditions \((M = 3.57, SD = .99)\), \(t(109) = .06, p = .95\). The article in the inverted pyramid error condition was perceived as having worse inverted pyramid structure \((M = 3.45, SD = .91)\) than the other articles \((M = 3.30, SD = .89)\); however, this difference was not significant, \(t(109) = .75, p = .46\).

For the Coral content, the article in the grammatical error condition was perceived as having significantly worse grammar \((M = 2.18, SD = 1.19)\) than the articles in the other conditions \((M = 3.40, SD = 1.14)\), \(t(109) = -4.853, p < .001, d = 1.05\). The article in the AP style error condition was perceived as having the same AP style \((M = 3.18, SD = 1.12)\) as the articles in other conditions \((M = 3.13, SD = 1.14)\), \(t(109) = .19, p = .85\). The article in the inverted pyramid error condition was perceived as having worse inverted pyramid structure \((M = 2.93, SD = .90)\) than the other articles \((M = 3.12, SD = .83)\); however this difference was not significant, \(t(109) = -1.03, p = .30\).

Based on the manipulation analysis, the Antibiotic article and Coral article were considered for removal from the full study; they were the two articles for which the AP style manipulation was not directionally acceptable. The Coral article, however, had a stronger level of significance for the grammar manipulation \((d = 1.05 \text{ vs. } d = .64)\), and it had a lower \(p\) value for the inverted pyramid manipulation \((p = .30 \text{ vs. } p = .46)\).
Additionally, the Antibiotic article was the most complex. Its Flesch Reading ease score was 30, which was lower (i.e., more complex) than the other three: 40, 41, and 47. Because the significance levels were lower and the article was more complex, the Antibiotic article was deemed to be the outlier among the four. Therefore, the Coral article was retained, and the Antibiotic article was dropped.

Despite so many non-significant differences, the three selected articles – Space, Brain, and Coral – were considered appropriate for analysis. The sample size on Pilot Test 2 was relatively small for a study with four experimental conditions ($N = 111$), and most relationships were directionally acceptable; therefore, the non-significant differences were assumed to be a power issue, rather than an indication of improper manipulations. To confirm this supposition — that non-significant effects were due to power issues, rather than poor manipulations — the three selected articles were merged and considered altogether. As expected, the relationships were directionally appropriate and stronger when observed in this way. When the three selected articles were merged, the articles in the grammatical error condition were perceived as having significantly worse grammar ($M = 3.23$, $SD = .78$) than the articles in the other conditions ($M = 3.54$, $SD = .79$), $t(84) = 1.72$, $p = .09$, $d = .39$. The articles in the AP style error condition was perceived as worse AP style ($M = 3.23$, $SD = .79$) than the articles in other conditions ($M = 3.45$, $SD = .91$); however, this difference was not significant, $t(80) = 1.13$, $p = .26$. The articles in the inverted pyramid error condition were perceived as having significantly worse inverted pyramid structure ($M = 2.96$, $SD = .50$) than the other articles ($M = 3.23$, $SD = .62$), $t(80) = 2.17$, $p = .03$, $d = .48$. In other words, pilot testing of the three selected
articles showed that the manipulations were all directionally effective but that people tended to be least attentive to the AP style errors.

Participants in the core story were, thus, randomly assigned to read one of these three articles: Space \((n = 170)\), Brain \((n = 169)\), and Coral \((n = 165)\).

For the main study, the articles were designed to look like they were on a news website. They were each given the following identical heading:

*Posted: Monday, Sept. 1, 2014, 2:30 p.m.*
*From Staff Reports*

**Washington, D.C. —**

**Motivation.** Motivation was manipulated by telling a random sample of the participants that they would receive higher compensation for a perfect recall score. After being shown the standard instructions, which told them they would read a news article and answer questions about it, this subset of participants were shown the following motivation-inducing message: “If you correctly answer all 5 of the multiple-choice recall questions, then we will pay you an extra 50 cents through MTurk.” In reality, all participants with a perfect recall score received the additional 50 cents, but knowing about the bonus ahead of time induced high motivation for a random subset of the sample.\(^1\) The logistics of this bonus are discussed in the *Recall* section.

**Construct validity.** Because this method of motivation manipulation as a measure of information-processing type was new, this study also included an established measure of systematic processing for use in construct validity testing. This measure is based on

\(^1\) It was later discovered that MTurk “workers” have several online forums (e.g., Reddit, Turkopticon) for sharing information about “hits,” which could have compromised this manipulation. This potential limitation is explored in the *Discussion* section.
Neuwirth, Frederick, and Mayo (2002). The following portion of their scale, which tests for systematic processing, was modified to be state-based rather than trait-based. Participants were asked to indicate the extent to which they agreed with five statements, from $1 = \text{strongly disagree}$ to $7 = \text{strongly agree}$: I thought about how the information in the news article related to other things I know; I found myself making connections between the information I got from the news article and information I got elsewhere; I tried to relate the ideas the news article to my own life; Based on the information I received from the news article, I thought about what actions should be taken by policy-makers; I tried to think of the practical applications of the information I got from the news article. These statements had high internal reliability ($\alpha = .85$), so the measure was constructed as an average of the five ($M = 4.50$, $SD = 1.24$).

**Measured Variables**

**Pretest.** A pretest of graduate students ($N = 16$) confirmed the clarity of the study and, in particular, the clarity of the measured variables. It confirmed that a code generator employed by the researcher accurately reflected recall scores, and it suggested the motivation manipulation be written as shown above. Additionally, several of the message and source measures originally had anchors of describes very well and describes very poorly; the non-native English speakers in the pretest separately commented that those original anchors were confusing, so they were rewritten in the language shown below.

**Processing time.** Qualtrics was set up to record the amount of time that participants spent on the page that displayed their news article. This served as a proxy for processing time. This measure was used by Appelman and Bolls (2011). The amount of time participants spent on news pages ranged from 1.32 seconds to 1,902.64 seconds.
(*M* = 152.09 seconds, *SD* = 167.04 seconds). However, the initial skewness (4.96) and kurtosis (36.66) levels were inappropriate. A box-plot flagged all values 332 seconds and above (*n* = 33) as significant outliers.² The values were, therefore, capped as 332 seconds; this translates to 5.50 minutes. This cap was data-driven, but it also makes sense as a logical maximum time to spend reading an approximately 500-word article. Once the cap was implemented, the variable had appropriate skewness (.88) and kurtosis (.22) levels, (*M* = 132.48 seconds, *SD* = 84.30 seconds).³

**Recall.** Recall was one of the hypothesized mediating variables in the study. Participants were asked five multiple-choice questions about what they remember of the content of the news article. These questions were specific to the content of the article (e.g., What NASA-related project was the focus of the article?) and, therefore, were written by the researcher. The three topics each, necessarily, required different questions. Appendix E shows the five multiple-choice questions for each of the three news articles in the full study. The correct answers are highlighted here for clarity; in the actual study, the answers were not identified. As shown in Appendix E, Space and Brain both had two correct answers for Question 3; therefore, participants who read those articles could conceivably have a recall score of six.

Qualtrics was programmed by the researcher such that correct answers were scored as 1 and incorrect answers as 0. A randomly generated code was provided for each participant at the end of the study: participants with a recall score of 5 or more were

---

² By default, SPSS flags data points in boxplots that are at least 1.5 times and 3 times the interquartile range from the other values. For each variable with normality concerns, these were the data points the researcher looked at as potential outliers.

³ Corrections for skewness and kurtosis here and throughout the Methods section allowed for more accurate reports of means and standard deviations. Although this caused the coefficients and *p*-values of relationships reported in the Results section to slightly change, it did not affect any conclusions: no insignificant relationships became significant, and no significant relationships became insignificant.
given a random six-digit number starting with 1, 2, 3, or 4; participants with a recall score of less than 5 were given a random six-digit number starting with 5, 6, 7, 8, or 9. At the conclusion of the study, participants were shown their code and asked to enter it into MTurk for compensation (Code entry is a common method of indicating study completion in MTurk.). All participants were paid $1 for having a six-digit code; those with a code that reflected a perfect recall score (n = 321) were manually given a 50 cent bonus by the researcher, in line with the motivation manipulation.

Most participants did fairly well on the recall questions (M = 4.47, SD = .82). However, the measure’s skewness (-1.51) and kurtosis (2.03) indicated non-normal distribution. One participant scored zero and was identified as a significant outlier. The participants who scored two (n = 12) were identified as lesser outliers, as well. Therefore, participants with recall scores less than three were recoded to be three; the new range of recall scores (3 to 5) resulted in appropriate skewness (1.09) and kurtosis (-.337), (M = 4.49, SD = .74).

**Writing quality.** Writing quality was one of the hypothesized mediating variables in the study. Participants were given the prompt “The content of the article was …” and asked to indicate how well each of four adjectives completed the sentence, from 1 = strongly disagree to 7 = strongly agree: coherent, clear, concise, and well-written. This measure is based on Sundar (1999). These adjectives had high internal reliability (α = .84), so the measure was constructed as an average of the four (M = 5.43, SD = .94).

**Message credibility.** Message credibility was one of the hypothesized mediating variables in the study. Participants were given the prompt “The content of the article was …” and asked to indicate how well each of three adjectives completed the sentence, from
I = strongly disagree to 7 = strongly agree: accurate, believable, and authentic. This measure is based on a measurement validity study by Appelman and Sundar (forthcoming). These adjectives had high internal reliability (α = .81), so the measure was constructed as an average of the three (M = 5.38, SD = .91).

Informativeness. Informativeness was one of the hypothesized mediating variables in the study. Participants were asked to indicate the extent to which they agreed with six statements, from I = strongly disagree to 7 = strongly agree: In general, I found the article to be informative; I learned new things about public issues from this article; This article gave me food for thought; If I needed information about public issues, I would visit this website; I felt like I got to know about the public issue by reading this article; As a result of reading this article, I am more comfortable talking to friends about this issue. This measure is based on Mutz and Reeves (2005). These statements had high internal reliability (α = .89), so the measure was constructed as an average of the six (M = 5.19, SD = 1.01). However, normality measures indicated a potential problem (skewness = -.88, kurtosis = 1.31). A box-plot flagged three values (1.00, 1.17, and 1.33) as extreme outliers. The measure was, thus, recoded such that values lower than two were recoded as two; this fixed the normality concern (skewness = -.74, kurtosis = .64), (M = 5.19, SD = .99).

Source credibility. Source credibility was one of the potential mediating variables in the study. Participants were given the prompt “The source of the article was …” and asked to indicate how well each of 18 adjectives completed the sentence, from I = strongly disagree to 7 = strongly agree: intelligent, untrained, inexpert, informed, incompetent, bright, caring (revised from “cares about me”), interested in me (revised
from “has my interest at heart”), self-centered, concerned with me, insensitive, not understanding, honest, untrustworthy, honorable, moral, unethical, and phoney.

The adjectives were meant to represent three constructs of source credibility: competence, goodwill, and trustworthiness. Once the negative items were reversed (i.e., untrained, inexpert, incompetent, self-centered, insensitive, not understanding, untrustworthy, unethical, and phoney), these adjectives had high internal reliability ($\alpha = .92$), so the measure was constructed as an average of the 18 ($M = 4.99$, $SD = .68$).

These were based on three sets of semantic differential scales created by McCroskey and Teven (1999) to represent source credibility.

**Source-related message credibility.** Source-related message credibility was one of the potential mediating variables in the study. Participants were given the prompt “The content of the article was …” and asked to indicate how well each of four adjectives completed the sentence, from $1 = strongly disagree$ to $7 = strongly agree$: authoritative, reliable, reputable, and trustworthy (Appelman & Sundar, forthcoming). These adjectives had high internal reliability ($\alpha = .80$); however, the Cronbach’s alpha level significantly increased when authoritative was removed (to $\alpha = .90$), so the measure was constructed as an average of the three ($M = 5.11$, $SD = .94$).

**Message cues.** Message cues were potential mediating variables in the study. Participants were given the prompt “The content of the article was …” and asked to indicate how well each of four adjectives completed the sentence, from $1 = strongly disagree$ to $7 = strongly agree$: fair ($M = 5.22$, $SD = 1.08$), professional ($M = 5.67$, $SD = 1.06$), familiar ($M = 3.53$, $SD = 1.51$), and emotional ($M = 2.97$, $SD = 1.53$).
The professional measure suggested normality concerns (skewness = -1.19, kurtosis = 1.96), and a box-plot flagged one significant outlier with a value of *strongly disagree* and seven with a value of *disagree*. This measure was recoded to be lower-bound at 3 (i.e., *somewhat disagree*), and the resulting measure was normally distributed (skewness = -.87, kurtosis = .58), \((M = 5.68, SD = 1.00)\).

The same anchors were used for five measures of newsworthiness: important, informative, interesting, serious, disturbing (Mayo & Leshner, 2000). These adjectives had low internal reliability \((\alpha = .57)\), but this value improved dramatically when “disturbing” was eliminated \((\alpha = .75)\), so the measure was constructed as an average of the four remaining adjectives \((M = 5.57, SD = .88)\). Normality concerns were found (skewness = -.76, kurtosis = 1.35), and a box-plot flagged one participant with a value of 1.75, as well as eight participants with values less than three. This measure was recoded to be lower-bound at 3 (i.e., *somewhat disagree*), and the resulting measure was normally distributed (skewness = -.52, kurtosis = .25), \((M = 5.58, SD = .85)\).

**Enjoyment.** Enjoyment was measured as a potential confounding variable. Participants were given the prompt “The *content* of the article was …” and asked to indicate how well each of three adjectives completed the sentence, from \(1 = strongly disagree\) to \(7 = strongly agree\): enjoyable, boring (reversed), and entertaining (Oliver, Weaver, & Sargent, 2000). These statements had high internal reliability \((\alpha = .85)\), so the measure was constructed as an average of the three \((M = 4.52, SD = 1.35)\).

**Quality.** Quality was measured as the dependent variable for structural equation modeling in this study. This measure was meant to represent a final, comprehensive judgment of the article. Participants were asked to move a slider from \(\theta = low\) to
100 = high in response to the following question: Overall, how would you rank the quality of the article? For all participants, the slider’s starting position was set at 50. This measure was created by the researcher for this study. The participants’ rankings ranged from 8 to 100, but most thought the articles were of relatively high quality ($M = 74.38$, $SD = 18.39$). Normality concerns were, again, found (skewness = -1.37, kurtosis = 1.83), and a box-plot flagged 24 people ranked the article as 40 or below; with the measure recoded to be lower-bound at 40, the resulting measure was normally distributed (skewness = -.84, kurtosis = .07), ($M = 75.44$, $SD = 15.76$).
Chapter 4

Results

Manipulation Check: Perceived Motivation

A manipulation check was intended confirmed the motivation manipulation. Participants were asked three questions on a scale of 1 = strongly disagree to 7 = strongly agree: I was motivated to carefully read the article; I paid close attention to the article; I wanted to understand the material. These measures were created by the researcher for this study.

The first measure ($M = 5.31, SD = .66$) had normality concerns, skewness = -1.47, kurtosis = 6.01. A box-plot flagged 10 outliers: one participant marked strongly disagree, two marked disagree, and seven marked somewhat disagree. The measure was, thus, recoded to be lower-bound at neutral, and the resulting measure was normally distributed, skewness = -.91, kurtosis = .58 ($M = 6.19, SD = .80$). The second measure ($M = 5.40, SD = .61$) had normality concerns, skewness = -.90, kurtosis = 1.87. A box-plot, again, showed only eight participants marked somewhat disagree, so the measure was again recoded to be lower-bound at neutral. This resulted in a normally distributed measure, skewness = -.95, kurtosis = .54 ($M = 6.28, SD = .78$). The third measure ($M = 5.33, SD = .68$) had normality concerns, skewness = -1.17, kurtosis = 2.90. A box-plot, again, flagged participants who marked disagree ($n = 2$) or somewhat disagree ($n = 9$), so the measure was again recoded to be lower-bound at neutral. This resulted in a normally distributed measure, skewness = -.98, kurtosis = .38 ($M = 6.19, SD = .86$).

The three newly capped measures had high internal reliability ($\alpha = .90$), so perceived motivation was constructed as an average of the three ($M = 5.36, SD = .55$).
An independent-samples $t$-test, however, revealed that participants in the high-motivation condition reported as much motivation ($M = 5.37, SD = .55$) as those in the low-motivation condition ($M = 5.35, SD = .55$), $t(502) = -.51, p = .61, d = 1.27$. This suggests the manipulation did not work.

As discussed in the Methods section, there were other variables that were related to perceived motivation, as well: the self-reported systematic processing measure and the record of time spent reading the article. The systematic processing measure (Neuwirth, Frederick, & Mayo, 2002) was correlated with perceived motivation, $r = .25, p < .001$, which means participants who self-reported systematic processing also self-reported high motivation. However, systematic processing was negatively correlated with time ($r = -.08, p = .07$), which suggests that people who said they were deeply processing the information actually spent less time reading the article. Analysis showed no significant relationship between the motivation manipulation, itself, and systematic processing, $t(502) = 1.82, p = .18$, or processing time, $t(502) = .57, p = .57$, which further suggests the motivation manipulation was not successful. This limitation is addressed in the Discussion section.

Because the motivation analysis, itself, did not seem to affect perceptions, these related measures – perceived motivation, systematic processing, and processing time – were used as proxies for motivation in analyses after the initial hypothesis testing.

**Descriptive Statistics**

First, the measured variables were analyzed and considered. Correlations between them are shown in Table 2, and descriptive statistics for hypothesis testing are shown in Table 3.
Table 2  
Pairwise Correlations between Measured Variables

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>.15**</td>
<td>.11*</td>
<td>.05</td>
<td>.65**</td>
<td>.61**</td>
<td>.55**</td>
<td>.04</td>
<td>.10*</td>
<td>.03</td>
<td>- .03</td>
<td>- .00</td>
<td>.64**</td>
<td>.69**</td>
</tr>
<tr>
<td>3</td>
<td>.04</td>
<td>.03</td>
<td>.04</td>
<td>.66**</td>
<td>.63**</td>
<td>.53**</td>
<td>.66**</td>
<td>.03</td>
<td>.10*</td>
<td>-.10*</td>
<td>.00</td>
<td>.14*</td>
<td>.20**</td>
</tr>
<tr>
<td>4</td>
<td>.11*</td>
<td>.05</td>
<td>.04</td>
<td>.69**</td>
<td>.63**</td>
<td>.53**</td>
<td>.66**</td>
<td>.04</td>
<td>.07</td>
<td>-.13*</td>
<td>.13*</td>
<td>.07</td>
<td>.25**</td>
</tr>
<tr>
<td>5</td>
<td>-.09</td>
<td>-.09</td>
<td>.06</td>
<td>.62**</td>
<td>.57**</td>
<td>.42**</td>
<td>.61**</td>
<td>.05</td>
<td>.07</td>
<td>-.01</td>
<td>-.01</td>
<td>.53**</td>
<td>.40**</td>
</tr>
<tr>
<td>6</td>
<td>.04</td>
<td>.04</td>
<td>.07</td>
<td>.67**</td>
<td>.65**</td>
<td>.68**</td>
<td>.63**</td>
<td>.61**</td>
<td>.41**</td>
<td>.20**</td>
<td>.17**</td>
<td>.56**</td>
<td>.56**</td>
</tr>
<tr>
<td>7</td>
<td>.01</td>
<td>.01</td>
<td>- .01</td>
<td>.53**</td>
<td>.40**</td>
<td>.57**</td>
<td>.41**</td>
<td>.36**</td>
<td>.28**</td>
<td>.27**</td>
<td>.20**</td>
<td>.29**</td>
<td>.64**</td>
</tr>
<tr>
<td>8</td>
<td>.02</td>
<td>.09*</td>
<td>.68**</td>
<td>.59**</td>
<td>.63**</td>
<td>.65**</td>
<td>.58**</td>
<td>.34**</td>
<td>.22**</td>
<td>.20**</td>
<td>.51**</td>
<td>.61**</td>
<td>.50**</td>
</tr>
</tbody>
</table>

Notes: N = 504, * p < .05, two-tailed. **p < .001, two-tailed.
1 = Processing Time. 2 = Recall. 3 = Writing Quality. 4 = Message Credibility. 5 = Informativeness.
6 = Source Credibility. 7 = Source-related Message Credibility. 8 = Fairness. 9 = Familiarity.

Table 3  
Descriptive Statistics for Perceptions by Error Type and Motivation

<table>
<thead>
<tr>
<th>Motivation</th>
<th>Error Type</th>
<th>N</th>
<th>Recall</th>
<th>Writing Quality</th>
<th>Message Credibility</th>
<th>Informativeness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>None</td>
<td>64</td>
<td>4.34(.74)</td>
<td>5.31(.94)</td>
<td>5.15(1.00)</td>
<td>4.95(1.00)</td>
</tr>
<tr>
<td></td>
<td>Grammar</td>
<td>63</td>
<td>4.40(.79)</td>
<td>5.28(1.03)</td>
<td>5.34(1.07)</td>
<td>5.28(1.02)</td>
</tr>
<tr>
<td></td>
<td>Associated Press</td>
<td>60</td>
<td>4.38(.85)</td>
<td>5.60(1.07)</td>
<td>5.49(1.01)</td>
<td>5.34(1.00)</td>
</tr>
<tr>
<td></td>
<td>Inverted Pyramid</td>
<td>64</td>
<td>4.41(.79)</td>
<td>5.39(1.09)</td>
<td>5.31(1.08)</td>
<td>5.13(1.05)</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>251</td>
<td>4.38(.79)</td>
<td>5.39(.99)</td>
<td>5.32(.97)</td>
<td>5.17(1.00)</td>
</tr>
<tr>
<td>High</td>
<td>None</td>
<td>63</td>
<td>4.63(.68)</td>
<td>5.46(.92)</td>
<td>5.42(.93)</td>
<td>5.22(1.15)</td>
</tr>
<tr>
<td></td>
<td>Grammar</td>
<td>59</td>
<td>4.56(.73)</td>
<td>5.50(1.04)</td>
<td>5.51(.95)</td>
<td>5.23(1.98)</td>
</tr>
<tr>
<td></td>
<td>Associated Press</td>
<td>66</td>
<td>4.74(.56)</td>
<td>5.39(.78)</td>
<td>5.44(.63)</td>
<td>5.26(1.95)</td>
</tr>
<tr>
<td></td>
<td>Inverted Pyramid</td>
<td>65</td>
<td>4.48(.71)</td>
<td>5.45(1.84)</td>
<td>5.40(.87)</td>
<td>5.17(1.85)</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>253</td>
<td>4.60(.67)</td>
<td>5.45(.89)</td>
<td>5.44(.84)</td>
<td>5.22(1.98)</td>
</tr>
<tr>
<td>Total</td>
<td>None</td>
<td>127</td>
<td>4.49(.72)</td>
<td>5.39(.93)</td>
<td>5.28(.97)</td>
<td>5.08(1.08)</td>
</tr>
<tr>
<td></td>
<td>Grammar</td>
<td>122</td>
<td>4.48(.76)</td>
<td>5.39(1.04)</td>
<td>5.42(.96)</td>
<td>5.25(1.99)</td>
</tr>
<tr>
<td></td>
<td>Associated Press</td>
<td>126</td>
<td>4.57(.73)</td>
<td>5.49(.83)</td>
<td>5.47(.77)</td>
<td>5.30(1.92)</td>
</tr>
<tr>
<td></td>
<td>Inverted Pyramid</td>
<td>129</td>
<td>4.44(.75)</td>
<td>5.42(.97)</td>
<td>5.35(.92)</td>
<td>5.15(1.95)</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>504</td>
<td>4.49(.74)</td>
<td>5.42(.94)</td>
<td>5.38(.91)</td>
<td>5.19(1.99)</td>
</tr>
</tbody>
</table>

Notes: All values listed as M(SD).
Recall scores represent correct answers, from less-than-or-equal-to 3 to 5.
Writing quality and message credibility were self-reported, from 1 to 7.
Informativeness was self-reported, recoded to range from less-than-or-equal-to 2 to 5.
Hypothesis Testing

ANOVA As were run to test for manipulation effects on key dependent variables as well as potential moderating effects of motivation, which would offer insights into the mechanisms driving any effects.

As discussed in motivation manipulation as analysis, this study suggests the hypotheses can be tested as follows: If errors matter because they are inhibitors, motivation should not moderate their effect on media judgments (i.e., there should be a main effect of error on judgments); if errors matter because they are heuristic cues, then errors should make more of a difference in low-motivation conditions, as compared to high-motivation conditions (i.e., there should be an interaction effect of error and motivation on judgments).

**H1. Main effect of errors on recall.** Analysis showed no significant effect of errors on recall scores, $F(3, 500) = .70, p = .55, \eta^2_p = .00$. Recall scores were not significantly different regardless of whether the article had no errors ($M = 4.48, SD = .72$), grammar errors ($M = 4.48, SD = .76$), Associated Press errors ($M = 4.57, SD = .73$), or inverted pyramid errors ($M = 4.44, SD = .75$). Therefore, $H1$ was not supported; the study did not find that errors inhibited processing with regards to recall.

Additionally, analysis showed no significant interaction between motivation and condition in predicting recall scores, $F(3, 496) = .98, p = .40, \eta^2_p = .01$. In other words, errors did not serve as heuristic cues in affecting recall. However, the main effect of motivation on recall was significant. Recall was higher in the high-motivation condition ($M = 4.61, SD = .68$) than in the low-motivation condition ($M = 4.38, SD = .79$),
This seems to show that the motivation manipulation did have an effect, in at least causing participants to attend to the facts.

**H2. Main effect of errors on writing quality.** Analysis showed no significant relationship between error condition and perceived writing quality, \( F(3, 500) = .36, p = .78, \eta^2_p = .00 \). Writing quality perceptions were all approximately equal regardless of whether the article had no errors (\( M = 5.39, SD = .93 \)), grammar errors (\( M = 5.39, SD = 1.04 \)), Associated Press errors (\( M = 5.49, SD = .83 \)), or inverted pyramid errors (\( M = 5.42, SD = .97 \)). Therefore, H2 is not supported; the study did not find that errors negatively affected perceived writing quality.

Additionally, analysis showed no significant interaction between motivation and condition in predicting perceptions of writing quality, \( F(3, 496) = 1.25, p = .29, \eta^2_p = .01 \). In other words, motivation did not moderate perceptions of writing quality. The main effect of motivation on writing quality also was not significant. Writing quality perceptions were slightly but not significantly higher in the high-motivation condition (\( M = 5.45, SD = .89 \)) than in the low-motivation condition (\( M = 5.39, SD = .99 \)), \( F(1, 496) = .41, p = .52, \eta^2_p = .00 \).

**H3. Moderating effect of motivation and errors on message credibility.** Analysis showed no significant interaction between motivation and condition in predicting perceptions of message credibility, \( F(3, 496) = .71, p = .55, \eta^2_p = .00 \). Therefore, H3 was not supported; the study did not find that the effects of errors on perceived message credibility were moderated by motivation.

Additionally, analysis showed no significant relationship between error condition and perceived message credibility, \( F(3, 496) = .98, p = .40, \eta^2_p = .01 \). Message credibility
perceptions were all approximately equal regardless of whether the article had no errors ($M = 5.28$, $SD = .97$), grammar errors ($M = 5.42$, $SD = .96$), Associated Press errors ($M = 5.47$, $SD = .77$), or inverted pyramid errors ($M = 5.35$, $SD = .92$). That is, errors did not inhibit processing with regards to perceived message credibility. The main effect of motivation on message credibility approached significance. Message credibility perceptions were higher in the high-motivation condition ($M = 5.44$, $SD = .84$) than in the low-motivation condition ($M = 5.32$, $SD = .97$), $F(1, 496) = 2.14$, $p = .14$, $\eta^2_p = .00$.

**H4.** Moderating effect of motivation and errors on informativeness. Analysis showed no significant interaction between motivation and condition in predicting perceived informativeness, $F(3, 496) = .83$, $p = .48$, $\eta^2_p = .01$. Therefore, H4 was not supported; the study did not find that the effects of errors on informativeness were moderated by motivation.

Additionally, analysis showed no significant relationship between error condition and perceived informativeness, $F(3, 496) = 1.21$, $p = .30$, $\eta^2_p = .01$. Informativeness perceptions were all approximately equal regardless of whether the article had no errors ($M = 5.08$, $SD = 1.08$), grammar errors ($M = 5.25$, $SD = .99$), Associated Press errors ($M = 5.30$, $SD = .92$), or inverted pyramid errors ($M = 5.14$, $SD = .95$). That is, errors did not inhibit processing with regards to perceived informativeness. The main effect of motivation on message credibility also was not significant. Informativeness perceptions were the same in the high-motivation condition ($M = 5.22$, $SD = .98$) and the low-motivation condition ($M = 5.17$, $SD = 1.00$), $F(1, 496) = .23$, $p = .63$, $\eta^2_p = .00$. 
Post-hoc analyses of these four hypotheses show that collapsing the three error conditions into one (i.e., creating presence of errors and absence of errors conditions) does not change the hypothesis findings. The ANOVA results are shown in Table 4.

Table 4

| Source Table for 2(Motivation) x 4(Error Type) Between-Subjects ANOVA |
|---------------------------------|-------|---------|----------------|----------|----------|
| Source                          | df    | Sum of Squares | Mean Square | F        | p        | \( \eta^2 \) |
| Motivation                      |       |               |             |          |          |              |
| Recall                          | 1     | 6.14          | 6.14       | 11.40    | .001     | .022         |
| Writing Quality                 | 1     | .37           | .37        | .41      | .52      | .001         |
| Message Credibility             | 1     | 1.77          | 1.77       | 2.14     | .14      | .004         |
| Informativeness                 | 1     | .28           | .28        | .27      | .60      | .001         |
| Condition                       |       |               |             |          |          |              |
| Recall                          | 3     | .98           | .33        | .61      | .61      | .004         |
| Writing Quality                 | 3     | 1.01          | .34        | .38      | .77      | .002         |
| Message Credibility             | 3     | 2.43          | .81        | .98      | .40      | .006         |
| Informativeness                 | 3     | 3.72          | 1.24       | 1.22     | .30      | .007         |
| Motivation*Condition            |       |               |             |          |          |              |
| Recall                          | 3     | 1.59          | .53        | .98      | .40      | .006         |
| Writing Quality                 | 3     | 3.34          | 1.11       | 1.25     | .29      | .008         |
| Message Credibility             | 3     | 1.75          | .58        | .71      | .55      | .004         |
| Informativeness                 | 3     | 2.72          | .91        | .89      | .45      | .005         |
| Error                           |       |               |             |          |          |              |
| Recall                          | 496   | 267.12        | .54        |          |          |              |
| Writing Quality                 | 496   | 440.86        | .89        |          |          |              |
| Message Credibility             | 496   | 409.98        | .83        |          |          |              |
| Informativeness                 | 496   | 503.90        | 1.02       |          |          |              |
| Corrected Total                 |       |               |             |          |          |              |
| Recall                          | 503   | 275.98        |            |          |          |              |
| Writing Quality                 | 503   | 445.51        |            |          |          |              |
| Message Credibility             | 503   | 415.90        |            |          |          |              |
| Informativeness                 | 503   | 510.62        |            |          |          |              |

**Alternative tests for H3 and H4.** After initial testing, writing quality was considered as a proxy for the error condition to determine if writing quality was the heuristic that cued perceived message credibility and perceived informativeness.

The following regressions were conducted hierarchically; the individual measures were entered in the first step and the interaction term in the second. After looking at the interaction results, the individual measures were examined in their unique step so as to observe their significance without marginal or conditional effects of the interaction term.

Regression analysis showed a significant main effect of writing quality, prior to the inclusion of the interaction term, in predicting message credibility, \( \beta = .64, t(503) = \)
18.86, \( p < .001 \); however, there was no significant interaction between writing quality and motivation in predicting message credibility, \( \beta = .02, t(503) = .11, p = .92 \). Similarly, regression analysis showed a significant main effect of writing quality, prior to the inclusion of the interaction term, in predicting informativeness, \( \beta = .61, t(503) = 17.03, p < .001 \). However, there was no significant interaction between writing quality and motivation in predicting informativeness, \( \beta = .02, t(503) = .07, p = .94 \).

Thus, though the perceived writing quality did affect these perceptions, it did not interact with the motivation measure. Because it did not interact with motivation or react to errors, it could be that writing quality was a judgment that influenced message credibility and informativeness but not, as predicted, a heuristic that was cued by errors. This possibility is explored in the \textit{Discussion}.

\textbf{Article differences}. The three individual articles – Space, Brain, and Coral – were then analyzed to determine whether the condition effects differed by article. Analysis showed the study’s results were consistent across the articles; in other words, there were no significant interactions between article type and any effects. Across all three articles, as with the group, there were main effects of motivation on recall, no main effects of motivation on the other outcomes, no main effects of errors on any outcomes, and no interactions between errors and motivation on the other outcomes.

For illustrative purposes, the descriptive statistics for individual articles are shown in Table 5.
Table 5
Descriptive Statistics for Perceptions by Error Type and Motivation for Individual Articles

<table>
<thead>
<tr>
<th>Article</th>
<th>Motivation</th>
<th>Error Type</th>
<th>N</th>
<th>Recall (SD)</th>
<th>Writing Quality (SD)</th>
<th>Message Credibility (SD)</th>
<th>Informativeness (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Space</td>
<td>Low</td>
<td>None</td>
<td>26</td>
<td>4.38 (.75)</td>
<td>5.62 (.82)</td>
<td>5.37 (.83)</td>
<td>5.13 (.13)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Grammar</td>
<td>18</td>
<td>4.67 (.59)</td>
<td>5.53 (.80)</td>
<td>5.39 (.101)</td>
<td>5.37 (.81)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Associated Press</td>
<td>22</td>
<td>4.50 (.80)</td>
<td>5.88 (.66)</td>
<td>5.70 (.67)</td>
<td>5.56 (.95)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Inverted Pyramid</td>
<td>19</td>
<td>4.47 (.84)</td>
<td>5.54 (.105)</td>
<td>5.32 (.100)</td>
<td>5.06 (.75)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Total</td>
<td>85</td>
<td>4.49 (.75)</td>
<td>5.65 (.83)</td>
<td>5.45 (.87)</td>
<td>5.28 (.91)</td>
</tr>
<tr>
<td></td>
<td>High</td>
<td>None</td>
<td>16</td>
<td>4.56 (.73)</td>
<td>5.81 (.55)</td>
<td>5.69 (.52)</td>
<td>5.65 (.71)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Grammar</td>
<td>24</td>
<td>4.88 (.34)</td>
<td>5.84 (.72)</td>
<td>5.71 (.93)</td>
<td>5.29 (.93)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Associated Press</td>
<td>21</td>
<td>4.76 (.54)</td>
<td>5.45 (.68)</td>
<td>5.44 (.66)</td>
<td>5.27 (.69)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Inverted Pyramid</td>
<td>24</td>
<td>4.58 (.65)</td>
<td>5.63 (.75)</td>
<td>5.40 (.102)</td>
<td>5.29 (.93)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Total</td>
<td>85</td>
<td>4.71 (.57)</td>
<td>5.68 (.70)</td>
<td>5.55 (.83)</td>
<td>5.33 (.83)</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>None</td>
<td>42</td>
<td>4.45 (.74)</td>
<td>5.69 (.73)</td>
<td>5.49 (.74)</td>
<td>5.33 (.94)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Grammar</td>
<td>42</td>
<td>4.79 (.47)</td>
<td>5.71 (.76)</td>
<td>5.57 (.97)</td>
<td>5.33 (.87)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Associated Press</td>
<td>43</td>
<td>4.63 (.69)</td>
<td>5.67 (.69)</td>
<td>5.57 (.67)</td>
<td>5.42 (.84)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Inverted Pyramid</td>
<td>43</td>
<td>4.53 (.74)</td>
<td>5.39 (.88)</td>
<td>5.36 (.100)</td>
<td>5.19 (.85)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Total</td>
<td>170</td>
<td>4.60 (.67)</td>
<td>5.66 (.76)</td>
<td>5.50 (.85)</td>
<td>5.31 (.87)</td>
</tr>
<tr>
<td>Brain</td>
<td>Low</td>
<td>None</td>
<td>23</td>
<td>4.35 (.71)</td>
<td>4.97 (.100)</td>
<td>4.91 (.129)</td>
<td>4.58 (.103)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Grammar</td>
<td>25</td>
<td>4.48 (.77)</td>
<td>5.04 (.102)</td>
<td>5.16 (.84)</td>
<td>5.04 (.102)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Associated Press</td>
<td>21</td>
<td>4.52 (.81)</td>
<td>5.06 (.103)</td>
<td>5.10 (.111)</td>
<td>4.97 (.95)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Inverted Pyramid</td>
<td>24</td>
<td>4.42 (.72)</td>
<td>4.93 (.117)</td>
<td>4.94 (.101)</td>
<td>4.88 (.130)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Total</td>
<td>93</td>
<td>4.44 (.74)</td>
<td>5.00 (.104)</td>
<td>5.03 (.106)</td>
<td>4.87 (.108)</td>
</tr>
<tr>
<td></td>
<td>High</td>
<td>None</td>
<td>20</td>
<td>4.70 (.66)</td>
<td>5.07 (.87)</td>
<td>5.17 (.95)</td>
<td>4.73 (.12)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Grammar</td>
<td>15</td>
<td>4.40 (.83)</td>
<td>5.32 (.116)</td>
<td>5.20 (.101)</td>
<td>5.10 (.108)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Associated Press</td>
<td>22</td>
<td>4.82 (.50)</td>
<td>5.43 (.72)</td>
<td>5.27 (.64)</td>
<td>5.09 (.116)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Inverted Pyramid</td>
<td>19</td>
<td>4.63 (.68)</td>
<td>5.20 (.82)</td>
<td>5.37 (.79)</td>
<td>4.91 (.86)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Total</td>
<td>76</td>
<td>4.66 (.66)</td>
<td>5.26 (.88)</td>
<td>5.25 (.83)</td>
<td>4.93 (.106)</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>None</td>
<td>43</td>
<td>4.51 (.70)</td>
<td>5.02 (.93)</td>
<td>5.03 (.114)</td>
<td>4.65 (.107)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Grammar</td>
<td>40</td>
<td>4.45 (.78)</td>
<td>5.14 (.107)</td>
<td>5.18 (.90)</td>
<td>5.06 (.103)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Associated Press</td>
<td>43</td>
<td>4.67 (.68)</td>
<td>5.25 (.89)</td>
<td>5.19 (.89)</td>
<td>5.03 (.103)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Inverted Pyramid</td>
<td>43</td>
<td>4.51 (.70)</td>
<td>5.05 (.103)</td>
<td>5.13 (.93)</td>
<td>4.90 (.111)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Total</td>
<td>169</td>
<td>4.54 (.72)</td>
<td>5.11 (.98)</td>
<td>5.13 (.97)</td>
<td>4.94 (.107)</td>
</tr>
<tr>
<td>FL</td>
<td>Low</td>
<td>None</td>
<td>15</td>
<td>4.27 (.80)</td>
<td>5.32 (.92)</td>
<td>5.13 (.71)</td>
<td>5.27 (.74)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Grammar</td>
<td>20</td>
<td>4.05 (.89)</td>
<td>5.36 (.120)</td>
<td>5.52 (.108)</td>
<td>5.50 (.115)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Associated Press</td>
<td>17</td>
<td>4.06 (.90)</td>
<td>5.93 (.58)</td>
<td>5.73 (.77)</td>
<td>5.52 (.63)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Inverted Pyramid</td>
<td>21</td>
<td>4.33 (.86)</td>
<td>5.79 (.87)</td>
<td>5.71 (.80)</td>
<td>5.47 (.93)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Total</td>
<td>73</td>
<td>4.18 (.86)</td>
<td>5.61 (.95)</td>
<td>5.54 (.88)</td>
<td>5.44 (.89)</td>
</tr>
<tr>
<td></td>
<td>High</td>
<td>None</td>
<td>27</td>
<td>4.63 (.69)</td>
<td>5.54 (.104)</td>
<td>5.44 (.107)</td>
<td>5.33 (.128)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Grammar</td>
<td>20</td>
<td>4.30 (.86)</td>
<td>5.24 (.120)</td>
<td>5.50 (.90)</td>
<td>5.24 (.99)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Associated Press</td>
<td>23</td>
<td>4.65 (.65)</td>
<td>5.30 (.93)</td>
<td>5.59 (.59)</td>
<td>5.40 (.94)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Inverted Pyramid</td>
<td>22</td>
<td>4.23 (.75)</td>
<td>5.47 (.92)</td>
<td>5.42 (.78)</td>
<td>5.26 (.74)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Total</td>
<td>92</td>
<td>4.47 (.75)</td>
<td>5.40 (.101)</td>
<td>5.49 (.85)</td>
<td>5.31 (.101)</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>None</td>
<td>42</td>
<td>4.50 (.74)</td>
<td>5.46 (.99)</td>
<td>5.33 (.96)</td>
<td>5.29 (.111)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Grammar</td>
<td>40</td>
<td>4.18 (.87)</td>
<td>5.30 (.119)</td>
<td>5.51 (.98)</td>
<td>5.37 (.107)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Associated Press</td>
<td>40</td>
<td>4.40 (.81)</td>
<td>5.57 (.85)</td>
<td>5.65 (.67)</td>
<td>5.45 (.81)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Inverted Pyramid</td>
<td>43</td>
<td>4.28 (.80)</td>
<td>5.62 (.90)</td>
<td>5.57 (.79)</td>
<td>5.36 (.84)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Total</td>
<td>165</td>
<td>4.34 (.81)</td>
<td>5.49 (.99)</td>
<td>5.51 (.86)</td>
<td>5.37 (.96)</td>
</tr>
</tbody>
</table>

Notes: All values listed as M(SD).  
Recall scores represent correct answers, from less-than-or-equal-to 3 to 5.  
Writing quality and message credibility were self-reported measures from 1 to 7.  
Informativeness was self-reported, recoded to range from less-than-or-equal-to 2 to 5.  
Articles listed based on topics: Space (Chang, 2014), Brain (Gorman, 2014), and Coral (Alvarez, 2014).
**H5. Main effect of errors on overall news quality.** Analysis showed no direct effect of *error condition* on quality, \( t(502) = -0.79, p = .43 \); participants who saw errors thought the articles were of the same quality (\( M = 75.76, SD = 16.01 \)) as participants who did not see errors (\( M = 74.48, SD = 15.03 \)). Therefore, *H5* was not supported; the study did not find that the errors negatively affected overall news quality.

However, Hayes and Matthes (2009) discuss the possibility for mediation effects without main effects, so mediation analysis was still conducted, to address *RQ1.*

**RQ1. Perceptions that mediate the relationship between errors and overall news quality.** For mediation and moderation analyses, *error condition* was dummy coded such that 0 represented the error-free condition and 1 represented the error-filled condition (i.e., grammar, AP, inverted pyramid). This dichotomization allowed for mediation analysis and made sense given the predictions; the unique error types were not expected to differ in terms of psychological processes. They were separated for hypothesis testing to allow for practical implications to be drawn about the relative effects of error types on readers’ perceptions; here, no such implications were assessed.

Bootstrapping procedures using 2,000 bootstrap samples and bias-corrected confidence intervals were employed to test the direct and indirect effects of errors on quality through the mediating variables.

In addition to the core variables of interest (i.e., recall, writing quality, message credibility, and informativeness), eight new variables were included in the model: enjoyment, newsworthiness, source-related message credibility, source credibility, fairness, professionalism, familiarity, and emotion. This approach shows the effects of individual mediators and, additionally, controls for other potential confounding variables.
in the model. Enjoyment, for example, was considered a potential confounding variable in these processes; as in Pilot Test 2, the main study showed significant differences in enjoyment based on error condition, \( t(502) = -2.33, p = .02, d = .23 \); participants who saw errors thought the articles were more enjoyable (\( M = 4.60, SD = 1.30 \)) than did participants who did not see errors (\( M = 4.28, SD = 1.44 \)). Despite largely holding enjoyment steady, then, there was potential for variation.

The standardized regression weights are shown with standard error measures in Table 6, and the full model tested in AMOS 21.0 is shown in Figure 1. Several paths in the model were significant, as shown. However, the standardized direct effect of error condition on quality was not significant (\( \beta = -.00, SE = .04, p = .96 \)), and the standardized indirect effect of error condition on quality through the 12 variables also was not significant (\( \beta = .05, SE = .05, p = .33 \)).

<table>
<thead>
<tr>
<th>Perception</th>
<th>Standardized Direct Effects, Error Condition ( \rightarrow ) Perception</th>
<th>( \beta )</th>
<th>SE</th>
<th>Standardized Direct Effects, Perception ( \rightarrow ) Quality</th>
<th>( \beta )</th>
<th>SE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recall</td>
<td></td>
<td>.005</td>
<td>.044</td>
<td>.076</td>
<td>.038</td>
<td></td>
</tr>
<tr>
<td>Writing Quality</td>
<td></td>
<td>.022</td>
<td>.045</td>
<td>.356**</td>
<td>.065</td>
<td></td>
</tr>
<tr>
<td>Message Credibility</td>
<td></td>
<td>.062</td>
<td>.046</td>
<td>.017</td>
<td>.084</td>
<td></td>
</tr>
<tr>
<td>Informativeness</td>
<td></td>
<td>.065</td>
<td>.047</td>
<td>.255**</td>
<td>.068</td>
<td></td>
</tr>
<tr>
<td>Enjoyment</td>
<td></td>
<td>.104*</td>
<td>.046</td>
<td>.110</td>
<td>.062</td>
<td></td>
</tr>
<tr>
<td>Newsworthiness</td>
<td></td>
<td>.102*</td>
<td>.047</td>
<td>.000</td>
<td>.072</td>
<td></td>
</tr>
<tr>
<td>Source/Message Credibility</td>
<td></td>
<td>.039</td>
<td>.045</td>
<td>.129</td>
<td>.082</td>
<td></td>
</tr>
<tr>
<td>Source Credibility</td>
<td></td>
<td>.022</td>
<td>.044</td>
<td>.286**</td>
<td>.081</td>
<td></td>
</tr>
<tr>
<td>Fairness</td>
<td></td>
<td>.084</td>
<td>.046</td>
<td>-.051</td>
<td>.043</td>
<td></td>
</tr>
<tr>
<td>Professionalism</td>
<td></td>
<td>-.014</td>
<td>.047</td>
<td>.062</td>
<td>.058</td>
<td></td>
</tr>
<tr>
<td>Familiarity</td>
<td></td>
<td>-.005</td>
<td>.046</td>
<td>.085*</td>
<td>.041</td>
<td></td>
</tr>
<tr>
<td>Emotion</td>
<td></td>
<td>.053</td>
<td>.043</td>
<td>.070</td>
<td>.042</td>
<td></td>
</tr>
</tbody>
</table>

Note: \* \( p < .05 \), \** \( p < .001 \)
Calculated using AMOS 21.0.

*Error condition* refers to the dichotomized measure of presence of errors and absence of errors.
Note: *p < .05, **p < .001

Error condition refers to the dichotomized measure of presence of errors and absence of errors.
For the credibility measures, M. = Message, S.M. = Source-related Message, and S. = Source.
Standardized indirect effect of error condition on quality was not significant (β = .047, SE = .046, p = .324).

The model was then redrawn with only the variables for which error condition had a significant effect. The only two variables to fit these criteria were enjoyment and newsworthiness. The standardized regression weights for the final model are shown in Table 7, and the model is shown in Figure 2. Here, the standardized direct effect of error condition on quality was still not significant (β = -.04, SE = .03, p = .25). However, the standardized indirect effect of error condition on quality through the two variables was significant (β = .08, SE = .03, p = .02). Both paths, interestingly, are in the positive
direction; errors make articles more enjoyable and more newsworthy, and, therefore, participants perceive them to be of higher quality.

Thus, the study determined positive responses to RQ1(a) and RQ1(b), and it found negative responses to RQ1(c to h).

Table 7
Standardized Regression Weights for Significant Mediating Effects of Errors on Overall News Quality

<table>
<thead>
<tr>
<th>Perception</th>
<th>Standardized Direct Effects, Error Condition → Perception</th>
<th>Standardized Direct Effects, Perception → Quality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enjoyment</td>
<td>( .104^* )</td>
<td>( .046 )</td>
</tr>
<tr>
<td>Newsworthiness</td>
<td>( .102^* )</td>
<td>( .047 )</td>
</tr>
</tbody>
</table>

*Note:* \( ^* p < .05, ^{**} p < .001 \)

Calculated using AMOS 21.0.

*Error condition* refers to the dichotomized measure of presence of errors and absence of errors.

![Path analysis](image)

Figure 2. Path analysis for significant mediating effects of errors on overall news quality.

*Notes:* \( ^* p < .05, ^{**} p < .001 \)

*Error condition* refers to the dichotomized measure of presence of errors and absence of errors.

Standardized indirect effect of error condition on quality was significant (\( \beta = .075, SE = .033, p = .015 \)).

Additional Analyses

**Motivation manipulation.** Two of the hypotheses – H3 and H4 – predicted moderating effects of motivation; however, as reported above, interaction effects were not found in either case. Is this because there is no effect? The unsuccessful manipulation check (i.e., the non-significant \( t \)-test results for motivation on perceived motivation) suggests this could just be the result of poor motivation manipulation. To combat this
limitation, other measures were considered as potential proxies for motivation—perceived motivation, systematic processing, and processing time—to determine whether some element of motivation affected the findings.

The following regressions were all conducted hierarchically; the individual measures were entered in the first step and the interaction term in the second. After looking at the interaction results, the coefficients for the individual measures from the first step are considered so as to observe their significance without marginal or conditional effects of the interaction term.

First, perceived motivation was considered. Maybe the manipulation, itself, was unsuccessful, but a person’s perception of the manipulation still had an effect. Regression analysis showed a significant main effect of perceived motivation, prior to the inclusion of the interaction term, in predicting recall score, $\beta = .15, t(503) = 3.38, p = .001$; however, there was no significant interaction between condition and perceived motivation in predicting recall score, $\beta = -.29, t(503) = -.69, p = .49$. Similarly, regression analysis showed a significant main effect of perceived motivation, prior to the inclusion of the interaction term, in predicting writing quality, $\beta = .36, t(503) = 8.54, p < .001$; however, there was no significant interaction between condition and perceived motivation in predicting writing quality, $\beta = -.40, t(503) = -.10, p = .32$. Regression analysis also showed a significant main effect of perceived motivation, prior to the inclusion of the interaction term, in predicting message credibility, $\beta = .05, t(503) = 7.74, p < .001$; however, there was no significant interaction between condition and perceived motivation in predicting message credibility, $\beta = .12, t(503) = .30, p = .77$. Finally, regression analysis showed a significant main effect of perceived motivation, prior to the
inclusion of the interaction term, in predicting informativeness, $\beta = .31$, $t(503) = .74$, $p < .001$; however, there was no significant interaction between condition and perceived motivation in predicting informativeness, $\beta = -.42$, $t(503) = -1.03$, $p = .30$. These perceived-motivation results – significant main effects and non-significant interaction effects – suggest that motivation (or a proxy, thereof) directly affected article perceptions, rather than doing so in relation to the errors. Participants who self-identified as more motivated were more likely to positively rate outcome measures; in other words, the perceived motivation, rather than errors, seemed to be affecting the readers.

To corroborate these findings, a second proxy for motivation was considered: the self-reported systematic processing measure. Maybe motivation subconsciously increased perceived systematic processing, which affected the outcomes. Regression analysis showed no significant main effect of the systematic processing measure, prior to the inclusion of the interaction term, in predicting recall score, $\beta = .04$, $t(503) = .79$, $p = .43$, and no significant interaction between condition and the systematic processing measure in predicting recall score, $\beta = -.04$, $t(503) = -1.21$, $p = .84$. However, for the other three outcome measures, the pattern reflects that of the perceived motivation measure. Regression analysis showed a significant main effect of the systematic processing measure, prior to the inclusion of the interaction term, in predicting writing quality, $\beta = .25$, $t(503) = 5.81$, $p < .001$; however, there was no significant interaction between condition and the systematic processing measure in predicting writing quality, $\beta = -.05$, $t(503) = -.26$, $p = .80$. Regression analysis showed a significant main effect of the systematic processing measure, prior to the inclusion of the interaction term, in predicting message credibility, $\beta = .22$, $t(503) = 4.94$, $p < .001$; however, there was no significant
interaction between condition and the systematic processing measure in predicting message credibility, $\beta = -.05$, $t(503) = -.29$, $p = .77$. Finally, regression analysis showed a significant main effect of the systematic processing measure, prior to the inclusion of the interaction term, in predicting informativeness, $\beta = .45$, $t(503) = 11.28$, $p < .001$; however, there was no significant interaction between condition and the systematic processing measure in predicting informativeness, $\beta = -.15$, $t(503) = -.85$, $p = .40$. Again, these results suggest that motivation (or a proxy, thereof) directly affected article perceptions, rather than doing so through in relation to the errors. In this case, it could mean that systematic processing was happening, rather than heuristic processing, or it could mean significant psychological effects of perceived systematic processing.

Next, processing time was considered. Maybe the motivation subconsciously increased attention, which caused people to spend more time reading the content and, therefore, to have stronger attitudes toward it. Regression analyses were conducted in the same way as with the previous variables, but the results were less conclusive. Regression analysis showed a significant main effect of processing time, prior to the inclusion of the interaction term, in predicting recall score, $\beta = .15$, $t(503) = 3.39$, $p = .001$, and no significant interaction between condition and processing time in predicting recall score, $\beta = .16$, $t(503) = 1.22$, $p = .22$. Regression analysis showed no significant main effect of processing time, prior to the inclusion of the interaction term, in predicting writing quality, $\beta = .04$, $t(503) = .97$, $p = .33$, and no significant interaction between condition and processing time in predicting writing quality, $\beta = -.02$, $t(503) = -.15$, $p = .88$. Regression analysis showed no significant main effect of processing time, prior to the inclusion of the interaction term, in predicting message credibility, $\beta = .03$, $t(503) = .73$, ...
$p = .47$, and no significant interaction between condition and processing time in predicting message credibility, $\beta = -.05, t(503) = -.39, p = .70$. Finally, regression analysis showed no significant main effect of processing time, prior to the inclusion of the interaction term, in predicting informativeness, $\beta = -.03, t(503) = -.67, p = .51$, and no significant interaction between condition and processing time in predicting informativeness, $\beta = -.03, t(503) = -.22, p = .83$. These tests confirm that more time spent reading means more information gained, but, otherwise, do not help us to better understand the role of motivation in the process.

These tests, in all, suggest that motivation (or a proxy, thereof) directly affected article perceptions, rather than doing so through in relation to the errors. Such possibilities are further explored in the Discussion.

**Demographics.** Gender and age were then examined. Previous research has found that women and young people are more likely than men and adults to experience information overload (Ji, Ha, Sypher, 2014; Williamson & Eaker, 2012). Therefore, it could be that gender and age moderate information processing. Several tests were conducted to determine whether this was the case.

**Gender.** There were significant effects of gender on the findings. First, there was a significant interaction effect of gender and error condition (i.e., presence vs. absence of errors) on writing quality, $F(1, 503) = 8.33, p = .004, \eta_p^2 = .02$. Post-hoc analysis with a Sidak correction shows that the effect of errors on perceived writing quality was reversed for men and women. For men, error-filled articles were perceived to be of higher quality ($M = 5.50, SE = .06$) than error-free articles ($M = 5.20, SE = .11, p = .02$); for women,
error-filled articles were perceived to be of lower quality ($M = 5.35$, $SE = .08$) than error-free articles ($M = 5.62$, $SE = .123$ $p = .07$).

There also was a significant interaction effect of gender and error condition (i.e., presence vs. absence of errors) on informativeness, $F(1, 503) = 3.45$, $p = .06$, $\eta^2_p = .01$, such that errors influenced men. Post-hoc analysis with a Sidak correction shows that men who saw error-filled articles thought they were more informative ($M = 5.26$, $SE = .07$) than those who saw the error-free articles ($M = 4.94$, $SE = .12$, $p = .02$); this relationship did not exist for women.

Additionally, there were significant interaction effects of gender and motivation on outcome measures. There was a significant effect of gender and motivation on recall, $F(1, 503) = 2.88$, $p = .09$, $\eta^2_p = .01$, such that motivation influenced women. Post-hoc analysis with a Sidak correction shows that women in the high-motivation condition remembered more ($M = 4.62$, $SE = .08$) than those in the low-motivation condition ($M = 4.23$, $SE = .08$, $p = .001$); this relationship did not exist for men. There also was a significant effect of gender and motivation on informativeness, $F(1, 503) = 4.12$, $p = .04$, $\eta^2_p = .01$, such that motivation influenced men. Post-hoc analysis with a Sidak correction shows that men in the high-motivation condition perceived the article as more informative ($M = 5.25$, $SE = .09$) than those in the low-motivation condition ($M = 4.94$, $SE = .10$, $p = .02$); this relationship did not exist for women.

Finally, there were main effects of gender on outcome variables. There was a main effect of gender on recall, such that, in all, men scored higher on the recall test ($M = 4.55$, $SD = .69$) than did women ($M = 4.42$, $SD = .80$), $t(501) = 1.96$, $p = .05$, $d = .17$. Additionally, women spent more time reading the articles ($M = 141.26$,
SD = 82.17) than did men (M = 126.51, SD = 84.85), t(501) = -1.95, p = .05, d = 18, which could be because women reported higher self-motivation (i.e., the manipulation check) (M = 5.41, SD = .53) than did men (M = 5.32, SD = .56), t(501) = -1.71, p = .09, d = .17.

These tests, in all, suggest that gender affected several article perceptions. This suggests that the error conditions and the motivation manipulation affected participants differently, based on gender. Implications are addressed in the Discussion.

**Age.** Age was significantly, positively correlated with processing time (r = .19, p = .004), such that older people spent more time reading the articles. Regression analyses were then run to determine interaction effects of age and error condition (i.e., presence vs. absence of errors), as well as age and motivation, on the outcome variables. However, no such interactions were found.
Chapter 5

Discussion

Interpretation of Findings

A between-subjects experiment ($N = 504$) was conducted to observe the relationship between errors and perceptions. Participants were shown news articles with various mistakes (i.e., grammatical errors, AP style errors, and inverted pyramid structure errors) and responded to questions about their perceptions. Participants also were put in either a high-motivation or low-motivation condition to determine whether perceptions differed based on attention. The predictions and findings are explored below. In all, errors were not shown to inhibit processing or to trigger heuristic cues.

**H1. Main effect of errors on recall.** The first hypothesis predicted a main effect of errors on recall, such that articles with errors would be associated with lower recall than articles without errors. This prediction was consistent with previous studies that show errors as ability constraints that inhibit comprehension (e.g., Appelman & Bolls, 2011; Reilly & Richey, 2011; Van Weert et al., 2011; Wolf et al., 2007). It was predicted, then, that errors would negatively impact participants’ ability to remember information.

Contrary to these predictions, there was no significant main effect of errors on recall; participants were no more or less likely to remember information from articles with errors as compared to articles without errors. Errors did not affect readers’ ability to remember information from the article. This could be because the recall questions were too easy; out of five possible correct answers, the average was 4.49. However, recall scores did vary based on the motivation manipulation; recall scores were higher in the high-motivation condition than in the low-motivation condition. This suggests that the
recall scores did systematically vary, but that they were based on motivation, rather than errors in journalistic writing.

**H2. Main effect of errors on writing quality.** The second hypothesis predicted a main effect of errors on perceived writing quality, such that articles with errors would be associated with lower writing quality than articles without errors. This prediction was based on previous studies that show errors as triggers of heuristic cues (e.g., Howard, 1997; Shtyrov et al., 2003; Aaltonen et al., 2008; Pulvermüller et al., 2008) and research that showed a direct effect of errors on writing quality (e.g., Appelman & Schmierbach, 2014). It was predicted, then, that errors would cue a *writing quality* heuristic that would, in turn, influence perceptions of message credibility and informativeness.

Half of this relationship was supported by the study. Contrary to predictions, there was no significant main effect of errors on perceived writing quality; participants perceived the articles with errors to be of just as high quality as those without errors. However, writing quality did predict perceived message credibility and perceived informativeness. It could be, then, that writing quality was not a heuristic cue but simply a component of positive attitudes toward content. On the other hand, maybe writing quality, in addition to being such a judgment, was, indeed, a heuristic that cued perceptions but all participants were engaged in systematic processing, which negated the effect of the heuristic cue. This possibility is supported by the relatively high motivation scores across conditions; maybe everyone was highly motivated and, thus, systematically processing.

Additionally, there might have been a source effect for readers. The articles were all modified from articles published in *The New York Times*, so the content, itself, was
well-reported and thorough, even once stylistic errors were inserted by the researcher. Most people rated the articles to be of relatively high quality, regardless of errors; on a scale of 1 to 7, the average article quality score was 5.43. Perhaps the well-reported and thorough news content overshadowed the manipulations in journalistic writing conventions. This could mean that content has a larger influence on perceived writing quality than does style. It also could mean that the quality questions were not presented in a way that clearly measured style over content.

**H3. Moderating effect of motivation and errors on message credibility.** The third hypothesis predicted that motivation would moderate the effect of errors on perceived message credibility, such that errors would only be associated with low message credibility in the low-motivation condition. The logic of this interaction-effect prediction was based on the Heuristic-Systematic Processing Model, which says heuristic cues only affect participants when their motivation is low; when motivation is high, they are processing systematically and overlooking heuristic cues. The directionality of this prediction was based on previous studies that show errors as triggers of heuristic cues (e.g., Howard, 1997; Shtyrov et al., 2003; Aaltonen et al., 2008; Pulvermüller et al., 2008) and research that showed a direct effect of errors on message credibility (e.g., Grazian, 1996; Charney & Rayman, 1989; McGoon, 1996; Sternadori & Thorson, 2009; Appelman & Bolls, 2011). It was predicted, then, that errors would negatively affect perceived message credibility when motivation is low (i.e., when participants were heuristically processing).

Contrary to these predictions, there was no interaction effect of motivation and errors on perceived message credibility; the effects of errors on perceived message
credibility were not moderated by motivation. This could be because the motivation manipulation was not strong enough for the interaction to be significant. This possibility is supported by the fact that the motivation manipulation check was unsuccessful (i.e., the non-significant t-test results for motivation on perceived motivation).

**H4. Moderating effect of motivation and errors on informativeness.** The fourth hypothesis predicted that motivation would moderate the effect of errors on perceived informativeness, such that errors would only be associated with low informativeness in the low-motivation condition. Again, the logic of this interaction-effect prediction was based on the Heuristic-Systematic Processing Model, which says heuristic cues only affect participants when their motivation is low; when motivation is high, they are processing systematically and overlooking heuristic cues. The directionality of this prediction was based on previous studies that show errors as triggers of heuristic cues (e.g., Howard, 1997; Shtyrov et al., 2003; Aaltonen et al., 2008; Pulvermüller et al., 2008) and research that showed a non-significant but directionally appropriate direct effect of errors on perceived informativeness (e.g., Appelman & Schmierbach, 2014). It was predicted, then, that errors would negatively affect perceived informativeness when motivation is low (i.e., when participants were heuristically processing).

Contrary to these predictions, there was no interaction effect of motivation and errors on perceived informativeness; the effects of errors on perceived informativeness were not moderated by motivation. Again, this could be because the motivation manipulation was not strong enough for the interaction to be significant. This possibility
is supported by the fact that the motivation manipulation check was unsuccessful (i.e., the non-significant $t$-test results for motivation on perceived motivation).

**H5. Main effect of errors on overall news quality.** This study also predicted a main effect on perceptions of overall news quality. Analysis showed no direct effect of error condition on quality; errors did not negatively affect overall news quality. However, Hayes and Matthes (2009) discuss the possibility for mediation effects without main effects, so mediation analysis was still conducted.

**RQ1. Perceptions that mediate the relationship between errors and overall news quality.** The study’s research question asked about potential mediators in the relationship between errors and overall news quality. The purpose was to determine potential effects of additional variables, including enjoyment, newsworthiness, source-related message credibility, source credibility, fairness, professionalism, familiarity, and emotion. The prediction was that errors would negatively affect these perceptions, which would, in turn, negatively affect perceived quality.

The model showed a significant indirect effect of error condition (i.e., presence vs. absence of errors) on overall news quality through enjoyment and newsworthiness. Contrary to expectations, the errors actually made the articles more enjoyable and more newsworthy; as expected, the positive mediators made participants perceive the articles to be of higher quality.

Why would the presence of errors make articles more enjoyable and more newsworthy? These indirect effects are difficult to interpret from the given theoretical framework. It could be that errors cause readers to spend more time processing articles, which, in turn, gives them more of an opportunity for enjoyment. It also could be that
readers associate errors with breaking news, which, in turn, causes them to identify the content as more newsworthy. Both results also could reflect a novelty effect; the unexpected errors provide increased enjoyment and newsworthiness, simply because they are unexpected. In other words, errors violate readers’ expectations, so they perceive the articles to be more edgy, so they develop increased positive attitudes. These possibilities would need to be corroborated with future analysis.

**Demographics.** Finally, non-hypothesized gender and age effects were measured, to determine whether the effects were conditional upon demographic variables. Did the manipulation affect people differently?

**Gender.** Analysis showed gender effects and found that, yes, errors affected men and women differently. For men, error-filled articles were perceived to be of higher writing quality than error-free ones; for women, this pattern was reversed. For men, error-filled articles were perceived to be more informative than error-free ones; no such effect was found for women. This could mean that men are just not as knowledgeable about grammar as women, so they do not notice errors the way women do. Alternatively, maybe men do know a lot about grammar and just don’t think it matters as much as women do. Perhaps men were paying less attention to the articles, and were, therefore, less careful about their judgments. This was supported by the fact that men, overall, spent less time reading the articles than women and reported less self-motivation than women. Additionally, there was a main effect of motivation for men, such that men in the high-motivation condition perceived articles to be more informative than those in the low-motivation condition; in this way, maybe men needed more overt motivation to attend to the articles.
An alternative explanation is that men thought the articles were too simple and, therefore, not worth their attention. Overall, men scored higher on the recall test than women. Additionally, motivation did not affect men’s recall scores, but it did for those of the women; women in the high-motivation condition scored higher on the recall test than those in the low-motivation condition. It could be, therefore, that men thought the articles were easy to understand and, therefore, spent less energy reading them, causing them to be less attentive to the effects of the errors. Previous research found that women are more likely than men to experience information overload (Ji, Ha, Sypher, 2014; Williamson & Eaker, 2012), so maybe men were just less distracted and better able to understand the content, in general.

Finally, it could be that biological sex is not the moderator here, but rather some other unmeasured variable. Perhaps gender is confounded with another trait, such as need for cognition, or another skill, such as grammar knowledge. A future study that measures such additional variables could determine whether this was the case.

**Age.** Finally, age effects were tested. Analysis showed that older participants spent more time reading the articles than did younger participants. This is consistent with expectations, in terms of older people being more thorough, detailed readers. Previous research found that young people are more likely than adults to experience information overload (Ji, Ha, Sypher, 2014; Williamson & Eaker, 2012), so maybe young people shy away from intense reading. However, there were no moderating effects of age regarding the core relationships between errors, motivation, and perceptions, so this tells us little about the unsupported hypotheses.
Limitations and Future Research

If only the first hypothesis was rejected, then the study could conclude that errors are always heuristic cues; if the first was supported but the second, third, and fourth were rejected, then the study could conclude that errors are always ability constraints. Because none of the four key hypotheses were supported, this study has to, instead, look toward the possibility that either (a) the phenomenon was not observed or (b) the phenomenon does not exist. This study was not set up to test null hypotheses, so, instead, it explores the possibility that the phenomenon existed but was not observed; in other words, that the theory was correct but the design or sampling was flawed. Potential flaws and future research directions are, thus, discussed.

The most obvious limitation of this study was the ineffectiveness of the motivation manipulation. Motivation was difficult to effectively manipulate with the MTurk sample. The motivation manipulation did not strongly influence participants; participants in the high-motivation condition reported as much motivation as those in the low-motivation condition. Additionally, there was no significant relationship between the motivation manipulation and the self-reported systematic processing measure or processing time. The average score on the perceived motivation question (i.e., the manipulation check) was relatively high ($M = 5.36$, on 3 to 7 scale), so this could be the result of a ceiling effect. Maybe MTurk participants have higher social desirability concerns or are just inherently more motivated that other participant samples.

It also could be that the self-reported systematic processing measure was inappropriately modified. We modified the scale from Neuwirth, Frederick, and Mayo (2002) to be state-based rather than trait-based (e.g., “I think about how the information
in news stories about the issue relates to other things I know” to “I thought about how the information in the news article related to other things I know). This shift may have changed the measure from one of processing and motivation to one of outcome and results. This might have been the reason for the scale’s mismatch with actual motivation.

Alternatively, it could be that some participants cheated and found out about the financial bonus from other participants. Post-hoc investigation revealed that MTurk “workers” have several online forums (e.g., Reddit, Turkopticon) for sharing information about “Hits.” This could mean that participants who received the recall bonus told other participants about it, which, in turn, would influence their attention to the content and their subsequent recall scores. In fact, a quick search revealed three Reddit postings on the current study, one of which warned participants of “MCs,” which refer to “memory checks.” This particular post and its replies are shown in Appendix F. The right side of that image shows the number of community members as 28,200. It is possible that the number of cheating participants was negligible; however, because it is unclear how many of this study’s participants read the post or participated in a similar forum, this could be a more significant limitation.

Either way, the apparent failure of the motivation manipulation was a significant limitation in this study. A future study that either offers a higher monetary motivation or otherwise better manipulates motivation could clarify this study’s findings. It should be noted, however, that recall was significantly higher in the high-motivation condition than in the low-motivation condition. This seems to show that the motivation manipulation did have some effect, in terms of causing participants to pay more attention to the facts; in that way, perhaps the motivation manipulation actually was more effective than the
manipulation check would indicate. Regardless, a future study would need to increase participant motivation, either with higher financial compensation or a more enticing incentive.

It also could be that the study was divided into too many conditions and the effects were, therefore, muddled. Maybe grammar errors were not different than AP errors, but both were different than no errors, for example. In previous studies (e.g., Appelman & Bolls, 2011), only one type of error was considered; maybe this study’s scope was too vast and resulted in Type II errors in interpretation (i.e., failure to detect true effects). However, post-hoc analyses showed that collapsing the error conditions into one (i.e., creating presence of errors and absence of errors conditions) does not change the hypotheses findings. This means that it was not the case that some errors had an effect while others did not; none of the errors had effects.

Additionally, it could be that the effects were not seen because of idiosyncrasies in the content; maybe the news article, itself, contained content that was somehow less affected by errors. However, analyzing the three stimulus articles individually does not change hypotheses findings. All three news articles show the same significant main effect of motivation on recall and all of the other non-significant effects. The Brain article was the only one to show a main effect of motivation on writing quality, but this single difference is more likely a fluke than a systematic, article-based difference. Thus, it also was not the case that some articles elicited an effect while others did not; none of the articles served as the content for effects.

It could be that the type of news article matters; the articles here varied in enjoyment, but they all were science-based. Maybe errors bother people more or less
depending on the subject-matter of the article, which is why results were not detected.

Science articles were selected for this study with the explicit intention of better detecting effects; science-based articles offer the possibility for learning, which is one of the core outcome measures of this study. In addition, science articles are fairly complicated, so deep processing could conceivably affect understanding. At the time, it was unclear why another type of article would elicit stronger results; if anything, it seemed that other topics would be less likely to elicit results. This topic selection, however, might have had unintentional consequences. The complexity of the topic could have caused participants to process systematically rather than heuristically; this could have distracted them from the writing convention manipulations, thereby skewing the results. In that sense, all three articles in this study might have been atypical, leading to the unexpected findings.

Another potential design flaw could be the error manipulation, itself. This possibility is supported by the fact that the outcome variables were highly correlated with each other (see Table 2) but unrelated to the error manipulation. Additionally, as predicted, the writing quality heuristic predicted perceived message credibility and perceived informativeness, but was, itself, not triggered by the errors. Why was this the case? We did not ask participants whether they noticed errors in the main study because we were worried about sensitization effects. Therefore, we look to the pilot tests for explanations. As described, Pilot Test 2 showed that some articles worked better than others. However, results in the final study were consistent across articles, so these minor differences in effectiveness would not be likely to massively affect the outcomes. Pilot Test 2 also showed that some style manipulations worked better than others, which is perhaps the more explanatory result. In particular, the AP style manipulation was not as
noticeable as the others during the pretest. Perhaps future studies with a more aggressive AP style manipulation would help to detect the true relationships. Even the manipulations that did work better in the pilot test had only small to moderate effect sizes: \( d = .39 \) for grammatical errors, \( d = .48 \) for inverted pyramid errors. It is unclear what an ideal effect size would be, but anything about .80 would indicate a more prominent effect. Why were errors not noticeable? In the absence of baserate error data, it is unclear how many of these types of errors are published in the average news article. There could be a non-linear effect, whereby errors only bother people after a certain threshold is reached.

Maybe one mistake per paragraph is forgivable but three per paragraph is not. It is unclear why this might be the case, or, if it is, what the psychological mechanism might be. Future research with varying numbers of errors could address this possibility that there were just not enough errors in these articles to elicit results. Similarly, a study could show these stimuli to journalists to determine whether the number of mistakes is ecologically valid for a media outlet, or to a journalism professor to determine what grade such an article would earn in his or her classroom. This could help us to determine whether there were enough errors in the stimuli to show ecologically valid results.

The placement of the errors might have been the problem, as well. Errors were specifically placed in the main text, rather than in the headlines or the quotes. The intention was to ensure that participants attributed errors to reporters and editors, not to designers or sources who might be responsible for errors in those locations. It could be, though, that this made the manipulation too subtle or otherwise skewed the results. Maybe readers do not care about in-text errors that they blame on the reporter but do care about headline- or quote-based errors that they blame on the media outlet or source.
Additionally, maybe the average reader is unfamiliar with the division of labor in a newsroom and does not distinguish between errors in different locations. A future study could account for this possibility.

The length of the articles also could be considered a design flaw. In Appelman and Bolls (2011), for example, the average article length was 135 words; in this study, the average was 445. Maybe mistakes bother people more in shorter articles; alternatively, maybe this study needed to add more errors to compensate for the additional length. However, these concerns can be refuted based on the results of Pilot Test 2; when shown the manipulated articles, most participants were able to correctly determine what kind of error had been inserted into the article, even if these findings were not statistically significant. The number of errors and length of the article, then, were likely not substantial limitations, although a future study with more errors and shorter stories could confirm this.

Additionally, the articles were modified for complexity (Table 1). This was meant to ensure consistency, but it could be that the high complexity was, itself, enough to block the ability to process. Maybe future studies with more complexity-variation would help in this manner. It could be that all participants had to systematically process the text because it was so complex; this possibility is supported by the finding that perceived systematic processing was a significant predictor of perceived writing quality, perceived message credibility, and perceived informativeness. However, these findings also could just mean that all of the perception outcomes were highly correlated with each other, which, in fact, they were (Table 2). Thus, this is more likely to indicate within-subjects consistency in survey responses than it is a true indicator of processing type.
There could have been, however, other effects of complexity modification, as well as effect of other efforts at increasing internal validity. This study removed source and medium cues to isolate the effects of message. In doing so, the study might have removed so much variance that it suppressed the findings. This study would need to be redesigned with fewer restrictions to determine whether this was, in fact, the case.

Insufficient sample size is often a concern for quantitative, social-scientific research. Maybe there just was not enough power to detect the true effect. However, the current study’s sample size ($N = 504$) should have been sufficient for path analysis (according to Holbert and Stephenson, 2002) and for detection of even small-sized effects (at $r = .25$, according to Faul et al., 2007; Cohen, 1992). Thus, it was not the case that sample size hid true effects.

If the sample size was appropriate, perhaps the sample itself was not. Maybe, one could argue, MTurk is a poor method of recruitment because it is an opt-in participant pool rather than a random sample. However, if anything, participants from a volunteer-based pool should be more motivated, rather than less motivated, to attend to survey protocol. By nature of the sampling technique, they are tech-savvy and frequent Internet users; because they have done these studies before, they also are likely more familiar with the format. Additionally, research shows that demographically, MTurk samples better represent the population than do college samples or other Internet samples (Paolacci, Chandler, & Ipeirotis, 2010, p. 414). If anything, then, maybe previous studies overstated relationships between errors and outcomes because the relationship is limited to those in university settings; maybe this study showed the true relationship (i.e., no relationship) whereas the others were showing a false positive.
However, there could be some unmeasured trait of MTurk participants that skewed the results of this study. For example, maybe MTurk workers are too efficiency-minded for studies of information processing. The fact that they create web forums to discuss ways of maximizing outcomes of online studies would speak to this notion. To probe this possibility, we compared reading speed for this sample to nationwide measures. In this study, the average article length was 445 words and the average reading time was 132 seconds, which means participants were reading an average of about 200 words per minute: the average American adult can read about 250 words per minute (Thomas, n.d.). Their slower-than-average reading speed suggests the possibility that participants were reading in an unusual way. In this case, perhaps it makes sense for academic researchers to use MTurk for political science studies, for example, where the goal might be demographic and political diversity; however, for psychological studies, the sample might not be representative enough of the population to be useful.

When proxies for the unsuccessful motivation manipulation were considered, it was discovered that perceived motivation had significant main effects on recall, writing quality, message credibility, and informativeness; processing time had a significant main effect on recall; and self-reported systematic processing had significant main effects on writing quality, message credibility, and informativeness. None of these proxies interacted with error condition in predicting outcome measures. In all, then, these tests suggest that article perceptions were affected by these psychological perceptions, rather than the presence or absence of errors.

These unsupported hypotheses suggest that the relationships were not hidden by design flaws but, rather, genuinely did not exist. That is, they suggest that writing-
convention errors do not affect processing or perceptions. Therefore, the implications of these unexpected null findings are now explored.

**Theoretical and Practical Implications**

As addressed above, this study was constructed to properly detect the predicted relationships between errors and perceptions. Motivation proxies were used for confirmation and construct validity. Pretests were conducted to verify the journalistic writing convention manipulation. Multiple news articles were used for generalizability. A sufficient sample size was recruited from a diverse participant pool. Despite these efforts, however, the predicted relationships were not found. This study was not set up to test null hypotheses. However, non-significant findings in the absence of design flaws leave us with the real possibility that the phenomenon in question — the relationship between errors and perceptions — does not exist.

Null findings are often overlooked by academic publications, but this omission has negative consequences for theoretical and practical advancement. As Ioannidis (2006) argues, “A credible scientific journal should publish all studies with ‘null’ results provided they acknowledge their limitations” (p. 186). This sentiment is echoed by Levine (2011): “Researchers, journal reviewers, and editors should not keep non-significant findings out of print just because they are not statistically significant. Non-findings are informative as long as the question being tested is interesting and the method is otherwise sound” (p. 118). This section discusses the potential theoretical and practical implications if this null conclusion is in fact correct.

**Errors as expectancy-violations.** First, these non-significant findings add nuance to our theoretical understanding of the effects of errors. Errors do not affect processing
for everyone all of the time; that is, these effects are not universal. Instead, their presence in some studies (e.g., Appelman & Bolls, 2011; Appelman & Schmierbach, 2014; Vultee, 2011) and absence in the current study suggest they depend on circumstances. They seem to affect readers only when they violate expectations.

For one, people might hold different publications to different standards. Maybe it is not the grammatical errors always bother people but that they bother people only when found in publications deemed to be “expert.” Maybe an error in a local newspaper bothers a reader less than an error in a national publication. Maybe professionalism is assumed from highly reputable outlets, in which case deviation from the rules is noticeable and considered negative. Maybe the expectations are lower for less reputable news organizations and, therefore, rule-deviation is ignored. This particular study did use articles from *The New York Times*, which is generally considered to be a highly reputable publication; however, all source cues were removed from the stimuli so that participants would focus on message cues. Maybe this decision had the unintended consequence of signaling an anonymous, low-reputation source; maybe the effects were suppressed because participants thought they were reading a non-expert source. A future study that maintained source cues could corroborate this possibility.

Similarly, maybe readers hold different media types to different standards. People might expect to see mistakes online and are, therefore, not bothered by them as much as if the errors were in print. Readers might assume that print publications have more time for editing and should be, therefore, held to higher standards. The current study was conducted online, so it could be that the results are unique to online news outlets, rather
than being generalizable to other media. Maybe print media and online media are held to different standards.

It also could be that a reader’s expectation of proper writing comes from prior media exposure. A person who primarily reads poorly written articles or unedited text messages or blog posts, for example, might not be bothered by poorly written articles, simply because they are accustomed to seeing mistakes. Maybe the effect of errors seen in previous research is based on an element of surprise. A future study that measured news media consumption habits could systematically test for this possibility.

Furthermore, it could be that there is a subset of people for whom errors trigger writing quality heuristics. Maybe the general public is not bothered as much as are people with background knowledge. For example, participant profession was not adequately measured in this study. Previous studies that found relationships had used journalism and communication students or professional journalists as participants; this study used the general public. Maybe the relationship between errors and perceptions only exists for people with prior knowledge of these conventions. This notion is supported by Appelman and Schmierbach (2014), which found results to be moderated by participants’ knowledge of grammar rules. As discussed in the Literature Review, there is not really one “American” standard of grammar; this study focuses on English standards enforced by journalists in their canonic literature. Maybe these rules are class- or region-based and, thus, only bother certain people. Maybe grammar errors only bother you when you know these grammar rules; maybe AP style violations only bother you when you know AP style rules; and maybe inverted pyramid structure only bothers you when you know inverted pyramid structure. These suppositions seem to reflect common sense, but they
call into question the contention that language is automatic and that our ears are trained to
sense language problems (e.g., Shtyrov et al., 2003; Aaltonen et al., 2008; and
Pulvermüller et al., 2008).

In these ways, then, our theoretical understanding of errors needs to change to
account for errors bothering people, not all of the time, but only when they appear in
unexpected sources or when a person is in a position to otherwise recognize them. The
expectancy violations, rather than the errors, themselves, seem to be the cause of negative
correct error effects.

Rethinking journalism education. These non-significant findings also add
nuance to our practical understanding of the effects of errors, in terms of how journalism
is practiced and taught. As discussed in the Introduction, journalism is clearly undergoing
significant changes and dilemmas; practitioners and academics are unsure how to
respond. As described by South (2012), the dilemma is whether to focus on traditional
conventions (e.g., AP style) or new technology (e.g., Storify). What do we spend our time
on in the newsroom? What about in the classroom?

The results of this study suggest a potential shift in the way journalism should be
taught. Based on the non-significant results of this study, modern news audiences do not
seem to care about journalistic writing conventions, or at least the same conventions that
matter to journalists. Again, this study analyzed the effects of journalist-selected errors
(through Pilot Test 1) on readers; maybe the average audience cares about different
mistakes than do journalists. Maybe this same study would have worked differently if
journalists were used as the participant sample or if non-journalists were used in the pilot
test sample.
One way of interpreting the results is to say that these conventions simply do not matter. Readers are not bothered by structure and style errors, so it is not worth our time to correct them. This is, understandably, a tough implication for journalists and instructors to consider because it suggests that their work is no longer valued. The other extreme is to say, not that traditional errors do not matter, but, rather, that readers just do not realize how important they are. Maybe journalists and instructors need to do a better job of teaching non-journalists these conventions; maybe it is not that the conventions do not matter, but that readers do not realize how much they matter. It could be that these styles are just not being taught well enough to lay people and that if we did a better job of teaching these rules to our readers, then readers would realize their importance and be appropriately bothered by deviations. This suggests we should keep teaching what we are teaching, but that we just need more teachers and more students.

However, as is often the case, the most likely implication lies not in one of the extremes but rather in the middle. Journalists clearly still serve a role; however, that role needs to evolve. In turn, the skills taught to journalism students needs to change. It is not enough to say that we should keep doing what we are doing, and it is also not prudent to throw up our hands and walk away. Instead, our goals would best be served by rethinking the skills that we are teaching our students and the rationales for those lessons.

Maybe we should forgo some grammar lessons in exchange for more web seminars. Maybe some social media training would be more beneficial than AP style quizzes. Based on the results of this study, errors in traditional writing conventions do not bother most readers; therefore, we might be focusing too heavily on them in the
newsroom and the classroom. Perhaps they need to be taught but need to be balanced with speed and other skills of importance to journalists.

On the other hand, perhaps style instruction has a carry-over effect in the classroom. Maybe students who learn *style* also learn *attention to detail*, which increases their likelihood of providing accurate information in their articles. Research has already shown that readers value accuracy (e.g., Appelman & Sundar, forthcoming; Kang & Yang, 2011; Sundar, 1999). It could be, then, that teaching style and teaching accuracy are not mutually exclusive but are, rather, connected. Future empirical tests would need to determine whether fact-checking and accuracy are taught at the expense of style or as part of a greater lesson in detail. It could be that the justification for teaching style and structure is that those lessons add a level of care to the writing and editing process.

Rather than discount these changes as radical, it can be helpful to remember that these “traditional” conventions were never based on empirical support to begin with. Journalism educators have taught AP style and inverted pyramid structure with the assumption that they help readers; however, we never knew if that was actually the case. Does that consistency of AP style help readers? Is inverted pyramid style easier to understand than articles written in chronological order? These questions were never tested. Instead, they were brought into classrooms because they were the on-the-job skills used by journalists in the early 20th century when journalism transitioned from apprentice-based to university-based. It is no more radical to replace them than it was to institute them in the first place.

Editing clearly still matters. However, this study’s results do not support our use of “comprehension” or “consistency” as justification for editing. Based on these results,
we want to proceed with caution when saying inverted pyramid structure is easier for readers to understand or AP style gives readers a sense of consistency; these relationships were not supported. This does not mean that journalistic writing conventions are useless but, rather, that we need to rethink their purpose. Maybe it is not about helping readers. Maybe they serve as a source of professional gatekeeping (e.g., “Someone applying for a job who knows our style is one of us.”). Or maybe they help journalists to maintain a sense of professional accountability (e.g., “That publication follows our style so we should respect them as peers.”). It is true that journalists also write for other journalists, so perhaps this is more about earning the respect of peers rather than of readers. Whatever the reason, editors need to rethink the way they describe and defend their work; the old standbys were not supported in this study.

Alternatively, maybe we needed these conventions in the early 1900s when we were trying to professionalize the work of journalism, but not anymore. These more prescriptive features of style came about as journalism was becoming an academic field of study; perhaps these style conventions were being used by journalists as a way to validate the field and impress related disciplines. Maybe now, more than 100 years later, when the AEJMC Directory lists 480 journalism and mass communication programs (Becker, Vlad, & Simpson, 2014) and the American Copy Editor Society (ACES) nears its 20th anniversary, these efforts are no longer needed.

Conclusion

Previous research found effects of journalistic writing errors on news perceptions; this study sought to determine the underlying psychological mechanism explaining these effects. Do errors cue other perceptions of the text, or do they prevent us from
understanding the content? In the language of dual-processing models of persuasion, are errors heuristic cues, or are they inhibitors that block in-depth processing?

This study did not find these previous effects and, therefore, was unable to determine the underlying mechanism. Instead, the findings suggest that these effects are either contingent on study design or no longer existent. Future studies, as outlined above, could determine whether such effects are still present. If, as this study suggests, they can no longer be found, then we could be observing a shift in the way readers view historic standards for writing. Based on this study, readers are more willing to overlook style and structure errors than has been found in previous research. Therefore, this study proposes that journalists and journalism educators in the 21st century need to rethink the emphasis on style in the newsroom and the classroom.
References


South, J. (2012). ‘Editing is everything,’ but it’s evolving. *Quill, 100*(6), 9.


Appendix A

Stimulus News Articles — Control Condition

1. NASA Extends the Life of the International Space Station
Citation: Chang, K., 2014. Word Count: 453. Flesch reading score: 41.

Instead of splashing into the Pacific Ocean in 2020 as planned, the International Space Station will continue circling Earth for at least an additional four years, NASA announced on Wednesday.

William H. Gerstenmaier, NASA’s associate administrator for human exploration and operations, said the extra time would lead more scientists to propose experiments, spur a young commercial space industry and allow the space agency to solve health and technical challenges before sending its astronauts on more distant missions.

If the station were destined for destruction just six years from now, scientists and companies would be more reluctant to invest their time and money, Gerstenmaier said. “Potentially, that creates a new economy in space,” he said.

Two companies, SpaceX and the Orbital Sciences Corporation, fly cargo to the space station, and NASA is looking to hire commercial “space taxis” to take astronauts beginning in 2017. Gerstenmaier said he thought that more companies would now also look at how they could take advantage of the weightlessness of space for profit.

Last year, NASA studied the station and concluded that it could last until 2028. The other space agencies participating in it, including the European Space Agency and those of Russia and Japan, have not decided whether they, too, will continue beyond 2020.

“In general, they’re all pretty supportive,” Gerstenmaier said, adding that if necessary, NASA would go it alone.

This is the second time the Obama administration has extended the life of the orbiting laboratory.

When President Obama entered the White House, NASA was working to send astronauts back to the moon, and the space station, which was still being built, was to be nudged to a fiery ocean crash in 2016. The cost of operating the station, about $3 billion a year, could then be devoted to the moon program.
However, it always seemed unlikely that the station, which was built at a cost of $100 billion and completed just three years ago, would be discarded that soon, and when the Obama administration announced it wanted to cancel the moon program, it gave the first extension, stretching the life of the station to 2020.

Gerstenmaier said that unexpected challenges would still pop up at the station.

“This is going to be life on the frontier,” he said. Last month, a valve in a cooling pump failed, curtailing work on the station until astronauts installed a spare pump, an incident that illustrated the need for more reliable systems for deep space missions.

A cargo rocket that was to be launched to the space station on Wednesday by Orbital Sciences was delayed until at least Thursday because of an explosion on the sun and the torrent of charged particles that could have interfered with the spacecraft’s electronics.

2. N.I.H. Seeks $4.5 Billion to Try to Crack the Code of How Brains Function

The National Institutes of Health set an ambitious $4.5 billion price tag on its part of President Obama’s Brain Initiative on Thursday, stamping it as an effort on the scale of the Human Genome Project.

The goals of the Brain Initiative were clearly grand when Obama announced it a year ago — nothing less than developing and applying new technology to crack the toughest unsolved puzzles of how the brains of humans and animals function. The hope is to lay a foundation for future advances in the medical treatment of brain disorders.

But the initiative began with $110 million budgeted for 2014, shared by three major entities: the National Science Foundation; the Defense Advanced Research Projects Agency; and the N.I.H., which has a $40 million share.

By calling for such a major commitment, to be spread over 12 years, the institutes answered concerns among neuroscientists about the initial level of funding.

“This is a realistic amount of money,” said Dr. Eric R. Kandel, director of the Kavli Institute for Brain Science at Columbia University, who, like some other neuroscientists, had been skeptical of what could be accomplished with the funding committed when the initiative was announced about a year ago.
Gerald Rubin, the executive director of the Janelia Farm Research Campus in Virginia, also found that this budget request allayed some of his concerns, but not all. “I am much more concerned about convincing Congress to fund the Brain Initiative at this level,” he said.

The institutes currently spend about $4.5 billion a year on existing neuroscience research. Dr. Francis S. Collins, director of the N.I.H., embraced the Brain Initiative with open arms, set up an all-star working group to study what research was needed and supported efforts to set a broad neuroscience agenda for the coming decade.

An interim report last September did just that, but did not present an estimate of the money needed. A final version presented to the institutes on Thursday tackled that issue.

The new funding recommendations, which Dr. Collins accepted today as goals for the institutes to pursue, would gradually increase until they reached $500 million a year in 2020 and stay at that level until 2025.

For the 2015 fiscal year, Dr. Collins said, the president has already asked for $100 million for the N.I.H.

The report, from a committee led by Cori Bargmann of Rockefeller University and William Newsome of Stanford University, emphasizes technology development for the first five years and use of the new technology in scientific inquiry in the second five years.

“It won’t be fast, it won’t be easy, and it won’t be cheap,” Dr. Collins said Thursday. But, he said, “It will be a pretty exciting ride.”

3. Single Dose of Antibiotic Found Effective in Quelling MRSA


A single infusion of an antibiotic can clear serious bacterial skin infections just as effectively as the 10-day regimen now used to treat patients, researchers reported Wednesday.

Many patients do not finish the complicated treatment for these infections, which requires two infusions of antibiotics daily, often in a hospital. Such incomplete treatments may breed resistance to antibiotics in surviving bacteria. A single-dose therapy may make it easier to treat these dangerous infections, said the authors of the new study, published in The New England Journal of Medicine.
The study was led by researchers at Duke University and designed and funded by the Medicines Company, the maker of the antibiotic, oritavancin.

The drug, to be sold as Orbactiv, may be approved by the Food and Drug Administration as early as August under a special fast-track process, the company said.

“This is a bit of a light at the end of a dismal tunnel in the development of new antibiotics,” said Dr. William Schaffner, an infectious disease specialist at Vanderbilt University, who was not involved in the new study.

In an editorial accompanying the research, Dr. Henry F. Chambers, a professor of medicine at the University of California, San Francisco, said the one-shot antibiotic infusion could transform the treatment of acute bacterial skin infections and alter how these infections are managed.

“These patients could potentially just get an antibiotic and not be admitted to the hospital at all,” Dr. Chambers said in a telephone interview.

Skin infections are among the most common reasons that doctors use intravenous antibiotics. Rates of infection have risen significantly in recent decades. There are an estimated 15 million cases in the United States each year, and they cause 870,000 hospitalizations.

In the new study, 475 patients with acute bacterial skin infections received a large single infusion of the antibiotic, while 479 similar patients received the current treatment, Oritavancin, twice a day for seven to 10 days. Response rates to the two regimens were similar: 82 percent and 78.9 percent, respectively, saw skin lesions shrink or stop spreading, and their fevers disappeared.

Oritavancin is effective because it persists in the body. Though none of the trial participants developed an allergic reaction to the drug, a serious reaction could result in extended illness, Dr. Chambers said.

There is no way to stop the treatment if a patient turns out not to have a drug-resistant infection, some experts noted.

And if patients are sent home after the treatment is administered, there is a chance that doctors could miss even more serious infections, said Dr. G. Ralph Corey of Duke University Medical Center, the lead author of the study.

4. Researchers Race to Save Coral in Miami


With many of the reefs in South Florida faltering — more than 90 percent of Miami’s coral has disappeared — there is new urgency in transplanting corals and researching why they are dying and how to help them thrive.
With dredging set to begin this weekend in the Port of Miami, researchers are scrambling to salvage a much larger than expected trove of delicate corals that were not in the original marine life rescue plan for the undersea area.

In the past few days, researchers say they have found thousands more specimens along sections of the channel beyond where the Army Corps of Engineers was obligated to remove large, important corals in preparation for the dredging. Most of those corals have been removed by the Army Corps of Engineers and, for the most part, have been transplanted onto other reefs.

But researchers recently secured permits for a second sweep of the channel, including outside the designated areas, in the hope of collecting smaller corals for scientific and educational purposes. During their dives, they said, they found much larger colonies of large, healthy stony corals than expected along the busy shipping channel.

The colonies are important, researchers said, because they show resilience and offer insight that could help save South Florida’s ailing coral reefs.

“They were allowed to go out and see what was left over,” said Rachel Silverstein, the executive director of Biscayne Bay Waterkeeper, which brought the original lawsuit to protect corals in Government Cut.

The Army Corps of Engineers has announced that it will begin dredging the port to allow access for larger vessels this weekend, six weeks sooner than researchers had expected.

Researchers said they had hoped to begin the difficult work, which involves coping with cruise ships, freighters, tides and poor visibility, earlier this year and finish in July.

Because of fluctuations in the Army Corps of Engineers’ construction schedule, the researchers were not issued a permit until May 26 and now have been told to clear out by Friday. Their efforts have also been hampered by bad weather.

Colin Foord, a marine biologist and one of the founders of Coral Morphologic, which combines art and science, said he and other divers had come across large specimens of coral that should have been removed by the Army Corps of Engineers.

The goal now is to try to safely transplant as much of the corals as possible by Friday, Foord said, adding that divers had not even reached a second reef.

“The corals are living in an industrial shipping lane,” Foord said. “We see them as perhaps being ideal research projects to find out why these corals are so adaptive.”
Appendix B

Pilot Test 1 Questionnaire — Grammatical Errors

Which five grammatical errors do you consider to be the most problematic?

1. incorrect spelling of proper nouns (e.g., Barak Obma)
2. incorrect spelling of common nouns (e.g., sissors)
3. incorrect subject-verb agreement (e.g., they is)
4. confused homophones (e.g., they’re vs. their vs. there)
5. double negatives (e.g., not never)
6. split infinitives (e.g., to quickly run)
7. run-on sentences (e.g., I am tired, and he likes TV, and they are broke.)
8. sentence fragments (e.g., Because of the rain.)
9. incorrect pronoun-antecedent agreement (e.g., A student raised their hand.)
10. misplaced modifiers (e.g., Walking to the store, her wallet fell.)
11. incorrect adverb placement (e.g., It has really been tough.)
12. comma splice (e.g., It was cold, it was raining.)
13. confusion of that and which (e.g., Dogs, which are old, tend to get sicker than young dogs.)
14. incorrect pronoun case (e.g., he vs. him vs. his)
15. incorrect verb tense (e.g., was vs. is vs. will be vs. had been)
16. incorrect verb mood (e.g. If it was me vs. If it were me)
17. confusion of like and as (e.g., He ran as a tiger.)
18. forgotten comma in compound sentences (e.g., She is sad and he is happy.)
19. added comma in complex sentences (e.g., He is happy, because she is sad.)
20. added comma in restrictive phrases/clauses (e.g., I have the answer, that she wants.)
21. forgotten comma in non-restrictive phrases/clauses (e.g., I have the answer which she wants.)
22. confusion of possessive and plural cases of nouns (e.g., buildings vs. building’s)
23. incorrect punctuation of compound modifiers (e.g., a well rehearsed presentation)
24. nonparallel construction (e.g., She is smart, sweet, and likes to read.)
25. confusion of a and an (e.g., a elephant)
Appendix C

Pilot Test 1 Questionnaire – AP Style Errors

Which five AP style errors do you consider to be the most problematic?

1. confusion between spelling out proper names and using abbreviations
2. using incorrect abbreviations
3. using periods in acronyms
4. confusion between spelling out numbers and using numerals
5. incorrect construction of p.m. and a.m.
6. confusion between postal code and AP state abbreviations
7. spelling out states with cities
8. abbreviating states when they stand alone
9. spelling out months with dates
10. abbreviating months when they stand alone
11. spelling out street/avenue/boulevard with numbered addresses
12. abbreviating street/avenue/boulevard when they stand alone
13. abbreviating alley/drive/road
14. incorrect punctuation of composition titles (e.g., The Associated Press Stylebook)
15. incorrect use of courtesy titles (e.g., Ms., Mr.)
16. adding a serial comma (e.g., one, two, three, and four)
17. placing commas and periods outside of quotation marks
18. inserting two spaces after a period
19. incorrect capitalization of formal titles
20. using trademarks as generic terms
21. using yesterday/today/tomorrow instead of days of the week
22. incorrect hyphenation of ages
23. confusion between using % and percent
24. adding an s to toward/backward/upward/forward
25. referring to a person by first and last name after first reference
Appendix D

Stimulus News Articles – Treatment Conditions

1. NASA Extends the Life of the International Space Station

Grammatical Error Condition.

Instead of splashing into the Pacific Ocean in 2020 as planned, the International Space Station will continue circling Earth for at least an additional four years, NASA announced on Wednesday.

William H. Gerstenmaier, NASA's associate administrator for human exploration and operations, said the extra time would lead more scientists to propose experiments, spur a young commercial space industry and allow the space agency to solve health and technical challenges before sending its astronauts on more distant missions.

If the station were destined for destruction just six years from now, scientist’s and companies would be more reluctant to invest there time and money, Gerstenmaier said. “Potentially, that creates a new economy in space,” he said.

Two companies, SpaceX and the Orbital Sciences Corporation, fly cargo to the space station, and NASA is looking to higher commercial “space taxis” to take astronauts beginning in 2017. Gerstenmaier said he thought that more companies would now also look at how they could take advantage of the weightlessness of space for profit.

Last year, NASA studied the station and concluded that it could last until 2028. The other space agencies participating in it, including the European Space Agency and those of Russia and Japan, has not decided whether they, too, will continue beyond 2020.

“In general, they’re all pretty supportive,” Gerstenmaier said, adding that if necessary, NASA would go it alone.

This is the second time the Obama administration has extended the life of the orbiting laboratory.

When President Obama entered the White House, NASA was working to send astronauts back to the moon, and the space station, which was still being built, was to be nudged to a fiery ocean crash in 2016. The cost of operating the station, about $3 billion a year, could
then be devoted to the moon program.

However, it always seemed unlikely that the station, which was built at a cost of $100 billion and completed just three years ago, would be discarded that soon, and when the Obama administration announced it wanted to cancel the moon program, they gave the first extension, stretching the life of the station to 2020.

Gerstenmaier said that unexpected challenges would still pop up at the station.

“This is going to be life on the frontier,” he said. Last month, a valve in a cooling pump failed, curtailing work on the station until astronauts installed a spare pump, an incident that illustrated the need for more reliable systems for deep space missions.

A cargo rocket that was to be launched to the space station on Wednesday by Orbital Sciences were delayed until at least Thursday because of an explosion on the sun and the torrent of charged particles that could have interfered with the spacecraft’s electronics.

**AP Style Error Condition.**

Instead of splashing into the Pacific Ocean in 2020 as planned, the International Space Station will continue circling Earth for at least an additional 4 years, NASA announced on Wednesday.

William H. Gerstenmaier, NASA’s Associate Administrator for Human Exploration and Operations, said the extra time would lead more scientists to propose experiments, spur a young commercial space industry, and allow the space agency to solve health and technical challenges before sending its astronauts on more distant missions.

If the station were destined for destruction just six years from now, scientists and companies would be more reluctant to invest their time and money, Gerstenmaier said. “Potentially, that creates a new economy in space”, he said.

Two companies, SpaceX and the Orbital Sciences Corporation, fly cargo to the space station, and NASA is looking to hire commercial “space taxis” to take astronauts beginning in 2017. Gerstenmaier said he thought that more companies would now also look at how they could take advantage of the weightlessness of space for profit.

Last year, the National Aeronautics and Space Administration studied the station and concluded that it could last until 2028. The other space agencies participating in it, including the European Space Agency and those of Russia and Japan, have not decided whether they, too, will continue beyond 2020.
“In general, they’re all pretty supportive,” Gerstenmaier said, adding that if necessary, NASA would go it alone.

This is the second time the Obama administration has extended the life of the orbiting laboratory.

When president Obama entered the White House, NASA was working to send astronauts back to the moon, and the space station, which was still being built, was to be nudged to a fiery ocean crash in 2016. The cost of operating the station, about $ three billion a year, could then be devoted to the moon program.

However, it always seemed unlikely that the station, which was built at a cost of $100 bill, and completed just three years ago, would be discarded that soon, and when the Obama administration announced it wanted to cancel the moon program, it gave the first extension, stretching the life of the station to 2020.

Gerstenmaier said that unexpected challenges would still pop up at the station.

“This is going to be life on the frontier,” he said. Last month, a valve in a cooling pump failed, curtailing work on the station until astronauts installed a spare pump, an incident that illustrated the need for more reliable systems for deep space missions.

A cargo rocket that was to be launched to the space station on Wednesday by Orbital Sciences was delayed until at least tomorrow because of an explosion on the sun and the torrent of charged particles that could have interfered with the spacecraft’s electronics.

Inverted Pyramid Structure Error Condition.

If the station were destined for destruction just six years from now, scientists and companies would be more reluctant to invest their time and money, Gerstenmaier said. “Potentially, that creates a new economy in space,” he said.

Two companies, SpaceX and the Orbital Sciences Corporation, fly cargo to the space station, and NASA is looking to hire commercial “space taxis” to take astronauts beginning in 2017. Gerstenmaier said he thought that more companies would now also look at how they could take advantage of the weightlessness of space for profit.

Last year, NASA studied the station and concluded that it could last until 2028. The other space agencies participating in it, including the European Space Agency and those of
Russia and Japan, have not decided whether they, too, will continue beyond 2020.

“In general, they’re all pretty supportive,” Gerstenmaier said, adding that if necessary, NASA would go it alone.

Instead of splashing into the Pacific Ocean in 2020 as planned, the International Space Station will continue circling Earth for at least an additional four years, NASA announced on Wednesday.

This is the second time the Obama administration has extended the life of the orbiting laboratory.

When President Obama entered the White House, NASA was working to send astronauts back to the moon, and the space station, which was still being built, was to be nudged to a fiery ocean crash in 2016. The cost of operating the station, about $3 billion a year, could then be devoted to the moon program.

However, it always seemed unlikely that the station, which was built at a cost of $100 billion and completed just three years ago, would be discarded that soon, and when the Obama administration announced it wanted to cancel the moon program, it gave the first extension, stretching the life of the station to 2020.

William H. Gerstenmaier, NASA’s associate administrator for human exploration and operations, said the extra time would lead more scientists to propose experiments, spur a young commercial space industry and allow the space agency to solve health and technical challenges before sending its astronauts on more distant missions.

Gerstenmaier said that unexpected challenges would still pop up at the station.

“This is going to be life on the frontier,” he said. Last month, a valve in a cooling pump failed, curtailing work on the station until astronauts installed a spare pump, an incident that illustrated the need for more reliable systems for deep space missions.

A cargo rocket that was to be launched to the space station on Wednesday by Orbital Sciences was delayed until at least Thursday because of an explosion on the sun and the torrent of charged particles that could have interfered with the spacecraft’s electronics.
2. N.I.H. Seeks $4.5 Billion to Try to Crack the Code of How Brains Function

Grammatical Error Condition.

The National Institutes of Health set an ambitious $4.5 billion price tag on their part of President Obama’s Brain Initiative on Thursday, stamping them as an effort on the scale of the Human Genome Project.

The goal’s of the Brain Initiative were clearly grand when Obama announced it a year ago — nothing less than developing and applying new technology to crack the toughest unsolved puzzles of how the brains of humans and animals function. The hope is to lay a foundation for future advances in the medical treatment of brain disorders.

But the initiative began with $110 million budgeted for 2014, shared by three major entities: the National Science Foundation; the Defense Advanced Research Projects Agency; and the N.I.H., which has a $40 million share.

By calling for such a major commitment, to be spread over 12 years, the institutes answered concerns among neuroscientists about the initial level of funding.

“This is a realistic amount of money,” said Dr. Eric R. Kandel, director of the Kavli Institute for Brain Science at Columbia University, who, like some other neuroscientists, had been skeptical of what could be accomplished with the funding committed when the initiative were announced about a year ago.

Gerald Rubin, the executive director of the Janelia Farm Research Campus in Virginia, also found that this budget request allayed sum of his concerns, but not all. “I am much more concerned about convincing Congress to fund the Brain Initiative at this level,” he said.

The institutes currently spend about $4.5 billion a year on existing neuroscience research. Dr. Francis S. Collins, director of the N.I.H., embraced the Brain Initiative with open arms, set up an all-star working group to study what research was kneaded and supported efforts to set a broad neuroscience agenda for the coming decade.

An interim report last September did just that but did not present an estimate of the money needed. A final version presented to the institutes on Thursday tackled that issue.
The new funding recommendations, which Dr. Collins accepted today as goals for the institutes to pursue, would gradually increase until they reached $500 million a year in 2020 and stay at that level until 2025.

For the 2015 fiscal year, Dr. Collins said, the president have already asked for $100 million for the N.I.H.

The report, from a committee led by Cori Bargmann of Rockefeller University and William Newsome of Stanford University, emphasizes technology development for the first five years and use of the new technology in scientific inquiry in the second five years.

“It won’t be fast, it won’t be easy, and it won’t be cheap,” Dr. Collins said Thursday. But, he said, “It will be a pretty exciting ride.”

AP Style Error Condition.

The National Institutes of Health set an ambitious $4.5 billion price tag on its part of President Obama’s Brain Initiative yesterday, stamping it as an effort on the scale of the Human Genome Project.

The goals of the Brain Init. were clearly grand when Obama announced it a year ago — nothing less than developing and applying new technology to crack the toughest unsolved puzzles of how the brains of humans and animals function. The hope is to lay a foundation for future advances in the medical treatment of brain disorders.

But the initiative began with $110 million budgeted for 2014, shared by three major entities: the National Science Foundation; the Defense Advanced Research Projects Agency; and the N.I.H., which has a $40 million share.

By calling for such a major commitment, to be spread over twelve years, the institutes answered concerns among neuroscientists about the initial level of funding.

“This is a realistic amount of money”, said Dr. Eric R. Kandel, Director of the Kavli Institute for Brain Science at Columbia University, who, like some other neuroscientists, had been skeptical of what could be accomplished with the funding committed when the initiative was announced about a year ago.

Gerald Rubin, the executive director of the Janelia Farm Research Campus in Virginia, also found that this budget request allayed some of his concerns, but not all. “I am much
more concerned about convincing Congress to fund the Brain Initiative at this level,” he said.

The institutes currently spend about $4.5 billion a year on existing neuroscience research. Dr. Francis S. Collins, director of the National Institutes of Health, embraced the Brain Initiative with open arms, set up an all-star working group to study what research was needed and supported efforts to set a broad neuroscience agenda for the coming decade.

An interim report last September did just that, but did not present an estimate of the money needed. A final version presented to the institutes on Thursday tackled that issue.

The new funding recommendations, which Dr. Collins accepted today as goals for the institutes to pursue, would gradually increase until they reached $500 million a year in 2020 and stay at that level until 2025.

For the 2015 fiscal year, Dr. Collins said, the president has already asked for $ one-hundred million for the N.I.H.

The report, from a committee led by Cori Bargmann of Rockefeller Univ. and William Newsome of Stanford University, emphasizes technology development for the first 5 years and use of the new technology in scientific inquiry in the second five years.

“It won’t be fast, it won’t be easy, and it won’t be cheap,” Dr. Collins said Thursday. But, he said, “It will be a pretty exciting ride”.

Inverted Pyramid Structure Error Condition.

But the initiative began with $110 million budgeted for 2014, shared by three major entities: the National Science Foundation; the Defense Advanced Research Projects Agency; and the N.I.H., which has a $40 million share.

By calling for such a major commitment, to be spread over 12 years, the institutes answered concerns among neuroscientists about the initial level of funding.

“This is a realistic amount of money,” said Dr. Eric R. Kandel, director of the Kavli Institute for Brain Science at Columbia University, who, like some other neuroscientists, had been skeptical of what could be accomplished with the funding committed when the initiative was announced about a year ago.
Gerald Rubin, the executive director of the Janelia Farm Research Campus in Virginia, also found that this budget request allayed some of his concerns, but not all. “I am much more concerned about convincing Congress to fund the Brain Initiative at this level,” he said.

The National Institutes of Health set an ambitious $4.5 billion price tag on its part of President Obama’s Brain Initiative on Thursday, stamping it as an effort on the scale of the Human Genome Project.

The institutes currently spend about $4.5 billion a year on existing neuroscience research. Dr. Francis S. Collins, director of the N.I.H., embraced the Brain Initiative with open arms, set up an all-star working group to study what research was needed and supported efforts to set a broad neuroscience agenda for the coming decade.

An interim report last September did just that, but did not present an estimate of the money needed. A final version presented to the institutes on Thursday tackled that issue.

The goals of the Brain Initiative were clearly grand when Obama announced it a year ago — nothing less than developing and applying new technology to crack the toughest unsolved puzzles of how the brains of humans and animals function. The hope is to lay a foundation for future advances in the medical treatment of brain disorders.

The new funding recommendations, which Dr. Collins accepted today as goals for the institutes to pursue, would gradually increase until they reached $500 million a year in 2020 and stay at that level until 2025.

For the 2015 fiscal year, Dr. Collins said, the president has already asked for $100 million for the N.I.H.

The report, from a committee led by Cori Bargmann of Rockefeller University and William Newsome of Stanford University, emphasizes technology development for the first five years and use of the new technology in scientific inquiry in the second five years.

“It won’t be fast, it won’t be easy, and it won’t be cheap,” Dr. Collins said Thursday. But, he said, “It will be a pretty exciting ride.”
3. Single Dose of Antibiotic Found Effective in Quelling MRSA

Grammatical Error Condition.

A single infusion of an antibiotic can clear serious bacterial skin infections just as effectively as the 10-day regimen now used to treat patients, researchers reported Wednesday.

Many patients do not finish the complicated treatment for these infections, which requires too infusions of antibiotics daily, often in a hospital. Such incomplete treatments may breed resistance to antibiotics in surviving bacteria. A single-dose therapy may make it easier to treat these dangerous infections, said the authors of the new study, published in The New England Journal of Medicine.

The study were led by researchers at Duke University and designed and funded by the Medicines Company, the maker of the antibiotic, oritavancin.

The drug, to be sold as Orbactiv, may be approved by the Food and Drug Administration as early as August under a special fast-track process, the company said.

“This is a bit of a light at the end of a dismal tunnel in the development of new antibiotics,” said Dr. William Schaffner, an infectious disease specialist at Vanderbilt University, who was not involved in the new study.

In an editorial accompanying the research, Dr. Henry F. Chambers, a professor of medicine at the University of California, San Francisco, said the one-shot antibiotic infusion could transform the treatment of acute bacterial skin infections and alter how these infection’s are managed.

“These patients could potentially just get an antibiotic and not be admitted to the hospital at all,” Dr. Chambers said in a telephone interview.

Skin infections is among the most common reasons that doctors use intravenous antibiotics. Rates of infection have risen significantly in recent decades. Their are an estimated 15 million cases in the United States each year, and they cause 870,000 hospitalizations.

In the new study, 475 patients with acute bacterial skin infections received a large single infusion of the antibiotic, while 479 similar patients received the current treatment, Oritavancin, twice a day for seven to 10 days. Response rates to the two regimens were similar: 82 percent and 78.9 percent, respectively, saw skin lesions shrink or stop spreading, and their fevers disappeared.

Oritavancin is effective because they persists in the body. Though none of the trial participants developed an allergic reaction to the drug, a serious reaction could result in extended illness, Dr. Chambers said.
There is no way to stop the treatment if a patient turns out not to have a drug-resistant infection, some experts noted.

And if patients are sent home after the treatment is administered, there is a chance that doctors could miss even more serious infections, said Dr. G. Ralph Corey of Duke University Medical Center, the lead author of the study.

**AP Style Error Condition.**

A single infusion of an antibiotic can clear serious bacterial skin infections just as effectively as the 10-day regimen now used to treat patients, researchers reported yesterday.

Many patients do not finish the complicated treatment for these infections, which requires two infusions of antibiotics daily, often in a hospital. Such incomplete treatments may breed resistance to antibiotics in surviving bacteria. A single-dose therapy may make it easier to treat these dangerous infections, said the authors of the new study, published in The New England Journal of Medicine.

The study was led by researchers at Duke University and designed and funded by the Medicines Company, the maker of the antibiotic, oritavancin.

The drug, to be sold as Orbactiv, may be approved by the Food and Drug Administration (FDRA) as early as August under a special fast-track process, the company said.

“This is a bit of a light at the end of a dismal tunnel in the development of new antibiotics,” said Dr. William Schaffner, an Infectious Disease Specialist at Vanderbilt University, who was not involved in the new study.

In an editorial accompanying the research, Dr. Henry F. Chambers, a professor of medicine at the University of California, San Francisco, said the one-shot antibiotic infusion could transform the treatment of acute bacterial skin infections and alter how these infections are managed.

“These patients could potentially just get an antibiotic and not be admitted to the hospital at all,” Dr. Chambers said in a telephone interview.

Skin infections are among the most common reasons that doctors use intravenous antibiotics. Rates of infection have risen significantly in recent decades. There are an estimated fifteen million cases in the United States each year, and they cause 870,000 hospitalizations.

In the new study, 475 patients with acute bacterial skin infections received a large single infusion of the antibiotic, while 479 similar patients received the current treatment, Oritavancin, twice a day for 7 to 10 days. Response rates to the two regimens were
similar: 82 percent and 78.9 percent, respectively, saw skin lesions shrink or stop spreading, and their fevers disappeared.

Oritavancin is effective because it persists in the body. Though none of the trial participants developed an allergic reaction to the drug, a serious reaction could result in extended illness, Doctor Chambers said.

There is no way to stop the treatment if a patient turns out not to have an drug-resistant infection, some experts noted.

And if patients are sent home after the treatment is administered, there is a chance that doctors could miss even more serious infections, said Dr. G. Ralph Corey of Duke University Medical Center, the Lead Author of the study.

Inverted Pyramid Structure Error Condition.

Many patients do not finish the complicated treatment for these infections, which requires two infusions of antibiotics daily, often in a hospital. Such incomplete treatments may breed resistance to antibiotics in surviving bacteria. A single-dose therapy may make it easier to treat these dangerous infections, said the authors of the new study, published in The New England Journal of Medicine.

The study was led by researchers at Duke University and designed and funded by the Medicines Company, the maker of the antibiotic, oritavancin.

The drug, to be sold as Orbactiv, may be approved by the Food and Drug Administration as early as August under a special fast-track process, the company said.

“This is a bit of a light at the end of a dismal tunnel in the development of new antibiotics,” said Dr. William Schaffner, an infectious disease specialist at Vanderbilt University, who was not involved in the new study.

In an editorial accompanying the research, Dr. Henry F. Chambers, a professor of medicine at the University of California, San Francisco, said the one-shot antibiotic infusion could transform the treatment of acute bacterial skin infections and alter how these infections are managed.

“These patients could potentially just get an antibiotic and not be admitted to the hospital at all,” Dr. Chambers said in a telephone interview.

A single infusion of an antibiotic can clear serious bacterial skin infections just as effectively as the 10-day regimen now used to treat patients, researchers reported Wednesday.

In the new study, 475 patients with acute bacterial skin infections received a large single infusion of the antibiotic, while 479 similar patients received the current treatment,
Oritavancin, twice a day for seven to 10 days. Response rates to the two regimens were similar: 82 percent and 78.9 percent, respectively, saw skin lesions shrink or stop spreading, and their fevers disappeared.

Oritavancin is effective because it persists in the body. Though none of the trial participants developed an allergic reaction to the drug, a serious reaction could result in extended illness, Dr. Chambers said.

There is no way to stop the treatment if a patient turns out not to have an drug-resistant infection, some experts noted.

Skin infections are among the most common reasons that doctors use intravenous antibiotics. Rates of infection have risen significantly in recent decades. There are an estimated 15 million cases in the United States each year, and they cause 870,000 hospitalizations.

And if patients are sent home after the treatment is administered, there is a chance that doctors could miss even more serious infections, said Dr. G. Ralph Corey of Duke University Medical Center, the lead author of the study.
4. Researchers Race to Save Coral in Miami

Grammatical Error Condition.

With many of the reefs in South Florida faltering — more than 90 percent of Miami's coral has disappeared — there are new urgency in transplanting corals and researching why they are dying and how to help them thrive.

With dredging set to begin this weekend in the Port of Miami, researchers are scrambling to salvage a much larger than expected trove of delicate corals that were not in the original marine life rescue plan for the undersea area.

In the past few days, researchers say they have found thousands more specimens along sections of the channel beyond where the Army Corps of Engineers was obligated to remove large, important corals in preparation for the dredging. Most of those corals have been removed by the Army Corps of Engineers and, for the most part, have been transplanted onto other reefs.

But researchers recently secured permits for a second sweep of the channel, including outside the designated areas, in the hope of collecting smaller corals for scientific and educational purposes. During their dives, they said, they found much larger colonies of large, healthy stony corals than expected along the busy shipping channel.

The colonies are important, researchers said, because they show resilience and offer insight that could help save South Florida's ailing coral reefs.

“They were allowed to go out and see what was left over,” said Rachel Silverstein, the executive director of Biscayne Bay Waterkeeper, which brought the original lawsuit to protect corals in Government Cut.

The Army Corps of Engineers has announced that they will begin dredging the port to allow access for larger vessels this weekend, six weeks sooner than researchers had expected.

Researchers said they had hoped to begin the difficult work, which involves coping with cruise ships, freighters, tides and poor visibility, earlier this year and finish in July.

Because of fluctuations in the Army Corps of Engineers’ construction schedule, the researchers were not issued a permit until May 26 and now have been told to clear out by Friday. Their efforts have also been hampered by bad weather.

Colin Foord, a marine biologist and one of the founders of Coral Morphologic, which combines art and science, said he and other divers had come across large specimens of coral that should have been removed by the Army Corps of Engineers.
The goal now is to try to safely transplant as much of the coral’s as possible by Friday, Foord said, adding that divers had not even reached a second reef.

“The coral's are living in an industrial shipping lane,” Foord said. “We see them as perhaps being ideal research projects to find out why these corals are so adaptive.”

**AP Style Error Condition.**

With many of the reefs in South Florida faltering — more than ninety percent of Miami’s coral has disappeared — there is new urgency in transplanting corals and researching why they are dying and how to help them thrive.

With dredging set to begin this weekend in the Port of Miami, researchers are scrambling to salvage a much larger than expected trove of delicate corals that were not in the original marine life rescue plan for the undersea area.

In the past few days, researchers say they have found 1,000s more specimens along sections of the channel beyond where the Army Corps of Engineers was obligated to remove large, important corals in preparation for the dredging. Most of those corals have been removed by the Army Corps of Engineers and, for the most part, have been transplanted onto other reefs.

But researchers recently secured permits for a second sweep of the channel, including outside the designated areas, in the hope of collecting smaller corals for scientific and educational purposes. During their dives, they said, they found much larger colonies of large, healthy stony corals than expected along the busy shipping channel.

The colonies are important, researchers said, because they show resilience and offer insight that could help save South Florida’s ailing coral reefs.

“They were allowed to go out and see what was left over”, said Rachel Silverstein, the Executive Director of Biscayne Bay Waterkeeper, which brought the original lawsuit to protect corals in Government Cut.

The Army Corps of Engineers has announced that it will begin dredging the port to allow access for larger vessels this weekend, six weeks sooner than researchers had expected.

Researchers said they had hoped to begin the difficult work, which involves coping with cruise ships, freighters, tides, and poor visibility, earlier this year and finish in July.

Because of fluctuations in the Army Corps of Engineers’ construction schedule, the researchers were not issued a permit until May 26 and now have been told to clear out by tomorrow. Their efforts have also been hampered by bad weather.
Colin Foord, a Marine Biologist and one of the founders of Coral Morphologic, which combines art and science, said he and other divers had come across large specimens of coral that should have been removed by the Army Corps of Engineers.

The goal now is to try to safely transplant as much of the corals as possible by Fri., Foord said, adding that divers had not even reached a second reef.

“The corals are living in an industrial shipping lane,” Foord said. “We see them as perhaps being ideal research projects to find out why these corals are so adaptive.”

**Inverted Pyramid Structure Error Condition.**

In the past few days, researchers say they have found thousands more specimens along sections of the channel beyond where the Army Corps of Engineers was obligated to remove large, important corals in preparation for the dredging. Most of those corals have been removed by the Army Corps of Engineers and, for the most part, have been transplanted onto other reefs.

But researchers recently secured permits for a second sweep of the channel, including outside the designated areas, in the hope of collecting smaller corals for scientific and educational purposes. During their dives, they said, they found much larger colonies of large, healthy stony corals than expected along the busy shipping channel.

The colonies are important, researchers said, because they show resilience and offer insight that could help save South Florida’s ailing coral reefs.

“They were allowed to go out and see what was left over,” said Rachel Silverstein, the executive director of Biscayne Bay Waterkeeper, which brought the original lawsuit to protect corals in Government Cut.

**With many of the reefs in South Florida faltering — more than 90 percent of Miami’s coral has disappeared — there is new urgency in transplanting corals and researching why they are dying and how to help them thrive.**

The Army Corps of Engineers has announced that it will begin dredging the port to allow access for larger vessels this weekend, six weeks sooner than researchers had expected.

Researchers said they had hoped to begin the difficult work, which involves coping with cruise ships, freighters, tides and poor visibility, earlier this year and finish in July.

Because of fluctuations in the Army Corps of Engineers’ construction schedule, the researchers were not issued a permit until May 26 and now have been told to clear out by Friday. Their efforts have also been hampered by bad weather.
With dredging set to begin this weekend in the Port of Miami, researchers are scrambling to salvage a much larger than expected trove of delicate corals that were not in the original marine life rescue plan for the undersea area.

Colin Foord, a marine biologist and one of the founders of Coral Morphologic, which combines art and science, said he and other divers had come across large specimens of coral that should have been removed by the Army Corps of Engineers.

The goal now is to try to safely transplant as much of the corals as possible by Friday, Foord said, adding that divers had not even reached a second reef.

“‘The corals are living in an industrial shipping lane,” Foord said. “We see them as perhaps being ideal research projects to find out why these corals are so adaptive.”
Appendix E

Recall Questions

NASA Extends the Life of the International Space Station.

1. What NASA-related project was the focus of the article? (International Space Station, Mars Rover, Spacewalk)
2. What was the deadline that was just extended? (2016, 2020, 2024, 2026)
3. Select the company (companies) that fly cargo to the space station (select all that apply). (SpaceX, Orbital Sciences Corporation, Rocket Lab, Mishaal Aerospace, Scorpius Space Launch Company)
4. Who was the NASA worker quoted throughout the article? (William H. Gerstenmaier, Jaiwon Shin, John Grunsfeld, Michael Gazarik)
5. This is the _____ time the Obama administration has extended the mission. (First, Second, Third, Fourth)

N.I.H. Seeks $4.5 Billion to Try to Crack the Code of How Brains Function.

1. What National Institutes of Health project was the focus of the article? (Brain Initiative, Human Genome Project, Stem Cell Research)
2. How much money does the NIH say it needs to complete the project? ($2.5 million, $10 million, $4.5 billion, $6 billion)
3. Select the group(s) initially involved in the project (select all that apply). (National Science Foundation, Defense Advanced Research Projects Agency, Centers for Disease Control and Prevention, Food and Drug Administration, Centers for Medicare and Medicaid Services)
4. Who was the N.I.H. director quoted throughout the article? (Dr. Francis Collins, Dr. Harold Varmus, Dr. Paul Sieving, Dr. Gary Gibbons)
5. Obama announced the goals of the initiative _____ year(s) ago. (One, Two, Three, Four)

Researchers Race to Save Coral in Miami.

1. What marine life-form was the focus of the article? (Coral, Jellyfish, Sponges)
2. How much of Miami’s coral has disappeared? (20 percent, 45 percent, 90 percent, 100 percent)
4. Which marine conservationist group was discussed in the article? (Coral Morphologic, Coral Reef Alliance, Reef Check, Planetary Coral Reef Foundation)
5. Researchers say they have found ____ of specimens along the channel. (Dozens, Hundreds, Thousands, Millions)
Appendix F

Discussions among MTurk Workers

119
Alyssa Appelman Curriculum Vitae

EDUCATION

Doctor of Philosophy (Ph.D.) in Mass Communications, 2015
The Pennsylvania State University, College of Communications

Master of Arts (M.A.) in Journalism, 2009
University of Missouri-Columbia, School of Journalism

Bachelor of Journalism (B.J.), Emphasis in Magazine, Minor in French, 2008
University of Missouri-Columbia, School of Journalism
Semesters at: Institut d’Études Politiques (Paris) and Macquarie University (Sydney)

RESEARCH


TEACHING

News Editing and Evaluation, Spring 2014, Fall 2014
News Writing and Reporting, Spring 2013, Fall 2013, Summer 2014, Spring 2015
Basic News Writing Skills, Summer 2012, Fall 2012

HONORS AND AWARDS

Certificate in Teaching with Technology, The Pennsylvania State University, 2014
Certificate in Graduate Teaching, The Pennsylvania State University, 2014
Certificate for Online Teaching, The Pennsylvania State University, 2013
University Graduate Fellowship, The Pennsylvania State University, 2011 – 2015
Summa Cum Laude, University of Missouri-Columbia, 2008
Certificate of General Honors, University of Missouri-Columbia, 2008