1. Problem

Statistical literacy matters. Each day, students confront messages that contain statistical language or reasoning in some form, be it at the doctor’s office, on social media, or in advertising, among other places. These encounters accumulate and have a nontrivial effect on the decisions we make (Fagerlin et al. 2007; Peters et al. 2007). In particular, a common message that students encounter involves risk caused by a specific factor or set of related factors. Headlines such as “Red wine could… prevent heart disease” (Bennett 2016) and “Fish oil supplements might cause cancer” (Salzberg 2013) are ubiquitous in modern news outlets, often reducing complex phenomena to a single antecedent cause. Such messages tend to portray a specific risk as related to one root, isolated from other factors. But while these stories provide for wonderful conversation, the writers—whether intentionally or not—often misrepresent the research they report on by leaving out specific details (Angell and Kassirer 1994; Condit 2004; Moynihan et al. 2000). These details are key for any realistic understanding of the ramifications, if any, of the research.

Statistical literacy involves critically interrogating this discourse, recognizing its reductionist nature and reading past “magical explanations” (Freire 2005, p. 13). That is, it entails evaluating others’ statistical arguments and giving cogent responses of one’s own. Unfortunately, this interrogative stance is not directly supported through current curriculum standards in mathematics (National Council of Teachers of Mathematics 2000; National Governors Association Center for Best Practices and Council of Chief State School Officers 2010); instead, the topic is arguably orphaned, covered only at the discretion of an instructor in K-12 classrooms. Because this topic appears to be absent within K-12 curricula for mathematics and statistics, in this research I sought to gain a preliminary snapshot of how students enrolled in a general education college mathematics course evaluated and made arguments surrounding media that contains messages of risk. The media article I chose for this research (see the online supplemental file), one on the relationship between processed meats and cancer (Johnson 2015), was a fruitful place to begin this exploration.

2. Purpose

Conceptions of risk form an expansive field of research (e.g., Breakwell 2007). Accordingly, the purpose for completing the present study was not to develop a theory for how college students interact with messages of risk; rather, my goal was to examine patterns in a specific group’s response to a common type of message—risk to one’s health, caused by particular factors (in this case, processed meats). My intention was to explore the nature of claims students draw from in such encounters, including the types of warrants (or assumptions) behind those
claims. The driving research questions in this study were the following:

1. In evaluating the evidence presented in a media message about risk in relation to one’s health, what is the nature of students’ arguments?
2. What assumptions do students use to justify their evaluations of media messages about risk in relation to one’s health?

Exploring these questions can aid in creating learning experiences for this group that foster an understanding of the complexity of risk and the extent to which statistical reasoning can shed light on it.

3. Literature Review

3.1. Statistical Literacy and Risk

Calls for attention to statistical literacy (or reasoning) have proliferated in the past three decades. Whether at the fringes of standards for mathematics (e.g., National Council of Teachers of Mathematics 2000; National Governors Association Center for Best Practices & Council of Chief State School Officers 2010) or as the central beacon call for an increase of statistics in the curriculum (e.g., American Statistical Association 2007, 2010), the notion of statistical literacy has emerged as a key component of one’s overall literacy (Gal 2002). Notwithstanding these differences, each call has alluded in some way to the need for high-school and college graduates to be prepared for the 21st century, wherein data—and accompanying statistical arguments—are ubiquitous. A common thread within the myriad of frameworks is the notion that a statistically literate individual should not only understand basic statistical and mathematical arguments in interacting with real-world media but also have a critical disposition to do so (Gal 2002). Without this disposition to interrogate and reason with media, one may take for granted its veracity or untruth. Accordingly, statistical knowledge is not sufficient and, in fact, though the foundations of statistical literacy (e.g., percentages, data collection, variability, etc.) develop in middle school, many have argued that high-school and postsecondary instructors have the responsibility to nurture statistical literacy growth as students develop (Hughes-Hallett 2003).

Because this learning is supposed to take place early on, the tradeoff is that someone should attend to this growth. The topic of interaction with media messages about risk and one’s health is problematic, as its place within various curricular guidelines is difficult to spot. Indeed, because its mathematical underpinnings (percentages and proportions) are simple, one might assume that students understand it. Nevertheless, because critical thinking is involved beyond knowing mathematical underpinnings, such an assumption is misguided. Moreover, with most standards in mathematics and statistics, there is a push to examine studies for their validity and conclusions—not necessarily mediating messages, those that are found in media or conversation, that one might encounter in daily experiences. It is problematic to solely examine studies, as it is the mediating message that is often incorrect.

This is problematic because the ability to understand risk in relation to one’s health is neither widespread among the general population nor a given among those with more years of education (e.g., Forrow, Taylor, and Arnold 1992; Perneger and Agoritsas 2011; Reyna et al. 2009). An unhelpful facet of the matter is that media and research outlets often translate research studies into easily consumable readings, which—though engaging—are not necessarily accurate (Angell and Kasirer 1994;condit 2004; Moynihan et al. 2000). There are a number of reasons journalists might do this, including the desire to make stories simple, ploys to bring in readers, and of course, honest misunderstandings of the scientific research process. The result is that if a person does not have an interrogative stance toward risk within media, they could easily be misled. This discussion highlights possible issues with respect to messages within media, but of course, these are not the only types of encounters students face in daily life. Conversation and social media, among other things, matter too. Information processed from such mediums is subject to the same issue.

3.2. Causal Understanding

In relation to information processing, research suggests that humans have an innate tendency to track, categorize, and develop causal explanations for events; however, our thinking is subject to a number of biases that—while useful in some instances—often lead us astray (Kahneman, Slovic, and Tversky 1982). One bias is that we often “grossly overestimate” our understanding of complex phenomena, mistaking “sketchy relationships for richer intuitive theories” (Keil 2003, pp. 666–667; Keil 2012). This notion accounts for the fact that when individuals are asked to describe a concept they feel confident about (e.g., how a helicopter works), they often struggle. Additionally, people employ a number of heuristics in tracking causal relations, including that of determining whether or not an individual or report should be considered authoritative, as well as paying greater attention to information that aligns with preexisting beliefs (Keil 2012). Indeed, concerning the latter heuristic, human tendency is to use previous background information and experiences in “coordinating expectations with evidence that may or may not support them” (Kuhn 2012, p. 330). I believe that a recognition of the complexity of causal relationships, as well as the ability to coordinate new information with our prior beliefs, are both key (albeit implicit) components of statistical literacy.

These heuristics and processes arise in part because, in processing information, we have a tendency to encode information both verbatim, as well as its gist; these encodings are independent of one another (Reyna 2004), and the distinction is the key underpinning of fuzzy-trace theory (Reyna and Brainerd 1991). Despite the fact that individuals encode verbatim and gist representations of information, they tend to use gist in maintaining and tracking categories and causal relations; in fact, specific verbatim details are often forgotten nearly immediately upon processing. This preference for gist increases with age, as children transition from using primarily verbatim information in reasoning to using gist as adults. As a concrete example, if a student reads that there is a 0.22% chance of developing a cancer if one drinks coffee, he or she is likely to forget the verbatim absolute risk immediately; what remains is the gist that the risk with coffee is low.
The implications of this research in the context of interpreting risk within media are both dismal and promising. On the one hand, if an article or message contains accurate information, such as that there is a small absolute reduction in cancer risk from eating chocolate, one will likely encode the gist of “low risk remains low risk,” leading to no change in habit—arguably a well-reasoned decision. On the other hand, if an article reporting the same study stated that eating chocolate halves one’s risk of developing cancer, an individual might encode the gist that chocolate is an important factor in cancer risk—a misguided interpretation. Individuals extract gist from a message, and the specific type of gist that the piece fosters has a nontrivial impact on how one acts upon reading the information. Moreover, if one takes such a message and then relays it to another individual, their own mediation—likely to include gist, rather than verbatim details—leads to another iteration of gist reasoning. The cycle becomes problematic when the source mediator begins with or originates misleading or incorrect interpretations.

3.3. Argumentation Framework

When evaluating original research or messages from a media source, students themselves, whether intentionally or not, are making arguments. For example, if a student sees a post on Facebook linking vaccinations with autism, and comments “Couldn’t agree more...I have never trusted medicine, anyway,” the student is claiming that they agree with the author, and for a stated reason. Though not stated outright, such a statement is an acknowledgment of what is true. Given my interest in seeing what elements are important to students as they interact with such media, it is important to analyze how they reason as they make arguments. A straightforward yet pragmatic tool for doing this is Toulmin’s (1958) framework for argumentation, which has utility for both written and oral arguments. Many scholars have found the framework useful, with areas of application ranging from argumentation in formal mathematics (Forman et al. 1998), to explanations teachers make in reasoning about their practice (Bieda, Sela, and Chazan 2014). In education research, Toulmin’s framework is the dominant theoretical tool for analyzing argumentation (Nussbaum 2011); it is particularly constructive for the short, written arguments that I analyze in this study.

While there exist adaptations of the framework for various purposes, in this study I use that from Simosi (2003), wherein arguments contain up to four elements: the claim itself (a conclusion), the data (any evidence alluded to), the warrant (a means of connecting the claim and data), and the backing (if applicable, an assumption supporting the warrant). As everyday arguments are often intricate and ill-defined, one should not expect every argument to contain all four elements explicitly; however, we can often still infer the warrant or backing from context (Sims 2003). As a simple example, consider the following statement: I will not go outside today because it is too hot outside. Here the claim is that one will not go out; the data are that it is too hot outside; the warrant—which in this case is implicit—is that the person does not want to be hot; and the backing (if any) is that being in a hot environment will lead one to become hot. Short statements of this type were present in most students’ responses in this study (discussed below), and tend to have ostensible claims, data, warrants, and backings (or at least three of the four), making Toulmin’s framework an appropriate tool for analysis.

4. Method

For this study, I focused on students who do not intend to go into mathematically intensive fields. To a large degree, such students are those who enroll in terminal mathematics or statistics courses, where terminal means that the course is likely the last mathematics or statistics course that the student will take. This population is useful for this study because—like that at most universities—it is a relatively large one at Michigan State University. As with many institutions of higher education across the country, Michigan State University has made strides over the past two decades to revise its mathematics graduation requirement to better meet students’ needs. In addition to offering several statistics courses students can take to fulfill the mathematics requirement, several mathematics and statistics faculties have worked to develop a new course, Quantitative Literacy, that centers on contexts relevant to students and the mathematics they might encounter or find useful in engaging in such contexts. Though the course is housed in the University’s Mathematics Department and does not center exclusively on statistical thinking and methods, the curriculum does include attention to medical testing, absolute and relative measures, and science in the media, the latter of which is often reported based on research using statistical methods. This attention suggests that an article on processed meats and cancer would be an appropriate one for students in the course to analyze. Indeed, of the five central learning goals for the course, the first is that students “Critically analyze quantitative information and recognize that mathematical and statistical methods have limits.” Thus, one would reasonably expect for the course to foster critical engagement with messages of risk and causation in media. All the while, given that students do not respond to real-world contexts in a vacuum (even when in the classroom), there is a need for us, as instructors and curriculum developers, to understand how students in the course are reasoning with articles outside of the expectation that they use course methods. Indeed, others have already noted that, for a variety of reasons, students do not immediately apply mathematical or statistical principles when faced with real-world issues (Enyedy and Mukhopadhyay 2007). Though not based on a pre- and postassessment, if this study finds that there is a need for such courses to include a greater focus on message interrogation, the Quantitative Literacy curriculum can adjust. I discuss potential adjustments based on my findings in Section 7.

4.1. Participants

Participating students were enrolled at Michigan State University in a 2015 pilot offering of Quantitative Literacy. Access was granted by the instructor and Michigan State University’s Institutional Review Board to interact with students in the course at the start of the semester, before they delved deeply into the course material; at the time, I was involved in curriculum
development, but not in the teaching of any of the course sections. This timing was intentional insofar as I did not seek to understand how the course was influencing students’ responses (a question that would make sense in subsequent research), but rather to explore where they were before going through the course. The course instructor asked students to respond to a task after reading an article that I chose about cancer risk in relation to various lifestyle choices (Johnson 2015). Students completed the task outside of class using an online survey tool, and received a small amount of extra credit for doing so, with 165 of the 200 students completing the task. Students knew that their responses were ungraded and would not be seen by the instructor; these characteristics of the task were key, as ideally they signaled to students that they could use reasoning as they might in daily life.

4.2. Task

Intended as a light and fun read for the consumer, the article contained many opinionated claims without addressing the complexity of the research it referenced. I felt that it was an exemplar of a message about risk in relation to one’s health, as it contained little quantitative information or interpretation of the study’s findings. The task required students to explain—in their own terms—whether they thought the author’s claims were justified and if they agreed with its message. Note that two students read an original draft of the questions to give feedback on the wording. The questions, bulleted below, were opened-ended and allowed students to give their views on the topic in relation to those of the author, specifically asking students to make claims about the article’s validity:

- Consider the following claim of the article: We should not worry about the daily influx of studies about health, as if we did, we would become overwhelmed and lose out on the joys of life. Instead, we should enjoy life without regrets and enjoy various things in moderation.
- Do you believe the author had sufficient evidence to support this claim? Explain.
- Do you agree with the message above? Why or why not?
- What did you take away from the article? Is there anything in addition to the claim above that you noted or believe warrants discussion?

A limitation inherent in the questions above is that I extracted the central claim for the respondents to engage with. Not all students may have taken the same message away; however, by asking if the author had evidence for this claim, whether they agreed with it, and allowing for a final, open-ended comment, I aimed to capture any departures from this claim that students found.

4.2.1. An “Expert” Response

Given the ubiquity of oversimplified messages in modern media, my prediction was that students would—to a large extent—agree with the author in the task. Though not the authoritative source on how one should interpret such an article, I certainly have ideas for how an expert would do so. Thus, to provide a starting point for later analysis (Breakwell 2007, p. 99), I include an “expert” response (written by me) to the first question here, before having read students’ responses.

An “Expert” Response: The author’s thesis—that we should not trouble over reports that seem to come out on a near daily basis—is attractive, but nonetheless flawed. Of course, it seems logical that we should, as Johnson (2015) points out, “try to lead a happy and productive life without focusing on how soon that which makes you happy will cause it to end.” But this is an oversimplification of reality. We seem to be given two options: to worry about each new study that reports risk in relation to our health, or to not care at all about any of them. This masks a middle ground wherein one can examine messages of risk with an awareness of both the complexity of the scientific process and the fact that studies contribute to our understanding of a given thing—not define it. In the case of processed meats and colorectal cancer, Johnson (2015) interprets the results as absolute truth, noting that eating such meats each day ensures one is “definitely doomed.” While the study Johnson reports on does suggest that eating large quantities of processed meats may increase one’s risk of cancer, this does not automatically imply that such meats are carcinogenic (a number of confounding factors are plausible here), nor that the risk of developing the cancer is large (the baseline risk of the disease goes unmentioned). Johnson then couples the risk from processed meats with a discussion of other known carcinogens, including the sun, uranium, and car exhausts, among other things. The argument from there is that all of these things have the potential to harm us in some fashion, and hence if we focused on avoiding them all, we would miss out on life. Again, this line of reasoning seems sensible, but it masks the distinctions among the degree to which these carcinogens occur in the population, how much exposure one would need to develop the cancer, as well as how dangerous the resulting cancers are (e.g., skin cancer versus colorectal cancer). All of these complexities are important to understand before one simply puts them under the same umbrella.

Important to note here is that one would not expect for students to write the response given above. There are a number of points here that students could draw from to craft their own thoughts; still yet, there are many points not included here that would be reasonable for one to make. What matters is that one recognizes the oversimplification of reality evidenced in the article; this recognition, coupled with a nod to absolute and relative measures of risk, is what students would ideally state upon having taken the course.

4.3. Data Analysis

4.3.1. Argumentation Framework

In preparation for the larger task of analyzing all of the students’ responses, using Simos’s interpretation of Toulmin’s framework, I took a random sample of 20 students’ answers to the task. I pooled all of each student’s answers to the task’s three questions into a single block of text, looking for a prominent claim about their evaluation of the author’s evidence. With that in hand, I applied the framework to tease out the student’s claim, data, warrant, and backing; if I could not find one of the four elements, I speculated what was meant to some degree. If too much speculation was necessary to understand an element, or if two or more argument elements were missing, I did not include the response within my results for analysis. This is the reason that only 152 of the 165 responses are included in subsequent analysis. Below is an example quote from a student’s response to the task, followed by an analysis in Table 1 underneath.

He read some articles on processed meats and the sun causing cancer. Although many people have gotten cancer just like
he said, he provided zero scientific evidence, like numbers or links, as evidence of his claims.

As an additional example, in Table 2 I have taken the “expert” response from earlier and separated it into its component parts. Notably, as we will see later, none of the data, warrants, or backings from Table 2 manifested themselves sufficiently in students’ responses, and so those in Table 2 are not seen in the coding scheme later.

I shared the 20 rows of argument breakdowns with two colleagues for their feedback, discussing places of disagreement and any potential overinterpretations of student responses.

### 4.3.2. Coding

With 20 rows of argument breakdowns and feedback from colleagues, I then developed a list of codes for the types of claims, data, and warrants present in the 20 sample analyses, with the intention of using those to look at the rest of the sample. I developed these codes by first going through the response rows one by one, adding a claim, data, warrant, or backing to my list when I recognized it; given the variety of codes I assembled after going through the 20 random rows, I then grouped together codes that were similar in nature, doing so until I felt that each of the codes was distinct from one another. The codes I developed for the four argument components, based on the random sample of 20 from the pool of responses, are listed in Table 3. Note that the columns of Table 3 are independent of one another; there are no rows.

With this list in hand, I then began the process of examining the rest of the codes, with the intention of adding to the list of codes only if a response deviated significantly from the current list.

### 5. Results

#### 5.1. Codes

An important note to keep in mind about Table 3 is that it only gives an idea of how students responded; students who fell into each code category may have intended to convey more (or less) than what the codes capture. With that said, given that I was careful not to make assumptions when reading students’ responses, I feel confident that the codes I assigned are meaningful representations of what the students intended to convey. As one can see, each of the 152 students made some sort of claim about the sufficiency of the author’s evidence; some felt it was enough, while others thought otherwise. Students anchored their response with a variety of data, ranging from information in the article (e.g., the author said uranium causes cancer) or common sense, to supposedly well-stated opinions, or—and quite importantly—a lack of references. The warrants connecting the data and the claims were straight-forward to tease out (though most bluntly noted it). As an example, with warrant A many students explained the sufficiency of evidence with the inclusion (or exclusion) of information in the article. The other claim-data connections used similar warrants. The backings—though often unstated—directly followed from the warrants. For instance, warrant B (that harkening to common sense provides for evidence in an argument) follows from the assumption that we can trust our common sense.

#### 5.2. Code Assignment

With the codes for each element of the argument in place, I assigned to each student’s response a code sequence corresponding to the codes they used in their argument. For instance, the code A1A1 means that, given that the article included information, and that information serves as evidence for one’s argument, the student felt the author had sufficient evidence; of course, that that information serves as evidence assumes that information in articles of this type is true.

I present common code variations, along with the overall percentage of students with such arguments, in Table 4. By common, I mean that at least five student responses used the codes associated with the variation. For clarity, I include an interpretation of the meaning of this argument structure.

This table contains an abundance of information, most notably being that the majority of students felt the author had sufficient support for the claim that we should live life to the fullest without concern for studies that implicate various things as potential carcinogens. Within this large group, most (the A1A1 variation) assumed that the author’s interpretation of various research was accurate; they agreed with his message and expressed no reason to assume otherwise. Hence, students drew from resources the article provided to come to a conclusion. Additionally, as I speculate later (Section 6.1), a possible reason for why students responded this way is that they may have held the same views as those of the author. But some students directly noted this, with those of the A12AB12 variation arguing that the author was merely restating information that is common sense. Others yet—taking a different approach—harkened to his reference to the World Health Organization (WHO) to say that his argument was coherent. Thus, there was a broad spectrum of reasons students has for their agreement with the author.

But of course, not everyone was in agreement. Interestingly enough, a large group of respondents noted that the author did not provide sufficient references or statistics to bolster his argument, with some (the B14AD13 variation) saying he had some (but not enough) references, and others yet (the C4D3 and C14D3 variations) noting he lacked them completely. This

---

**Table 1.** A Toulmin breakdown of a student’s response.

| Claim | Data | Warrant | Backing |
|-------|------|---------|---------|
| The author did not present sufficient evidence to support his claim. | The author did not present statistics or links to studies. | Recognition of complexity is important for argumentation. | All-or-nothing thinking is inappropriate for sound arguments. |

**Table 2.** A Toulmin breakdown of an “expert” response.

| Claim | Data | Warrant | Backing |
|-------|------|---------|---------|
| The author did not present sufficient evidence to support his claim. | The author’s comparison of various risks is flawed. | Different substances or activities pose different types of risk. | Comparing distinct things merits a discussion of their differences. |
result—that different levels of the same type of evidence are sufficient for different individuals—is thought-provoking and merits further discussion later. For now, it is worthwhile to note that the backings and warrants behind these assume that statistics and citations are necessary for a valid argument.

But even when these resources were present (or not), it appears that students’ evaluation of the author’s evidence did not align with whether or not they agreed with the central message. In Table 5, the 152 responses are broken down by the respondent’s evaluation of the author’s evidence, coupled with—based on their response—whether or not they demonstrated agreement with the author’s message.

Table 5 begs a question: does agreement with an author’s message—whether the message is substantive or not—have a strong impact on how one evaluates an author’s evidence? Conversely, does how one evaluates an author’s message have any impact on whether they agree or disagree with it? These are questions I explore later, though—as one might imagine—their answers are beyond the scope of this article.

6. Discussion

Before delving into a discussion, it is important to recall my driving research questions: In evaluating the evidence presented in a media message about risk in relation to one’s health, what is the nature of students’ arguments? What assumptions do students use to justify their evaluations of media messages about risk in relation to one’s health? Tables 2 and 3 begin to shed light on potential answers. Let us begin with the largest variation of response type: those beginning with “A,” meaning the student believed Johnson (2015) provided sufficient evidence for his claim.

6.1. Sufficient Evidence From the Author: The “A” Variety

The simplest of the argument types that begin with “A” were those of the A1A1 variety, comprising 42% of all responses. In essence, the A1A1 argument was that—by talking about a long list of things that “cause” cancer—the author provided sufficient evidence for his claim that we should not worry very much about studies that consistently come out about risk in relation to our health. The warrant in the A1A1 response is that a discussion of others’ work qualifies as sufficient evidence to make one’s own point. In and of itself, such a warrant is sensible, because without it, one would have to prove trivial propositions in order to make even minute arguments. On a grander level, this tendency to track relationships and arguments is a characteristic of evolution (Buss 1995; Cosmides and Tooby 2013). With that said, the backing for this warrant is that we can trust the author’s interpretations of previous research. In
many cases this is appropriate; however, as I discussed earlier, here an “expert’s” response would include questioning the author’s analysis. Most students in this variant simply took his word as truth. As an example, below is a portion of an A1A1 student’s response:

The author explains that research shows that bacon cause cancer. He then goes on to say that there are tons of things that can shorten our lives and that we cannot truly avoid every single one of them. This is sufficient evidence to me.

Problematic here, as seen in the quote above, is that students in the A1A1 category took the author’s interpretation as authoritative. In the context of my research question, it appears from this that a significant resource students draw upon is what others say, a heuristic humans commonly use (Keil 2012). Given that this was a published article on a major news outlet’s site, it is likely that students assumed the content within the piece was true. Again, this is not problematic in and of itself. The source itself is the issue. In relation to the resource of what others say, it is worthwhile to revisit the notion of gist reasoning. A common theme that manifested within students’ responses—especially those in the A1A1 category—was to repeat the gist of what the author stated, leaving out verbatim details. As an example, below are two quotes from separate A1A1 student responses.

The evidence in this article to support the claim is that sunlight is considered a carcinogen, meat is a carcinogen, cars give off carcinogens, and so does almost everything else in this world. We are going to die anyways, so it is up to us how we want to live out our days. I believe that there is sufficient evidence in this article because it convinced me that we do not need to be so extremely cautious of everything when we are eventually going to die anyways. Might as well live out our lives the way we want to.

I definitely believe that this is sufficient evidence because those are the most basic things we need for survival. If the light that touches your skin outside, the air you breathe, and the food you eat all cause cancer than what’s the point of trying to avoid things you enjoy that also “cause cancer.”

In both responses, the problematic gist is that sunlight, processed meats, car exhaust, etc. are equivalent as carcinogens, with the implication being that we should not avoid any of them. But of course, sunlight and processed meats are not dangerous in the same ways. Each has its own complexities that one should be aware of before reaching the conclusion that we should simply give in to all of them without recourse to moderation.

With all of this said, this discussion of resources students draw from—others’ views and the gist they get from such—is not meant to paint students’ understandings as wrong or lacking. Rather, it is meant to start a conversation, part of which concerns the responsibility of authors to be aware of the gist their work conveys. In this case, Johnson (2015) gave very few indications of the distinctions between the carcinogens discussed; accordingly, it is of little surprise that in reading his verbatim details, students’ resulting gist was that they were all the same. Hence, given this well-known tendency of media outlets to translate research into engaging and consumable news stories (Angell and Kasirerr1994; Condit 2004; Moynihan et al. 2000), this reminds us of the importance for curricula to foster a critical lens in the classroom.

The previous discussion applied to 42% of respondents—those in the A1A1 variation, believing the author had sufficient evidence based on his discussion. About 22% of total respondents believed the author had sufficient evidence for additional reasons. Among these students, nine (or 6% of total respondents) specifically noted that the author’s argument was well-known and common sense (the A12AB12 variant). The warrant in this variant is that well-known things are valid to employ in an argument, and the backing is that such things are true. Just like the warrant and backing of A1A1, this resource has the potential to be problematic, but nonetheless seems sensible from an evolutionary perspective (Kuhn 2012). Below I include a quote from a typical response in the A12AB12 category:

Processed meats are bad for you and lots of other things in life put you at risk but you can’t live your life being afraid of those things. The author cites the World Health Organization, and much of the other evidence is common knowledge. I believe this evidence is sufficient.

Just as with the A1A1 group, here we see the student has the gist that processed meats are “bad for you,” just as other things mentioned in the article are. Though the student does not state outright that they believe all are equally risky, this is certainly a plausible inference based on the student’s response. What distinguishes this student from those in the A1A1 group is the leveraging of “common knowledge” to argue one’s point. In addition to resorting to common knowledge, some students referenced the author’s use of statistics and citations as specific evidence in favor of his argument. Nine students (or 6% of total respondents) fell into the A1A13 category, as they noted the author’s use of references and statistics as specific evidence that the argument was valid. The warrant and backing behind these are problematic, because sources, statistics, and “official” reports are not necessarily credible. I discuss this in further detail in the coming section about a similar “B” variant.

Taking a step back, a result I take away from this section is that those who argued that the author had sufficient evidence tended to anchor their responses under a warrant and backing that presumed truth in the author’s writing, in common knowledge, or in statistics and references. That is, respondents tended to trust other sources. With that said, a complicating ingredient here is the potential confounding factor of prior agreement with a message; this makes it difficult to claim that students solely relied upon one of the aforementioned resources to come to the conclusion that the author had sufficient evidence. This point merits further discussion, which I include in Section 6.3. Regardless of what becomes of this potential confounding factor, from a statistical literacy lens, the results of this section call for an increased emphasis on having a critical lens as one interacts with messages that appear to be true.
6.2. Partial or Insufficient Evidence From the Author: The “B” and “C” Varieties

Not to be discounted, about 36% of respondents felt that the author’s evidence was lacking in some way, with roughly 28% of all students feeling he had some—but not enough—evidence for his claim. The other 8% felt he had no evidence for his claim at all. The largest subset of the former group, about 23% of total respondents, fell into the B14AD13 category. The basic sentiment of those in this category was that while the author made quality points in describing why we should live life to the fullest and without concern for studies that discuss carcinogens, there were simply not enough references or “hard” evidence to adequately support his argument. What makes this group interesting is that while an earlier group felt the author did present enough references and statistics, this group felt he did not. The warrant here is that quality arguments include more than just persuasive rhetoric—they also contain references and statistics, the backing of which is that such references are to be trusted.

Two example responses in this category are below:

The article had no evidence as in numbers, just more so stating how these things all have an effect and that something is going to kill us one day regardless.

There are no hard facts or references. I believe the only “evidence” is persuasion. Persuasion that yes, it’s possible there are many things that could cause your death to come sooner than expected, but a lot of things are dangerous. That doesn’t mean you just shouldn’t do it. I don’t know if I would call it sufficient.

Now, while respondents in this group—from an “expert’s” perspective—might have taken a better, more critical lens to the piece than that used by those who believed he had sufficient evidence, again we run into the potentially problematic assumption that references and statistics are to be trusted. On the one hand, it might seem blasphemous to quibble over a student’s searching for references, but on the other hand, a major facet of statistical literacy is to critically evaluate statistical claims, many of which often come in official, published formats. While promising, the notion that “hard” evidence is immediately more trustworthy than qualitative argumentation is a manifestation of our reliance on numbers as authoritative sources in our lives (Cohen 1999; Porter 1995). In spite of respondents’ note that the author did not have sufficient evidence, it is telling to again consider the gist surrounding carcinogens that students took away from the article. To that end, the first student quoted above noted “these things all have an effect” and the second that “a lot of things are dangerous.” The gist from these statements is that all the things the author discussed—including car exhausts, uranium, the sun, and processed meats—are dangerous in similar ways. Given that this is not true, again we run into support for the notion that authors should be wary of the gist their work reflects.

The remaining individuals who responded, those who felt the author had no evidence for his claim, comprised a minority of respondents in the “B” variant, with 12 individuals (8% of all responses). To a large extent, this group helps a similar belief as those discussed above, in that they felt the author had no evidence for the claim, as the majority of his discussion of carcinogens lacked references or insightful statistics. Below is a quote from a response akin to those in B14AD13 variant:

6.3. Sufficiency and Agreement

The discussion thus far has centered on the various resources students in this study called upon to evaluate the author’s argument. We saw that with the sufficient evidence group—those with variants beginning with an “A”—the main resource students used was the author’s words, with some coupling that with either his appeal to common sense or his links to various sources. With the “B” variant, a significantly smaller group than the one above, respondents called out the author for the dearth of “hard” evidence in the piece, noting his opinion alone was not quite sufficient for demonstrating his point. These students drew upon their presumptions about what qualifies as sufficient evidence for a valid argument. Finally, respondents within the “C” variant followed in a similar fashion, but instead noted that the argument was wholly incomplete, drawing upon
similar presumptions; only three in this group—those whom I quoted earlier—also used critical reasoning to interpret the author’s claims. With all of this said, a striking finding from these results is that notwithstanding their evaluation of the author’s evidence, most students still agreed with his claim; that is, whether or not he had evidence supporting his view, it still seemed logical that one should live life without regret, giving no credence to the scientific reports we consistently hear about.

As Table 4 demonstrates, 100% of students who believed the author had sufficient evidence for his claim also agreed with it. Given the high agreement rating, this raises the question of how the two variables interacted in the study. On the one hand, it is logical that if one believes another has argued their point well, then they are likely to agree with such a point; conversely, though, is it possible that individuals felt he had sufficient evidence because they agreed with what he stated? The notion of confirmation bias (Cox and Popken 2008) suggests so, but this is a speculative claim that I can only examine with further research. An example quote below from a student’s response portrays the difficulty inherent in trying to tease out the direction here of causality, if any:

To support its claim of indulging in moderation (which he does), the article begins by warning the reader of all the harm that red meats and the sun can cause you. However it quickly becomes overly cautious in its warning and it becomes sarcastic. This shows how silly it is to be overly cautious and not to live your life, but also to be careful to some extent. I do not believe this article will affect any of my habits or opinions, I already sort of had this mindset when it came to my life. The red meat caution was newish information but nothing groundbreaking to me.

In this student’s response, it is unclear if the agreement with the author’s message (“I already sort of had...”) influenced the student’s evaluation of the argument itself. For now, given that confirmation bias is a known phenomenon, it is worth noting that prior agreement is likely a resource students draw from in evaluating evidence.

Considering the other levels of sufficiency evaluation, we see that even among the respondents who noted his evidence was not fully sufficient, 93% (or 39 out of 42) still agreed with his message. This pattern waned with the smallest group, those who felt he had no evidence, but nonetheless the percent in agreement was 67% (or 8 out of 12). These latter two groups suggest that one’s agreement with another’s message is not wholly dependent upon their evaluation of that person’s evidence. Below is an example quote from a student in the “B” variant who agreed with the author’s message:

There is evidence of the proven facts that red meat, sun exposure, and automobile exhaust are all linked to causing cancer. I believe this evidence is somewhat sufficient because I have known these things cause cancer but I don’t believe they provided enough data to really support their claim. I’m aware of these cancerous causing things but that doesn’t mean I am going to change the way I live life because of it. I hold the same views as the author.

Here the student highlights well that one can criticize another’s evidence while still believing their overall message. This was general pattern among the students in the groups with variants beginning with “B” and “C”; with that said, it is worth noting that the three students who gave responses I viewed as “expert” all said they did not agree with the author’s message. This trend begs the question of how an article’s validity relates to a reader’s expertise and agreement (or lack thereof).

My purpose here in including this section of the discussion is not to discount students’ beliefs or label them as deficient. Indeed, the notion of living one’s life to the fullest is both intuitive and difficult to argue with. Moreover, it is certainly easy to imagine one being overwhelmed if they read (and followed) all the advice that news outlets give in reporting scientific studies. The obvious caveat, though, is that—assuming one wants to continue living—she or he must also be cognizant of living to the fullest in the future. The implication of this in the context of this study is that one should not ignore the reality of risks to one’s health—doing so reflects a troubling, antiscience attitude (Maddox 1995). Given that this is precisely what Johnson (2015) endorses, an expert is likely to disagree with the author’s message. Unfortunately, because the majority of students (145 out of 152) agreed with this view, this adds an additional layer of concern to what I have already discussed. That is, on top of the problematic assumptions inherent in some of the resources students used in evaluating the author’s argument, it may be the case that agreement with the author’s intuitive message was more important anyway.

7. Conclusion and Steps Forward for Educators

Research is a complex endeavor—a fact that often gets lost in translation when news outlets report published findings (Angell and Kasirer 1994; Condit 2004; Moynihan et al. 2000). My purpose in completing this study was to examine how a particular group of college students, those in Quantitative Literacy at Michigan State University, reasoned with and evaluated an opinionated article about risk in relation to one’s health. The large size of this group, coupled with the flexibility we have in shaping the course’s curriculum, were my primary motivation for examining this group in particular. The claims I have made throughout this project concern the students of the study, so any generalizations are left to the reader as they feel appropriate in making sense of the research for their own context.

The article students responded to was precisely the type wherein research was lost in translation; the basic message, that we need not concern ourselves over the steady influx of studies about our health, assumed that various things we hear about consistently (e.g., walk more, drink less alcohol, etc.) all contribute in an equally positive way to our overall health. The reality is that this is not the case, and accordingly, “experts” would see past this oversimplification. But how would students interact with such an intuitive message? Among other things, this is a matter of statistical literacy, which includes applying a critical lens to day-to-day messages of this type. My prediction was that—given the lack of standards which support such a critical lens, coupled with the intuitive message of the article—students would follow suit with the author.

With that said, this project was not about pointing out students’ “mistakes”; rather, my research question was what assumptions students drew upon to develop their evaluation of the sufficiency of the author’s evidence for his claim. To that end, with coding for students’ claim, data, warrants, and backings, I was able to examine the nuances of how students tended
to make sense of the author’s claim. For the 64% of students who felt he argued his claim well, the driving rationale was that elements of his writing—the facts used, its appeal to common sense, or use of references—were what made his evidence sufficient for the claim. All the while, those same elements were what others (the remaining 34%) called upon as rationale for why the author had not provided a valid argument.

Woven throughout the responses—whether claiming his evidence was sufficient or not—were assumptions about the trustworthiness of such elements, including that published writing, common sense, and statistics are to be trusted. The notion that published articles, from reputable sources or not, are inherently more trustworthy is problematic. So too is the notion that we can immediately view one’s argument as valid simply because of its inclusion of sources or statistics. And while these resources (coupled with their warrants and backings) are innocent enough to draw upon, they are only helpful insofar as they can guide one in search of the “truth”—they do not compensate for critical thinking. Only three respondents engaged with the article in a manner we might claim an “expert” would. Certainly, the author’s citations were scant and commentary pedestrian, but none of these things was the deal-breaker for these three respondents—the author’s logic was.

Moving beyond the three “expert” responses, a striking finding was that notwithstanding respondents’ evaluation of the author’s evidence, 145 out of 152 still agreed with the author’s message. That is, in spite of whether they felt the author argued his point well, most still agreed with his message. This disconnect between agreement and evidence evaluation merits further exploration. Indeed, while it seems logical that students might agree with the author’s message without explicitly stating why the author had not provided a valid argument. All the while, those same elements were considered sufficient evidence, unclear is the extent to which the converse relationship exists. To boot, given that 36% of students felt the author lacked evidence and yet still agreed with his message, one must question how much such resources students draw upon even matter. To that end, future work—which need not be of this demographic or content type—should hone in on the relationship between agreement and evidence evaluation. Other questions also arise from these findings, including: What role did the article’s lack of quantitative statements play in the way students responded to the article? Would other topics, such as those that students know little about or have little investment in, yield different responses? What would students’ responses have looked like at the end of the semester, or even years later?

Regardless of the answers to the above questions, the results of this article suggest that, assuming curriculum choices have a direct impact on student learning, we should foster an awareness of scientific complexity and one’s own self-bias throughout curricula that aim to foster quantitative or statistical literacy. It is not sufficient to teach students that media sources or quantitative statements are inherently valid or invalid—we must include active practice in interrogating the claims that we and others make (American Statistical Association 2016). While such interrogation can and should take place in English or science classrooms (for example), it should be a practice we foreground in mathematics and statistics classrooms.

As statistics and mathematics educators, I believe we can do this in a number of ways. One is to begin lessons or activities with specific moments dedicated to asking students to share their existing understandings or beliefs about a topic; as a lesson or activity comes to a close, we can have students reflect again, but on how their understandings are different or have shifted. This strategy calls attention to students’ lived experiences, and foregrounds the reality that students will exit the classroom with beliefs about that topic. In relation to that strategy, another is to choose datasets and topics for which students are likely to face dissonance as they examine it from a statistical perspective. Indeed, if we only use examples that are trivial (e.g., the correlation between ice cream sales and drownings), distant from students’ lives (e.g., widgets at a factory), or easy to agree with (e.g., climate change data), students may leave the course having little reason to believe that they should question statistics or their own preconceived notions about a given topic. Hence, not only should we have students foreground their own experiences and beliefs in thinking about a topic, but we should also be intentional in choosing topics themselves. One strategy I have found particularly useful for fostering dissonance and engagement is to have students pick media articles containing quantitative statements for us to examine as a class; depending on the class size, this often means having one or two students present at each course meeting. One reason that this strategy is especially useful is that the topics students choose tend to be those that they have strong feelings about. By modeling at the beginning of the semester the type of interrogation and self-reflection I go through in reading media articles, I find that students tend to engage similarly as the semester goes on. Ideally, this results in students potentially viewing topics of interest through a more critical lens. A goal for future work is to examine whether that is actually the case; I intend to use the results of this study to examine how the Quantitative Literacy course impacts students’ responses to a variety of articles.

Acknowledging and then grappling with complexity is a critical practice for all of us in the 21st century; as educators, we can be leaders in fostering this practice.

Acknowledgments

The author is grateful to Michigan State mathematics educators—including Kristen Bieda, Vincent Melﬁ, Jeffrey Craig, and Ralph Putnam—for their assistance in the formulation and execution of this project. The author is also indebted to Bronlyn Wassink for allowing him to obtain data from students in her course.

References

American Statistical Association (2007), Guidelines for Assessment and Instruction in Statistics Education (GAISE) Report: A Pre-K–12 Curriculum Framework. Available at http://www.amstat.org/education/gaise/gaisep-12_full.pdf

— (2016), Guidelines for Assessment and Instruction in Statistics Education: College Report 2016. Available at http://www.amstat.org/asa/files/pdfs/GAISE/GaiseCollege_Full.pdf

Angell, M., and Kassirer, J. P. (1994), “Clinical Research—What Should the Public Believe?” New England Journal of Medicine, 331, 189–190.

Bennett, A. (2016, April 5), “Red Wine is Actually Good for You Because it Could Prevent Heart Disease,” The Mirror. Available at http://www.mirror.co.uk/news/technology-science/red-wine-actually-good-you-7691136

Bieda, K., Sela, H., and Chazan, D. (2014), “You Are Learning Well My Dear’ Shifts in Novice Teachers’ Talk about Teaching During Their Internship,” Journal of Teacher Education, 66, 150–169.
