SMARTER COOKIES

An inquiry-based project to examine statistical claims encourages students to become more savvy media consumers.

Jeremy F. Strayer and Michael Todd Edwards

A news story claiming that the cream filling of a popular "double" cream sandwich cookie is not really "double" went viral in fall 2013. A high school mathematics teacher posted a blog entry describing how he and his students measured 20 cookies, analyzed the data, and concluded that the double cream cookies had only 1.86 times the filling of the regular variety (Anderson 2013). A media frenzy ensued. Countless blog sites and large media outlets—including CNN, *The Huffington Post*, ABC, *Time*, and *The Atlantic*—picked up the story, leading many to conclude that the double cream cookies "aren't all they're stuffed up to be" (Perreira and Payne 2013).

The news story broke while we were teaching statistics topics in entry-level undergraduate courses—an introductory statistics course and a methods course for secondary school mathematics teachers. As mathematics educators, we found the report both encouraging and troubling. We are genuinely excited whenever we see real-world applications play a prominent role in school classrooms. Nevertheless, investigative reports highlighting the use of very small, nonrepresentative samples and dubious data collection strategies are disturbing; they reinforce statistical misconceptions and misunderstandings.

In this sense, the story was immediately compelling and timely; it provided us with a vehicle for engaging students in a meaningful investigation of statistical claims in the popular media. Within a few weeks—and after many hours of collaborative planning—we developed a project that encouraged students to implement statistical experiments of their own design while confronting misunderstandings perpetuated by the news story. Here we describe that project—the Cookie Conundrum—through displays and discussions of our students' work. Although we conceived the Cookie Conundrum as one for use with undergraduates or preservice teachers, the project could certainly be

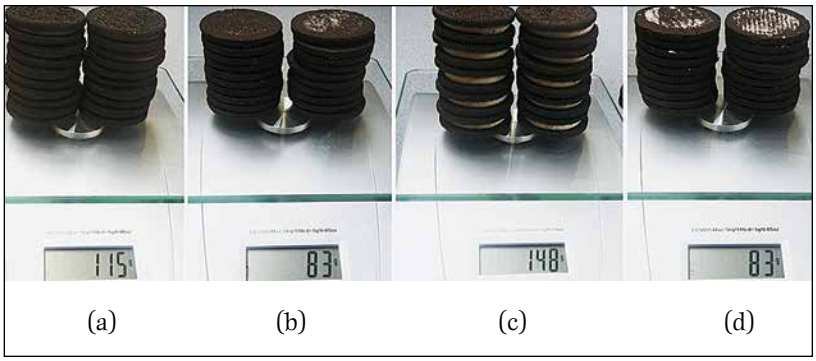

Fig. 1 Students weighed 10 regular cookies (115 g) (**a**), biscuits from the same 10 regular cookies (**b**), 10 double-cream cookies (148 g) (**c**), and biscuits from the same 10 double-cream cookies (**d**).

Fig. 2 Students scraped cream off the biscuit with a driver's license.

adapted for use with mathematics students in the middle school or secondary school grades.

According to the American Statistical Association's *GAISE* report (Franklin et al. 2007), students need repeated opportunities to engage in each of the four components of statistical investigations when learning statistics: (1) formulating questions, (2) collecting data, (3) analyzing data, and (4) interpreting results. We explicitly designed the Cookie Conundrum to engage students in all four GAISE components. In particular, we describe ways in which students formulated their own questions and designed their own research methods to answer their own questions, using appropriate statistical methods to critically examine unfounded claims involving double-cream cookies.

The Cookie Conundrum project consists of three phases. We conceived phase 1 and phase 2 to be completed in approximately one class meeting each, with phase 3 requiring several class meetings. However, we recognize that timing will vary according to the students' prior experiences and content knowledge.

- *Phase 1* (*Replication*): First, students are introduced to the original study. After watching a CNN video highlighting the basics of the news story, small groups replicate the data collection (GAISE component 2) and preliminary data analysis (GAISE component 3), gaining a more detailed understanding of the methods used as they produce comparison data.
- *Phase 2* (*Critique*): Students consider the original cookie investigation and ensuing media coverage more critically. First, they formulate their own research questions (GAISE component 1). Next, they critique the original data collection methods (GAISE component 2) and brainstorm alternatives.
- *Phase 3* (*Reformulation*): Last, students analyze and interpret a revised data set (GAISE components 3 and 4) to answer their own cookie questions.

Next we describe challenges and opportunities that we observed with our students as they engaged in each of the three phases of the project.

PHASE 1: REPLICATING THE ORIGINAL STUDY

We launch the project by showing students a CNN video (http://bit.ly/cnn_cookies) that describes the background of the cookie experiment and the ensuing controversy it prompted. Following the method reported in the video and the teacher's original blog post (Anderson 2013), our students weighed ten regular and ten double-cream cookies, with and without cream, as illustrated in **figure 1**.

For each type of cookie, small teams of three to four students separated cream from biscuits, simultaneously twisting the top and bottom in opposite

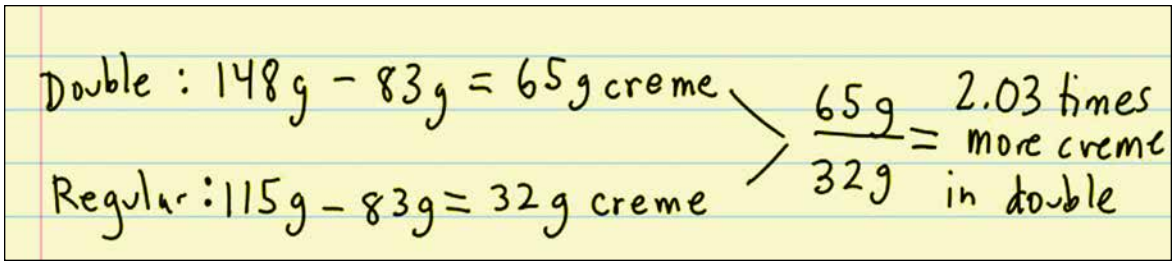

Fig. 3 A student calculated a cream ratio for regular and double-cream sandwich cookies.

directions. This approach left all the cream on one biscuit. As **figure 2** suggests, teams scraped as much of the remaining cream off the bottom biscuit as possible using readily available materials.

Next, teams subtracted the mass of 10 cookies without cream (i.e., the biscuits) from the mass of the same cookies with cream. This calculation provided an estimate of the mass of the cream, which students used to calculate the cream ratio for 10-cookie samples. Student work is shown in **figure 3**.

Similar calculations by five other teams are provided in **table 1**. As the table indicates, the teams generated conflicting results. Some cream ratios were larger than 2, whereas others were less. This variation, coupled with remarks from the cookie manufacturer stating that "our recipe . . . has double the stuff, or cream filling, when compared with our . . . original" (Perreira and Payne 2013), brought into question both the design of the reported investigation as a statistical study as well as the subsequent media coverage. Confronted with additional data, many students questioned whether newsroom editors and correspondents had taken the time to perform the cookie experiment before reporting their findings.

A prospective high school teacher in our methods course shared the following observations: "It is absolutely amazing to me that the cookie story went viral based on data obtained from 20 cookies in two bags. In 15 minutes, we obtained results that contradicted details presented in the online article. Personally, it's difficult for me to believe that others would report that double cookies have less than twice the icing using data collected this way" (Student A, personal communication, Nov. 21, 2013).

The simple act of repeating the data collection (GAISE component 2) and mathematical analysis (GAISE component 3) allowed our students to experience natural variation that exists in repeated processes, providing an important point of reference as they developed statistical data analysis methods in subsequent phases of the project.

PHASE 2: CRITIQUING THE ORIGINAL STUDY

In the next phase, we asked students to reflect on weaknesses of the original cookie investigation as they critiqued the data collection procedures from phase 1 and formulated their own research questions (GAISE component 1). To help focus classroom conversations and generate discussion, we divided students into a number of small "question" and "data" groups, each consisting of three to four students. Individual question groups were asked to state the aim of the original study as a one-sentence research question. Data groups brainstormed concerns related to data collection procedures. After some time, the small groups came together, forming large question and data groups. The larger groups shared initial ideas, consolidating their findings as they prepared a short presentation for the whole class. Below we provide observations that students shared in their whole-group presentations.

Question Formulation

The following research questions were generated by question groups in our classes.

1. Does a double-cream cookie have double the cream of a regular cookie?
2. Do double-cream cookies have double the cream when compared with regular cookies?
3. Is the average amount of cream in double-cream cookies equal to 2 times the average amount of cream in the regular cookies?

These three questions are related, but the scope of each query is markedly different. Question 1 is not a statistics question as defined by GAISE because it is not answered using data that vary; it is answered by comparing only two individual cookies. On the other hand, questions 2 and 3 are answered using multiple cookies of the same type, and the amount of cream in each will differ, even if only slightly. Moreover, questions 2 and 3 suggest generalization to the broader group of all regular and double-cream cookies. In addition, question 3

Table 1 Measures for 10 Regular and 10 Double-Cream Cookies

Group	Regular Cookies with Filling	Regular Cookies without Filling	Double-Cream Cookies with Filling	Double-Cream Cookies without Filling	Cream Ratio
Group 1	115	82	146	82	1.94
Group 2	116	83	146	82	1.94
Group 3	116	82	147	82	1.91
Group 4	114	85	146	81	2.24
Group 5	114	84	147	83	2.13

suggests comparing the mean (i.e., average) amount of cream, a population parameter, in the two cookie types. In whole-group conversations, students noted that questions 2 and 3 "ask the same sort of question" but that question 3 suggests "what to do" more clearly than question 2.

The comparison of student-authored research questions helped students—both those in methods courses and those in introductory statistics courses—recognize the importance of posing questions succinctly, in ways that are clearly measurable and testable. Further, the task emphasized that statistical studies require asking statistical questions—those that can be answered using data that vary.

Data Collection

The concerns of our data-group students focused on limitations associated with sample size. In addition, students identified problems concerning data variation, instrument precision, and random sampling. These concerns are summarized below.

Sample size and variation. The data-group students agreed that sample size in the original investigation was "too small," although they struggled to articulate what "too small" meant. Several groups recommended repeating the original investigation with 50 of each cookie rather than 10. This recommendation led to a series of interesting questions. For instance, why not 100? Why not 500? How many cookies are needed? Questions such as these sparked animated debates in each of our classes. Although some students agreed that 50 cookies would provide a more accurate overall measure of regular and double-cream varieties than 10, others understood that statistical questions need to be answered using data that vary (Franklin et al. 2007). By challenging students with the question "How many measurements were actually taken?" students came to the conclusion that finding the average weight of 10 or 50 (or even 5 million) cookies provides a single measure of weight and, as such, fails to reveal how weight in a sample tends to vary from one cookie to another. When we asked students, "Is taking one measurement enough?" they responded that the original investigation failed to provide information regarding variation of cream weight "from cookie to cookie" for each cookie type.

Instrument precision. A number of students in both classrooms expressed concerns regarding the precision of the scales used to measure cookies. Note that the scales in **figure 2** measure mass to the nearest gram. Hence, conflicting results may be due in part to instrument precision. For instance, one methods student asked her classmates to consider the measurements in the first row of **table 1**. She pointed out that measures of 146.4, 81.5, 114.5, and 82.4, when rounded to the nearest gram, yield a cream ratio identical to that provided in the table: 1.94. Yet the unrounded measures yield a ratio larger than 2 (i.e., 2.02). Students' observations such as these suggested the need for scales that measure to the nearest tenth or hundredth of a gram.

Random sampling. Last, our students noted that the cookies in the original investigation came from the same package. As such, the cookies were manufactured on the same day, at the same time, on the same machinery, at the same processing plant. As one preservice teacher commented, "What if the icing station at the double-cream factory was dispensing less cream than usual on that day? What if regional differences existed at different factories?" (Student B, personal communication, Nov. 21, 2013). By failing to randomly sample cookies, the original investigation failed to account for possible variation associated with the cookie manufacturing process.

PHASE 3: REFORMULATING THE STUDY

In the final phase, we asked students to reformulate the original study. In small teams, they developed their own research questions (GAISE component 1) and described their own procedures for collecting (GAISE component 2), analyzing (GAISE component 3), and interpreting (GAISE component 4) cookie data (see the **sidebar**). Phase 3 culminates with students sharing their analyses and findings in short presentations to classmates. Here we highlight issues critical to the success of this phase. Specifically, we describe how students addressed their concerns associated with data collection, sampling, precision, and variation. Last, we highlight a statistical investigation that was popular among introductory statistics students that aimed to compare regular and double-cream cookies more rigorously than the original cookie investigation.

Sampling Concerns

Students commonly believe that factors such as sample size are more important than representativeness, so at the beginning of phase 3 we emphasized the importance of collecting data from a sample that adequately represents the population. Through online research, students learned that this particular brand of regular and double-cream cookies is manufactured at two different factories in the United States. They also learned that strict controls

Variable Name	estimated mean	degrees of freedom	t	p	confidence interval percent	confidence interval of difference
original2x.no.outliers double.no.outliers	6.334379 6.514118	224.6721	-4.734654	3.885102e-06	95	-0.2545469 -0.104931

Fig. 4 Student results of a *t*-test comparing the mean of twice the mass of regular and double-cream data sets showed evidence that double-cream cookies have more than twice the cream of regular cookies.

are implemented to ensure uniformity at each factory through a Total Quality Management (TQM) process (Manley 2011; Lusas and Rooney 2001). On the basis of these findings, students concluded that it was reasonable to assume that cookies purchased from a few different stores at different times could adequately represent the population of all regular and double-cream cookies.

Precision and Variation Concerns

Students determined that a precision scale capable of measuring mass to the nearest hundredth of a gram should be used to collect mass data from individual cookies. Unlike the original investigation, which provided students with a single measure of mass, the revised approach should provide students with access to data that vary.

Data Collection Concerns

Because purchasing packages of cookies and precision scales for multiple groups is cost prohibitive and because significant time is required to collect individual cookie data, we elected to collect cookie data that our students could subsequently use in their follow-up work. We measured the mass of cream and biscuits for 180 regular and 180 double-cream cookies purchased from different stores in two different states. Our data collection was informed by concerns that students expressed in phase 2. We provided access to the data from an online spreadsheet available at http://bit.ly/cookie-data.

Student Investigation

After concerns regarding data collection, sampling, and precision were addressed, small groups of students formulated their own research questions in phase 3. One group's research question is provided in question 4:

4. Is the mean mass of cream in all double-cream cookies equal to twice the mean mass of cream in all regular cookies?

To address this question, the student group performed a hypothesis test with the provided data after deciding to identify and remove outliers. Using a two-sided (i.e., two-tailed) *t*-test, the group determined whether there was evidence that the mean of twice the mass of the regular cream is different from the mean double–cream mass. Using R, a freely available statistics analysis software (go to http://www.r-project.org), students analyzed data from http://bit.ly/cookie-data to generate the results provided in **figure 4**.

In their written work, students summarized their interpretations of these results, connecting the *t*-statistic back to regular and double-cream cookies. After a class discussion of these results, one team member wrote the following summary: "The *t*-test gives a *p*-value of 0.000004. This means that if we randomly collected 360 cookies many, many times, where the double-cream cookies are truly twice the regular cookies, only 0.0004% of those samples would produce the difference we observed in our data. Because this probability is so small and because the mean cream weight of the double-cream cookies in our sample was larger than twice the mean cream weight of our regular cookies, we can conclude that, on average, the

THE COOKIE CONUNDRUM: PHASE 3

Formulating a Question

1. What question does your group wish to answer with the revised cookie data? State the aim of your follow-up investigation as a one-sentence research question.

Collecting Data

2. What issues should you address when collecting data that will help you answer the question that you formulated?

Analyzing Data

3. Analyze the cookie data available at http://bit.ly/cookie-data, specifically addressing the research question 1. Describe the steps of your analysis in several complete sentences. For each step, provide a brief justification for its inclusion in your study.

Interpreting Results

4. What do the results of your analysis appear to indicate? Justify your conclusions using evidence from your data analysis. Include data displays (e.g., graphs, output from statistical software, etc.). In several complete paragraphs, give thoughtful interpretations of these displays and the other data analysis that you performed.

cream in double-cream cookies for this brand is actually *more than twice* the cream in the regular cookies" (Student C, personal communication, Nov. 21, 2013).

These findings were consistent with those of other student groups in both our classes. Without exception, students found reasonable evidence to conclude that double-cream cookies, in fact, had at least twice the cream of regular cookies. Arguably more important, students in both classrooms recognized that the media failed to tell the whole story.

NEXT STEPS AND IDEAS FOR FURTHER STUDY

Engaging Students at Various Levels

Connecting content from the mathematics classroom with news stories in the popular media provides students at any level with opportunities to use tools that are accessible to them to strengthen their understanding of statistics. Students in earlier grades can analyze cookie data using comparative displays (box plots and stem-and-leaf plots); measures of center (mean, median, and mode); measures of spread (variance, standard deviation, and interquartile range); and position (quartiles and percentiles). More advanced students can use resampling and bootstrapping methods or more traditional hypothesis testing such as that presented in the previous section.

Exploring Other Topics

The Cookie Conundrum project demonstrates ways in which new teaching and learning ideas can be generated in the space created by the overlap of popular media, teacher blogs, and instructional objectives. Certainly, such investigations are not limited to the study of cream-filled cookies. The three-phase model explored here may be used to investigate a myriad of other media claims in the mathematics classroom. Indeed, accounts in the popular media are rife with categories of examples worthy of exploration in school classrooms. Consider the following news stories and advertisements:

- A teen recently claimed that the U.S. government would save $400 million by changing fonts on government documents from Times New Roman to Garamond (Miller 2014). Is this true? What mathematical and statistical claims are made in the media story?
- A news report claims that students are not attending school dances because of their social media use (Pawlowski 2014). Is this claim warranted? What role did data, mathematics, and statistics play in this story?
- A popular sandwich chain was recently sued for selling footlong subs that were shorter than expected (Arumugam 2013). Was such a lawsuit warranted? What role might statistics play in the controversy?
- A popular brand of gum is marketed as producing "double" bubbles. Does the size of bubbles generated while chewing this brand of gum result in bubbles that are any larger than competing brands of bubble gum?

The Cookie Conundrum project provides powerful learning opportunities for students and their teachers. First and foremost, the project empowers students to question news reports that make use of mathematics and statistics. Because today's students live in an age of unparalleled access to data and technology, they need opportunities to grow as critical consumers of information. Classroom activities that encourage students to analyze data and statistics in media literacy contexts—news articles, blog posts, video analyses, statistical reports—address calls for instruction that is both "robust and relevant to the real world" (CCSSI 2010). As students complete the Cookie Conundrum project, they are surprised to uncover reporting that misrepresents facts. Moreover, they understand that journalists and readers need solid statistical understanding to report and respond to quantitative studies in an educated manner. Indeed, the project illustrates the need for quantitative literacy among all our citizenry.

Teachers who provide opportunities for students to make meaning of the complexities found in the four components of statistical studies (as outlined by the GAISE report) will help those students decipher warranted and unwarranted statistical claims—whether the claims are found in the news media, government reports, or the workplace. When we neglect to help students make meaning of these components, we unwittingly provide them with an artificial sense of the statistical design and investigation process. Through projects like the Cookie Conundrum, students are challenged to reason with and about data to reach well-reasoned conclusions in easily accessible and highly motivational contexts.

REFERENCES

Anderson, Dan. 2013. "Oreo: Original vs. Double vs. Mega." *A Recursive Process: Math Teacher Seeking*

Patterns. http://bit.ly/cookie_blog.
Arumugam, Nadia. 2013. "Why Lawsuits over Subway's Short Footlong Sandwiches Are Baloney." *Forbes*. http://bit.ly/frivolous_lawsuits.
Common Core State Standards Initiative (CCSSI). 2010. Common Core State Standards for Mathematics. Washington, DC: National Governors Association Center for Best Practices and the Council of Chief State School Officers. http://www.corestandards.org/wp-content/uploads/Math_Standards.pdf
Franklin, Christine, Gary Kader, Denise Mewborn, Jerry Moreno, Roxy Peck, Mike Perry, and Richard Scheaffer. 2007. *Guidelines for Assessment and Instruction in Statistics Education (GAISE) Report: A Pre-K–12 Curriculum Framework*. Alexandria, VA: American Statistical Association.
Lusas, Edmund W., and Lloyd W. Rooney. 2001. *Snack Foods Processing*. Boca Raton, FL: CRC Press.
Manley, Duncan. 2011. *Manley's Technology of Biscuits, Crackers, and Cookies*. 4th ed. Cambridge, UK: Woodhead Publishing Limited.
Miller, Michelle. 2014. "Teen Says Font Change Could Save Government Millions." *CBS News*. http://cbsn.ws/QyCtGi
Pawlowski, A. March 14, 2014. "Is Social Media Killing the School Dance? Schools Cancel Dances as Teens Opt Out." *Today*. http://tinyurl.com/SMKSD
Perreira, Michaela, and Ed Payne. 2013. "Oreos High School Experiment: Double and Mega Stuf Filling Doesn't Add Up." CNN U.S. Website. http://bit.ly/cnn_cookies

JEREMY F. STRAYER, jeremy.strayer@mtsu.edu, is an assistant professor of mathematics education at Middle Tennessee State University in Murfreesboro. He is interested in statistics education and integrating technology into standards-based mathematics instruction. **MICHAEL TODD EDWARDS,** edwardm2@miamioh.edu, is an associate professor of mathematics education at Miami University in Oxford, Ohio. His research interests focus on the teaching and learning of school mathematics with technology and writing as a vehicle to learn mathematics. The authors extend special thanks to Greg Foley and the QUANT project for making us aware of the CNN news story and to Grace E. Hoyt and Frances Hoyt for their inspiration and invaluable assistance in this project.