

Undergraduates' Graph Interpretation and Scientific Paper Reading Shift from Novice- to Expert-like as a Result of Participation in a Summer Research Experience: A Case Study

Anne U. Gold, *University of Colorado Boulder*
Rachel Atkins, *North Carolina State University*
Karen S. McNeal, *Auburn University*

Abstract

Research Experiences for Undergraduate (REU) programs often introduce students to scientific research and STEM career possibilities. However, the program impact on students and their research skill development is not well understood. In a case study with 10 REU students, the authors used eye-tracking and self-report data to determine student strategies for reading scientific papers and interpreting graphs at the beginning and end of the program. The strategies of REU students and science experts were then compared. The REU students changed their strategies and performed more like experts at posttest. These findings indicate that, during the REU, students acquired expert-like strategies necessary to engage with scientific articles and extract key information from graphs. The study demonstrates that eye-tracking can document skill growth in REU students.

Keywords: *community college students, eye-tracking, Research Experiences for Undergraduates, scientific graph interpretation, scientific paper reading*

doi: 10.18833/spur/5/2/2

The 2018 National Academies report on undergraduate STEM indicators highlights the need to involve students in authentic STEM practices (National Academies 2018). Research Experiences for Undergraduates (REU) programs offer such opportunities. Many benefits of REU programs have been described in the literature—from increased enrollment in STEM majors and continuation

toward a graduate degree (Brewer and Smith 2011; Graham et al. 2013; Cooper, Jabanoski, and Kaplan 2019) to gaining professional skills (Hunter et al. 2007; Thiry, Laursen, and Hunter 2011), developing a science identity (Seymour et al. 2004; Russell et al. 2007; Hunter et al. 2007; Thiry et al. 2011; Weston and Laursen 2015; Cooper et al. 2019), and research skill building (Seymour et al. 2004; Hunter et al. 2007; Thiry et al. 2011; Linn et al. 2015), among others. The National Science Foundation (2019) describes research experience programs as “one of the most effective avenues for attracting students to and retaining students in science and engineering, and for preparing them for careers in these fields.” REU participants tend to express high satisfaction with and strong learning gains from their experience. However, using self-report data from students as an impact measure for REU programs has been criticized, and researchers have called for reliable, robust assessment of students’ learning gains from REU programs to evaluate the efficacy of REU programs, justify their cost, and help improve their design (Linn et al. 2015).

Here, findings are reported from a case study with a cohort of REU students ($n = 10$) that examined student eye-movements as a representation of the strategies employed by students in reading scientific papers and interpreting graphs before and after their REU experience at the beginning and at the end of the program and compared them to experts. A cohort of experienced researchers (“experts”; $n = 16$) was recruited to participate in the same exercises as a comparison group. Study participants were recruited from the Research Experiences for Community College

Students (RECCS) program—a nine-week summer REU program that pairs Colorado community college students with researchers from a large research institute or federal research lab. The authors describe below self-reported data on students' perceived changes in confidence in and ability to read scientific papers and graphs as well as report on eye-tracking data and responses from students and experts while they engaged in paper reading and graph interpretation—all through the lens of the following two research questions:

1. In which ways does participation in an REU program impact students' strategies in engaging with scientific papers and graphs?

More specifically, eye-tracking technology was used to explore how students approach graph and scientific paper reading at the beginning and the end of their research experience, how students differ from or are similar to experts, and whether or not REU student participants move toward an expert-like approach for these tasks.

2. In which ways do students' self-reported confidence and ability of engaging with scientific papers and graphs change through the participation in an REU program?

Importance of Reading Scientific Papers and Interpreting Graphs

Scientific publications are the primary means of communication between scientists. Paper and graph reading are critical skills that researchers use in their work; thus, improving students' strategies for and skills in reading scientific papers and interpreting graphs is an important steppingstone on the path to becoming a scientist (Willmott, Clark, and Harrison 2003). Scientific papers usually follow a prescribed structure to build the case for the research findings (Dean 2013). Papers usually include an introduction, in which the study is motivated and prior work is discussed, a methods section, a results section (often with images or tables), a section in which the results are discussed in a larger context, and a summary of the take-away messages in the conclusion. Other typical paper components are an abstract, list of references, figures, and tables. Because papers are primarily written to inform the scientific community, scientific papers are often dense, include technical terms and jargon, and are sometimes hard to access for a novice (Willmott et al. 2003). Such professional publications are often read using a variety of strategies by experts such as skimming the text, jumping back and forth within the text, focusing on figures and tables, and comparing tables/figures to text among others (Wyatt et al. 1993). Thus, reading professional papers requires different strategies and skills that differ from typical text-reading skills. Students that are pursuing a scientific career need to learn how to read professional articles as part of their academic training.

Similarly, reading and interpreting scientific graphs is a skill that students acquire over time and with training (Freedman and Shah 2002). As with professional articles, graphs have distinct elements (e.g., axes, units, title, key, data, caption) that are important to communicate the graph's message. Graph interpretation requires critical thinking skills to interpret and draw conclusions from the graphically represented data (Malamitsa, Kokkotas, and Kasoutas 2008). Students often make systematic errors when asked to interpret graphs (Guthrie, Weber, and Kimmerly, 1993; Leinhardt, Zaslavsky, and Stein 1990; Shah and Carpenter 1995; Shah and Hoeffner 2002). Thus, like reading papers, students need to learn how to interpret graphs as part of their training as a scientist.

The way in which readers allocate their attention during reading tasks depends on characteristics of both text and reader. Attention can be affected by text organization, verbal complexity, and typographical prompts that indicate relations (van den Broek 2010). The reader's attention is also influenced by available working memory and the reader's relevant background knowledge (McNamara et al. 1996). During reading comprehension studies, Mayer (2001, 2005) found that individuals learn better from words and images rather than from words alone ("dual multimedia principle"). According to the theory of dual coding, the use of both pictures and text makes it easier to remember and retrieve information (Paivio 1990). This phenomenon can be explained through neurological processes; during internal processing, the brain draws on different subsystems to generate a text-based model and a picture-based model (Mason et al. 2013). Most scientific papers include a combination of pictures and text; thus, readers need to generate both text- and picture-based models as they encode the information.

Numerous research studies have been conducted to explore how students best learn using a variety of types of images and animations, text, and the combination of text and images (Butcher 2006; Moreno and Mayer 1999; Mason et al. 2013). In addition, readers' attention is affected by their existing background knowledge of the text or graph content (Curcio 1987). Bertin (1983, translated from French by Peterman et al. 2015) identified three critical steps for graph reading: The first step is "External Identification"—students explore external factors of the graph only (e.g., title, axes labels, and scale of each axis). The second step is "Internal Identification"—students explore the internal factors ("data") of the graph (e.g., lines, dots, bars). The third step is the "Perception of Pertinent Correspondences"—students combine the information to understand the content being displayed through the interaction of the external and internal graph features. Only after the completion of these three steps are students able to extract information, identify trends, and draw conclusions (Bertin 1983). Thus, studying the strategies of the initial steps of

graph reading can provide evidence of how novices and experts navigate graph reading and graph interpretation.

Even though graph interpretation and reading of scientific papers are critical professional skills for a scientist, only few studies explored the skills in detail (Glazer 2011). Although assessing comprehension of paper or graph content can be conducted through postsurveys or interviews (e.g., Peterman et al. 2015), studying the different strategies employed by subjects in their paper reading and graph interpretation is more challenging.

Cognitive processes during reading can be investigated by eye-tracking devices (van den Broek 2010). Eye-tracking methodologies have been applied to multimedia learning (Hyönä 2010; Mayer 2010; van Gog and Scheiter 2010) in which the effectiveness of animations (Boucheix and Lowe 2010), complex graphics (Canham and Hegarty 2010), dynamic stimuli (Jarodzka et al. 2010), and cognitive load (Mitra, McNeal, and Bondell 2017) have been examined. Various studies have used eye-tracking to explore the strategies of participants when reading or interpreting scientific papers (Rayner 1998, 2009), graphs (Atkins and McNeal 2018; Ho et al. 2014), combinations of text and graphs (Ho et al. 2014), decision support systems (Maudlin et al. 2020), and maps (Ooms et al. 2012; Ooms, De Maeyer, and Fack 2014). In nearly all of these studies, findings have suggested that experts when compared to novices tend to allocate their attention differently and were able to interpret the provided representations (e.g., maps, graphs, online tools) more efficiently and with greater accuracy (Ooms et al. 2012; Ho et al. 2014; Atkins and McNeal 2018). Of particular interest to this study, Ho and colleagues (2014) employed an experiment that contained a combination of graphics and text (similar to this study) and found that students with higher prior knowledge fixated longer on the graphics instead of the text and that they tended to have more interscanning transitions between the text and graphics, suggesting that students with higher prior knowledge were able to integrate text and graphic information more readily than students with low prior knowledge. As such, eye-tracking is a promising tool to address the research question that aims to understand how experts and novices navigate scientific papers and graphs.

Program Description

Research experiences engage students in working on a clearly defined authentic research project under the guidance of a mentor. REU students conduct an entire research project—from the definition or refinement of the research question to data collection, data analysis, interpretation of results, and presentation of the results. Among the skills gained by REU students throughout their research experience are some that are specific to their individual research projects (e.g., certain lab, fieldwork, or programming skills); however, they also develop skills in reading

scientific papers, as well as reading and interpreting graphs (Yeoman and Zamorski 2008).

The RECCS program offers students from Colorado community colleges a nine-week, paid summer research experience in environmental or earth science. Student researchers are matched with a mentor or a mentor team at a large research institution or in a federal research lab. Throughout the summer, students work toward two program deliverables: a poster presentation and a 10-minute scientific talk.

The RECCS program team provides weekly training for the students in the development of research questions, reading scientific papers, learning graph reading and interpretation, and developing skills in poster and talk preparation (see Table 1). Weekly assignment deadlines are aligned with components of scientific papers, keep the students moving toward their program deliverables, and allow mentors and program staff to provide timely feedback. Reading scientific papers, as well as creating and interpreting scientific graphs, are part of the instruction. Students learn to identify the different components of scientific papers and graphs as well as to dissect the strengths and weaknesses of examples through a total of about 10 hours of dedicated instruction (see Table 1). However, students also read scientific papers as background to their own research topic throughout the summer on an almost daily basis, initially to learn about the project and methods and later to put their findings in the broader scientific context. Mentors often assign papers for students to read (which usually include graphs) and discuss them with the student later. In preparation for their final products, they draft text following the structure of a scientific paper (see Table 1), and once they start analyzing their data, they have to develop meaningful graphs that summarize their findings.

Description of Participants

Participating in the 2015 RECCS program were 10 students from 4 of the 13 Colorado community colleges. Of the 10 students, 7 were female, and 4 were first-generation college students. Six students self-identified as white; two as Hispanic/Latino; two as Asian; and one as Native Hawaiian/Pacific Islander/Filipino/a (one student identified as multiracial). The mean GPA of the students in the semester before the program was 3.7 (SD = 0.37), with a range of 2.8 to 4. All summer research students received a weekly stipend and were expected to work 40 hours per week during the 9-week research experience program. All students completed the program, presented their work in a scientific poster session, and gave a 10-minute scientific talk.

Experts were recruited from a convenience sample of graduate students, researchers, and faculty members. All experts were earth scientists; five of the experts were females, three were graduate students, three were faculty

TABLE 1. Training Activities around Graph and Paper Reading throughout the RECCS Program

Graph reading in training workshop		Paper reading in training workshop	Graph and paper reading in research experience activities
Week 1	Development of simple graphs based on data collection. Presentation of graphs in teams, critiquing of graphs.	Identify components of scientific posters/talks as examples of science communication. Exercise: How to read a scientific paper. Group work. Report out.	Ongoing, almost daily practice in reading scientific papers, reading and interpreting scientific graphs, as well as creating graphs
Week 2		How do scientists communicate?— Analysis of structure of scientific papers/talks.	
Week 3		Reading about and complete assignment of developing research title of paper/poster.	
Week 4	Reading graphs in example papers.	Learn about reference lists and working with references. Reading critical references.	
Week 5		Reading about and complete assignment—method section of paper/poster.	
Week 6		Reading about and complete assignment—results section of paper/poster.	
Week 7	Discussion about creation of effective graphs, practice of making graphs for their own data.	Reading about and complete assignment—results section of paper/poster.	
Week 8	Continue creating graphs from own data. Interpreting graphs from results from others	Reading about and complete assignment—discussion section of paper/poster.	
Week 9	Presenting own data in talk and poster including discussing own graphs.		

members, and two were research scientists. Of those that completed the pretest, four were geoscientists, three were climate/atmosphere scientists, and one was an ecologist. Those that completed the posttest included five geoscientists and three climate/atmosphere scientists.

Study Design and Methodology

In this case study, eye-tracking data was collected from a RECCS student cohort (“novices”; $n = 10$) at the beginning and at the end of the nine-week program, as well as from a cohort of experts ($n = 16$, see details above), while they engaged in graph interpretation and scientific paper reading tasks. Half of the experts ($n = 8$) were randomly assigned the pretest paper, and the other half was randomly assigned the posttest paper. None of the experts was an expert in either of the topics of the article. Data for each of the subgroups (e.g., pretest paper, posttest paper) were averaged. The authors compared the expert performance to validate the methodological approach and found no significant differences between expert performance on the articles; thus, there is no differentiation between the expert groups.

Eye-Tracking

Eye trackers are non-invasive instruments that determine the location and duration of users' visual fixations. Eye movements represent an involuntary response that can be measured and analyzed to determine engagement with a

visual stimulus such as text or a graph (Atkins and McNeal 2018). While engaging with the scientific paper and graph, participants' eye movements were tracked using a Tobii X2-60 eye tracker with a sampling rate of 60Hz. Calibration was completed for each participant at the beginning of the data collection to ensure accuracy, precision, and consistency within participant groups and allowed participants to wear corrective lenses without affecting results. Participants sat approximately 65 cm from the monitor and gazed at the computer screen to view the provided images with an unobstructed view, allowing natural eye movements to be captured. Participants were given a prompt for both the paper-reading and the graph-reading exercise (see details below) and were asked to think aloud and share their thought and analysis process. No specific training on thinking aloud was provided, although the authors explained the concept to the participants. The eye-tracking sessions with participants were recorded and transcribed.

Paper-Reading Task

For the scientific paper reading task, short-format publications with a geoscience focus were selected from the magazine *Science* that included all key components of a scientific publication (i.e., title, introduction, data, a graph of results, and references) and that fit on one screen for eye-tracking feasibility and consistency. For the pretest, students were provided with a different article (Hoorn and

Flantua 2015) than during the posttest (Conrad 2015) to avoid recognition of the content; however, both articles were of similar difficulty, style, and length. They both include one figure and about 1000 words of text and, according to a readability index amalgamator (webfx.com), both articles score an average grade level of 13 (“18–19-year-olds should easily understand the text”).

Participants were asked to explore the article in a way that would enable them to answer comprehension questions (Prompt: “Please read this article, which will be followed by a few short questions.”). Reported here are the data collected while participants explored the paper during a free-exploration reading period. There was no time limit for participants to respond.

Graph-Reading Task

A line graph plotting “Global Greenhouse Gas Concentrations Over Time” (data from US Environmental Protection Agency 2014) was used for the graph-reading task. Following the methodology outlined in Atkins and McNeal (2018), the graph was first shown for a 5-second free viewing period; then, the participants were presented with the question about the graph’s content and were asked to think out loud as they answered the question. There was no time limit to answer the question. The same graph was used for the pretest and the posttest. Given that pretest and posttest were almost two months apart, the research team decided that it would be unlikely that the subjects would remember the details of the graph and that the benefits of being able to compare the same graph for both tests exceeded concerns of memorization of the graph.

Data Analysis

Eye-tracking data were analyzed using Tobii Studio 2012 and Microsoft Excel 2013. The focus was on two metrics: the *visit count* provided insight into how many times a participant visually revisited areas of interest (AOIs) within the article or the graph (i.e., the number of times their gaze left and re-entered an AOI), whereas the *visit duration* metric captured the total amount of time spent by the participant in visually examining an AOI. To systematically analyze the gaze, the scientific articles were split into five AOIs using Tobii software:

1. Title/description
2. Article text
3. Figure
4. Figure caption
5. References

The graph was split into four AOIs:

1. Title
2. Axes
3. Question
4. Data

Each AOI had a different size; therefore, the authors calculated the area of each AOI and normalized the data across each AOI by taking an average metric spent viewing an AOI and dividing it by the area of the AOI.

The students’ utterances from the pre-program and post-program eye-tracking transcripts were analyzed. A thematic analysis of the transcript (Braun and Clarke 2006) was conducted to explore in which ways students narrated the eye-tracking tasks. During the coding of the student data, the authors outlined four broad themes (see Table 2). These themes were the following:

1. Prior experience or training with respect to paper/graph reading (or lack thereof)
2. Self-reported ability/inability in graph/paper reading or comprehension of graph/paper
3. Confidence (lack thereof) in graph/paper reading or graph/paper comprehension
4. Comments on any strategies used during the eye-tracking tasks around graph or paper reading/comprehension

After coding of the student data was complete, the prevalence of all themes in each data source was analyzed. Table 2 shows the number of instances each code was identified in the two eye-tracking sessions and examples for each code.

Assessment of Scientific Graphing and Engagement with Scientific Literature

As a complement to the identification of reading strategies, the authors collected self-report data on graph- and paper-reading confidence and ability. At the beginning and the end of the program, students completed a short self-assessment (multiple-choice questionnaire) on their confidence and skill levels in creating, reading, and interpreting graphs (see Atkins and McNeal 2018 for questions). Data were analyzed using descriptive statistics.

Results

Eye-tracking methodologies were used to study strategies that a cohort of RECCS students employed while reading a scientific article and interpreting graphs at the beginning (pretest) and at the end (posttest) of the summer research experience. Figure 1 shows heat maps and gaze plots of RECCS students (“novices”), as well as those of experts reading scientific papers. During the pretest assessment, novice eye movements show a reading pattern of the paper from the beginning to the end (see Panel A), similar to that of reading a book, by following the text line-by-line with little eye movement between the different article components. The RECCS students spent little time on the figure and rarely referenced between the text and the figure, as illustrated by the gaze plot (see Panel C). In contrast, when reading the post-program article, students appeared to read selected sections of the article and spend more time on the

TABLE 2. Prevalence of Themes in Transcripts

Coded theme	Description	Pre-program eye-tracking session	Post-program eye-tracking session
		Number of student utterances	Number of student utterances
Prior experience or training	Any mention of students' sense of prior experience or training (or lack of preparedness) around graph/paper reading or graph/paper comprehension	0	0
Self-reported ability/inability	Any mention of students' self-reported ability/inability around graph/paper reading or graph/paper comprehension Student: "This article confused me [laughs] for some reason."	2	4
Confidence	Any mention of students' self-reported confidence (or lack thereof) around graph/paper reading or graph/paper comprehension Student: "I think [the graphs] were easy to understand. The graphs were very straightforward"	1	19
Strategies	Any mention of strategies used during the eye-tracking sessions around graph/paper reading or comprehension Student: "I started off looking at the beginnings of the paragraphs, and then I jumped into the middle."	0	10

Note: Data from student pre-program and post-program eye-tracking sessions around each graph/paper reading ($n = 10$ students).

figure; the gaze plot shows strong integration between the text and other elements of the article (see Panels E, G). When reviewing the paper-reading patterns of the experts (see Panels B, D, F, H), they appeared to read selected sections of the paper and jumped between different components of the paper, especially the graphic. Multiple revisits to different areas of the paper (visit counts) indicate a deeper synthesis and integration of content. RECCS students, after their program completion (posttest; see Panels E, G), revisited all parts of the paper more frequently than on the pretest, integrating different paper components. Thus, RECCS students appear to adopt more expert-like paper-reading strategies at the posttest. Visit counts (see Figure 2) describe these patterns quantitatively. All changes from novice from pretest to posttest were significant ($p < 0.05$) with large effect sizes (Cohen's $d > 1.1$).

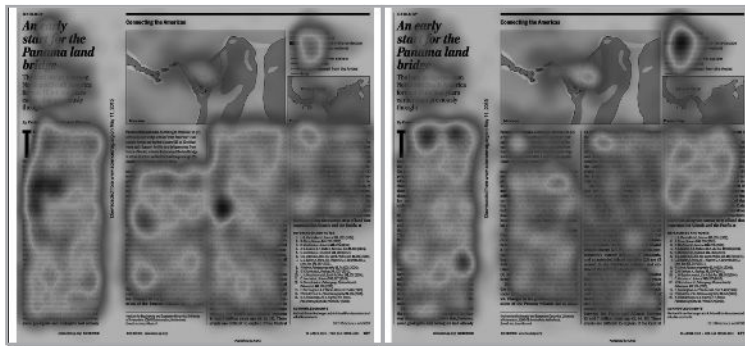
Novices took about six minutes to read the pretest article ($t = 332$ sec.), whereas experts took about a minute less ($t = 280$ sec.). Novices and experts both took about eight minutes to read the posttest article ($t_{\text{novice}} = 484$ sec.; $t_{\text{expert}} = 453$ sec.). Although the novice-expert gap was 52 sec. at pretest, there was no difference between experts and RECCS students at posttest.

Similar to the exploration of students' strategies for scientific paper reading, eye-tracking allowed for the exploration eye movement while reading and interpreting

scientific graphs. Heat maps from the exploration of a scientific graph showed that RECCS students (novices) at the beginning of the program attempted to learn about the main idea of the graph by reading the title and spending very little time looking at the data (see Panels A, B in Figure 3). Experts, on the other hand, focused on the data displayed in the graph to learn about the main idea of the graph (about double the time; see Panels A, C, E) and spent less time reading the title (about half the time; Panel A). On the posttest, RECCS students' eye movement (see Panels A, D) was very similar to that of experts (see Panels A, C, E); both groups focused on the data and the legend. This change in eye-movement pattern suggests that students developed more expert-like behavior and moved toward proficiency toward the end of their research experience. Panel A shows the visit duration for the four AOIs. From pretest to posttest, novices spend significantly less time looking at the title and the axis (large effect size, Cohen's $d > 1$), more similar to experts. Changes in the time spent on data and legend appear nominally different and more "expert-like," but because of large variation, the change is not significant.

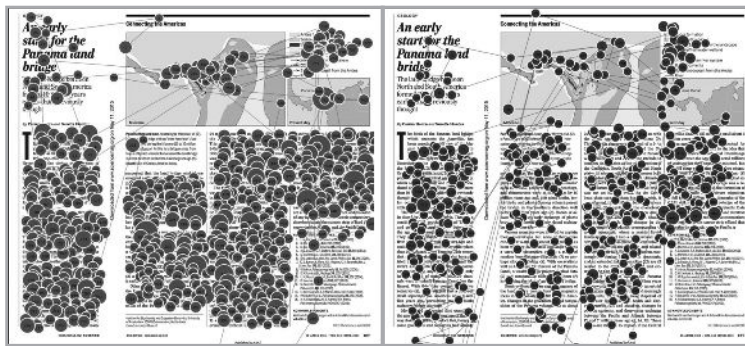
Studying the transcripts from the eye-tracking sessions shows that students described more readily strategies they used and made comments about their confidence on the tasks during the post-program eye-tracking session, and a few more students commented on their ability level (see

FIGURE 1. Eye-Tracking Data (Heatmaps and Gaze-Plots) of Novices and Experts Reading Scientific Papers



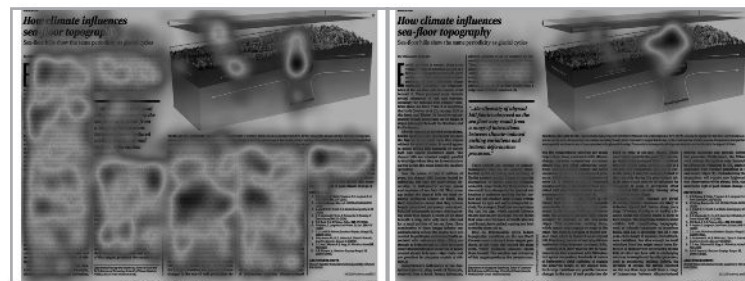
Panel A: Heatmap; novices reading the pretest article ($n = 10$)

Panel B: Heatmap; experts reading the pretest article ($n = 8$)



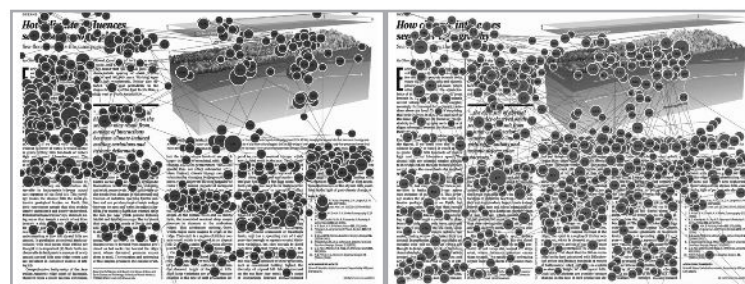
Panel C: Gaze plot; novice reading the pretest article ($n = 1$, same student pretest and posttest)

Panel D: Gaze plot; expert reading the pretest article ($n = 1$)



Panel E: Heatmap; novices reading the posttest article ($n = 10$)

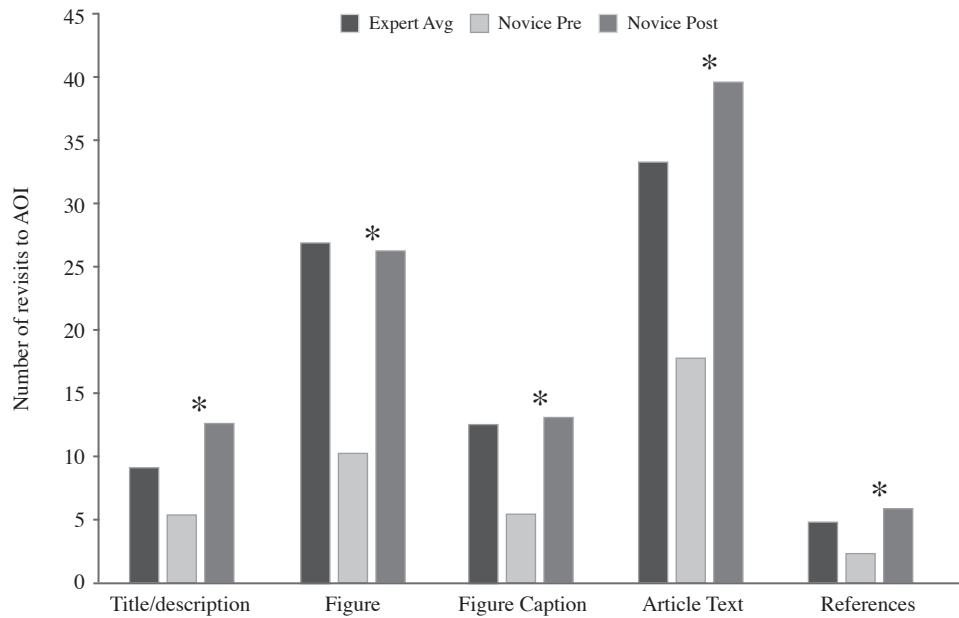
Panel F: Heatmap; expert reading the posttest article ($n = 8$)



Panel G: Gaze plot; novice reading the test article ($n = 1$, same student pretest and posttest)

Panel H: Gaze plot; expert reading the posttest article ($n = 1$)

Note: Novices' data were collected before (pretest) and after (posttest) their participation in the research experience. Gaze-plots show exemplary data from one study subject. Note that heatmaps are not normalized for time spent on task. More saturated gray indicates more visual attention. Color versions of the heatmaps can be viewed on the RECCS program page (CIRES n.d.)

FIGURE 2. Visit Count by Area of Interest (AOI) for the Scientific Article

Note: Data from the beginning of program (novice pre), the end of program (novice post), and from experts. Data are averaged across each group. Asterisks indicate that change of novices from pretest to posttest was significant ($p < 0.05$), each with a large effect size (Cohen's d).

Table 2). The analysis of the transcripts also shows that students increased mentioning of their confidence in the tasks. However, during neither the pre-program nor the post-program eye-tracking session did any of the students comment on prior training or experience or a lack thereof.

When asked about their confidence in understanding scientific papers, four students increased their confidence from the pretest to the posttest, five students reported the same confidence level, and one student reported a lower confidence after the completion of the summer research experience (see Table 3). Two students reported feeling “very confident” about reading scientific papers after the program; no students felt “very confident” at the beginning of the program. When asked to rate their ability in understanding scientific papers (“How good are you at...?”), seven students expressed greater confidence at the end of the REU program, whereas three perceived a lower ability after the completion of the summer research experience. Overall, four students reported that they felt “Proficient” in understanding of scientific papers at the end of the summer, whereas only one felt “Proficient” at the beginning.

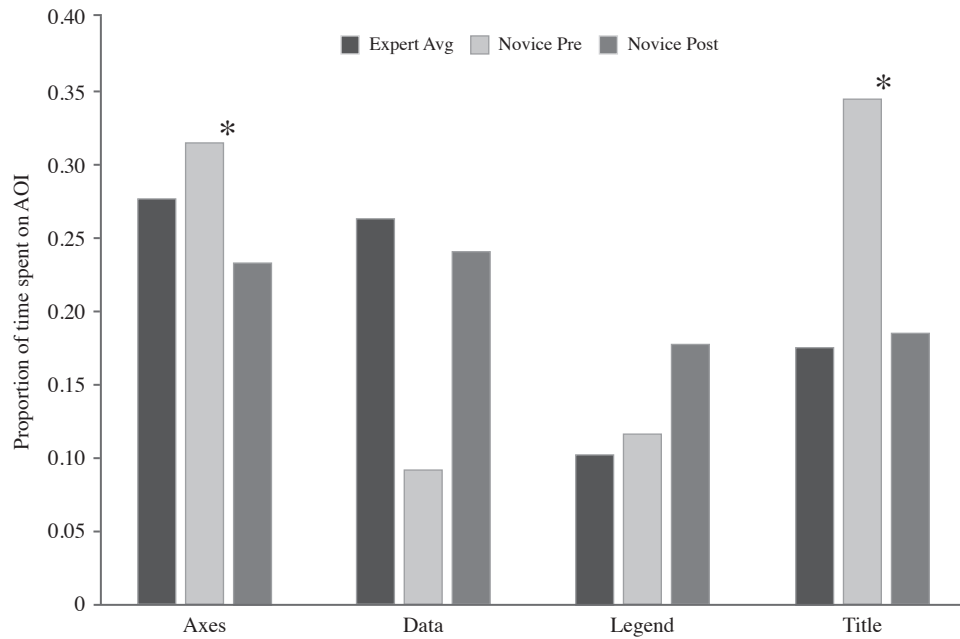
When asked about their confidence in interpreting scientific graphs on the pretest and posttest, five students reported an increase in their level of confidence from the pretest to the posttest, four students rated their confidence the same, whereas one reported a lower confidence after the completion of the summer research experience (see Table

3). Overall, five students reported feeling “very confident” about interpreting graphs at the end of the program, whereas only one reported feeling “very confident” at the beginning. When asked about their ability to interpret scientific graphs, five students increased the rating of their ability of interpreting graphs from the pretest to the posttest, whereas four perceived a lower ability after the completion of the program. Overall, two students reported feeling “Proficient” in interpreting scientific graphs at the end of the summer, whereas three felt “Proficient” at the beginning.

Discussion and Conclusion

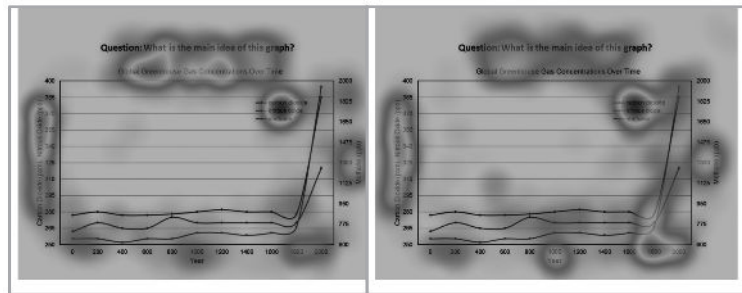
Findings from this case study suggest that participation in an authentic research experience during an REU program resulted in measurable improvement of strategies for reading scientific papers and interpreting scientific graphs for the cohort of students that was examined. Students moved toward more expert-like behavior when interacting with scientific papers or graphs through their participation in the RECCS summer research experience. Although the number of study subjects was small, the findings are an important contribution to the discussion about the value and impact of REU programs. REU program evaluations are often based on self-report data that have been criticized as unreliable measures of impact (Linn et al. 2015); thus, this study provides an additional layer of information on program impacts through quantifying students' growth in important research strategies such as engagement with scientific papers and graphs. Although quantifiable impact measures are an important addition to the available body

FIGURE 3. Eye-Tracking Data (Heatmaps, Visit Counts) of Novices and Experts while Reading and Interpreting a Scientific Graph



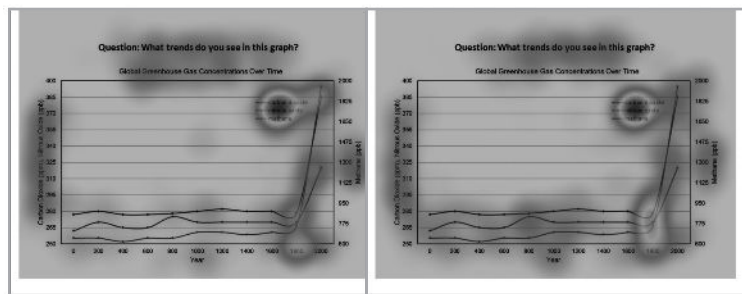
Note: Data from the beginning of program (novice pre), and the end of program (novice post), and from experts. Data are averaged across each group. Stars indicate that change of novices from pre to posttest was significant ($p < 0.05$), each with a large effect size.

Panel A: Normalized visit duration for novices and experts while reading and interpreting scientific graph.



Panel B: Heatmap; novices interpreting the pretest graph ($n = 10$), following prompt: *What is the main idea of this graph?*

Panel C: Heatmap; experts interpreting the pretest graph ($n = 8$), following prompt: *What is the main idea of this graph?*



Panel D: Heatmap; novices interpreting the posttest graph ($n = 10$), following prompt: *What trends do you see in the graph?*

Panel E: Heatmap; experts interpreting the posttest graph ($n = 8$), following prompt: *What trends do you see in the graph?*

Note: Novices' data were collected before (pretest) and after (posttest) their participation in the research experience. Note that heatmaps are not normalized for time spent on task. More saturated gray indicates more visual attention. Color versions of the heatmaps can be viewed on the RECCS program page (CIRES n.d.).

TABLE 3. RECCS Student Self-Assessment of Their Ability and Confidence in Understanding Scientific Papers and Interpreting Scientific Graphs

<i>How confident are you at interpreting scientific graphs?</i>						<i>How confident are you in understanding scientific papers?</i>					
Pretest Response	Very confident					1					
	Somewhat confident		1	3	4			1	5	2	
	Somewhat unconfident			1					2		
	Very unconfident										
		Very unconfident	Somewhat unconfident	Somewhat confident	Very confident		Very unconfident	Somewhat unconfident	Somewhat confident	Very confident	
	Posttest Response						Posttest Response				
Pretest Response	Expert										
	Proficient		2	1					1		
	Competent		1	1	1			2		3	
	Advanced Beginner			1	1				1	1	
	Novice			2					2		
		Novice	Advanced Beginner	Competent	Proficient	Expert		Novice	Advanced Beginner	Competent	Proficient
	Posttest Response						Posttest Response				

Note: Data from student pretest and posttest. The gray-shaded boxes indicate no change from pretest to posttest; any values below the gray-shaded vertical are gains from pretest to posttest; any values above the gray-shaded vertical indicate a decrease from pretest to posttest.

of understanding of the impact of REU programs, perceived student gains around self-confidence, professional identity as scientists, and other self-reported gains from their research experiences (Seymour et al. 2004; Russell et al. 2007; Hunter et al. 2007; Thiry et al. 2011) are critical for students' persistence in STEM. The independent measure of eye tracking on strategies was combined with asking students for their perceived learning gains around reading graphs and papers. Both indicators point into the same direction, reporting increasing skills and confidence in interacting with graphs and papers.

This study indicates that eye-tracking technology facilitates an exploration of expert-like strategies around scientific paper reading and graph interpretation. The data suggest that, when asked to read a scientific paper, experts read selected parts of the paper while their eyes move between different components of the paper, especially between text and graphics. In contrast, novices with limited or no experience in reading scientific papers follow a standard line-by-line reading pattern, starting from the beginning of the text, with little eye movement between different parts

of the paper. When asked to interpret graphs, experts in the study focus on the data displayed in the graph, the axes, and the legend, whereas novices attempt to find information about the graph through the graph title and look at data less. This finding is similar to Ho and colleagues (2014) who used eye-tracking approaches to examine high and low prior knowledge students while they engaged with text and figures. These authors found the students with more prior knowledge engaged more with the figures and that they had more interactions between text and figures than students with low prior knowledge. Additionally, Ooms and colleagues (2012) also used eye-tracking to measure expert and novice subjects during map-reading tasks and found that experts were able to interpret the map contents more efficiently than novice users. During a study aimed to investigate the viewer's navigation of scientific graphs using eye-tracking by Atkins and McNeal (2018), low-performers spent less time viewing important graph features that represented the data and spent more time looking at the text features of the graph. Thus, novice-like behavior differs from expert patterns when viewing scientific papers and graphs; however, REU

students move measurably toward expert-like behavior as a result of participation in the RECCS program. The study findings allow the answering of research question 1 (In which ways does participation in an REU program impact students' strategies in engaging with scientific papers and graphs?). It was found that students change the pattern of their reading of scientific papers and scientific graphs toward expert-like behavior. Thus, the improvements in students' strategies around scientific paper and graph reading are measurable over the duration of just a nine-week research experience program. With respect to graph reading, students appear to spend less time on graph components (Bertin's step 1) on the posttest and more on data (Bertin's step 2) and synthesis (Bertin's step 3) when compared to the pretest.

Although science communication training during the RECCS program exposes the REU students to scientific papers and trains them in thinking about graph components, the RECCS team did not explicitly highlight to novices during the training workshops that there are inherent differences between expert and novice behavior, which alone could influence the performance of novices on the posttest. Instead, instruction focused on discussion of strategies of reading papers and graphs. In the future, it will be possible to point out these different patterns and, through a metacognitive process, to support students even more in their learning of such research skills.

Both the self-report and the eye-tracking data showed a growth toward improved confidence and proficiency in students' research tasks around paper reading and graph interpretation (our research question 2 (In which ways do students' self-reported confidence and ability of engaging with scientific papers and graphs change through the participation in an REU program?). Although a small number of students decreased the ranking of their confidence and self-reported ability, it could be argued that these students did not actually decrease their ability, but rather they have gained a more realistic understanding of what is entailed by engaging with scientific papers and graphs. Thus, even though they improved their skills, the relative rating may have dropped. Self-report bias has been described in the literature previously (Stone et al. 1999). Eye-tracking allows researchers to quantify the changes in the strategies for these tasks as it focuses on cognitive processes and thus provides a powerful tool in exploring patterns of engagement with scientific papers and graphs.

This case study adds to the literature on exploring undergraduate graph-reading strategies (Atkins and McNeal 2018; Wang et al. 2012) and constitutes the first study to explore both expert and novice patterns of reading scientific papers, while measuring the REU participants' improvements in graph and paper reading as a result of their participation.

Limitations and Future Work

This case study was conducted with a small sample size. Increasing the number of novices and experts would provide more statistical power for exploring the data. Experts were recruited based on a convenience sampling. The authors recognize that most scientific papers are multi-page papers, and studying strategies for reading a one-page paper may have some limitations. However, it is assumed that if a subject does not display expert reading strategies in a one-page paper, it is not likely to display in a multi-page paper, and the eye-tracking research design described here has a limit of a one-page design. In a repeat of this study, a more strategic sampling of experts with regard to their expertise would help standardize the expert baseline. Future work could also include an exploration of the trajectory of these REU students and their research strategies as they advance in their academic career.

Acknowledgments

This work was supported by the National Science Foundation under Grant EAR-1757930 and EAR-1461281 (REU site grants) and through the Broader Impact component of the Critical Zone Observatory grants EAR-1239281 and EAR-1331828. Part of the work was also supported by the NOAA Cooperative Agreement with CIRES, NA17OAR4320101. The RECCS program would not have been possible without the participation and enthusiasm of all RECCS participants, mentors, and RECCS program personnel. The authors thank Suzanne Anderson, Lesley Smith, Rebecca Bachelor, Valerie Sloan, Rebecca Hacker, Jennifer Taylor, Barbara Monday, Renee Curry, Lorena Medina Medina Luna, Amanda Morton, Patrick Shabram, and many other advisers and supporters for their work on the RECCS program. Coding of the transcripts was conducted by Quinlain Swain. Suggestions from anonymous reviewers significantly strengthened the article.

References

- Atkins, Rachel M., and Karen S. McNeal. 2018. "Exploring Differences among Student Populations during Climate Graph Reading Tasks: An Eye Tracking Study." *Journal of Astronomy and Earth Sciences Education (JAESE)* 5(2): 85–114. doi: 10.19030/jaese.v5i2.10219
- Bertin, Jacques. 1983. *Semiology of Graphics: Diagrams, Networks, Maps*. Trans. W. J. Berg. Madison: University of Wisconsin Press.
- Boucheix, Jean-Michel, and Richard K. Lowe. 2010. "An Eye Tracking Comparison of External Pointing Cues and Internal Continuous Cues in Learning with Complex Animations." *Learning and Instruction* 20(2): 123–135. doi: 10.1016/j.learninstruc.2009.02.015
- Braun, Virginia, and Victoria Clarke. 2006. "Using Thematic Analysis in Psychology." *Qualitative Research in Psychology* 3(2): 77–101. doi: 10.1191/1478088706qp063oa
- Brewer, Carol A., and Diane Smith. 2011. *Vision and Change in*

- Undergraduate Biology Education: A Call to Action*. Washington, DC: American Association for the Advancement of Science.
- Butcher, Kirsten R. 2006. "Learning from Text with Diagrams: Promoting Mental Model Development and Inference Generation." *Journal of Educational Psychology* 98(1): 182–197.
- Canham, Matt, and Mary Hegarty. 2010. "Effects of Knowledge and Display Design on Comprehension of Complex Graphics." *Learning and Instruction* 20(2): 155–166. doi: 10.1016/j.learninstruc.2009.02.014
- CIRES, University of Colorado Boulder. n.d. "Research Experience for Community College Students (RECCS)." Accessed June 23, 2021. <https://cires.colorado.edu/outreach/reccs>
- Conrad, Clinton P. 2015. "How Climate Influences Sea-Floor Topography." *Science* 347: 1204–1205. doi: 10.1126/science.aaa6813
- Cooper, Jessica, Kristen Jabanoski, and Marlene Kaplan. 2019. "Exploring Experiential Opportunity Impacts on Undergraduate Outcomes in the Geosciences." *Journal of Geoscience Education* 67: 249–265. doi: 10.1080/10899995.2019.1581394
- Curcio, Frances R. 1987. "Comprehension of Mathematical Relationships Expressed in Graphs." *Journal for Research in Mathematics Education* 18: 382–393. doi: 10.5951/jresemathe-duc.18.5.0382
- Dean, Rachel. 2013. "How to Read a Paper and Appraise the Evidence." *In Practice* 35: 282–285. doi: 10.1136/inp.f1760
- Freedman, Eric G., and Priti Shah. 2002. "Toward a Model of Knowledge-Based Graph Comprehension." In *Diagrammatic Representation and Inference. Diagrams 2002. Lecture Notes in Computer Science*, ed. Mary Hegarty, Bernd Meyer, and N. Hari Narayanan, 18–30. Berlin: Springer. doi: 10.1007/3-540-46037-3_3
- Glazer, Nirit. 2011. "Challenges with Graph Interpretation: A Review of the Literature." *Studies in Science Education* 47: 183–210. doi: 10.1080/03057267.2011.605307
- Graham, Mark J., Jennifer Frederick, Angela Byars-Winston, Anne-Barrie Hunter, and Jo Handelsman. 2013. "Increasing Persistence of College Students in STEM." *Science* 341: 1455–1456. doi: 10.1126/science.1240487
- Guthrie, John. T., Shelley Weber, and Nancy Kimmerly. 1993. "Searching Documents: Cognitive Processes and Deficits in Understanding Graphs, Tables, and Illustrations." *Contemporary Educational Psychology* 18: 186–221. doi: 10.1006/ceps.1993.1017
- Ho, Hsin J., Mengy-Jung Tsai, Ching-Yeh Wang, and Chin-Chung Tsai. 2014. "Prior Knowledge and Online Inquiry-Based Science Reading: Evidence from Eye Tracking." *International Journal of Science and Mathematics Education* 12: 525–554. doi: 10.1007/s10763-013-9489-6
- Hoorn, Carina, and Suzette Flantua. 2015. "An Early Start for the Panama Land Bridge." *Science* 348 186–187. doi: 10.1126/science.aab0099
- Hunter, Anne-Barrie, Sandra L. Laursen, and Elaine Seymour. 2007. "Becoming a Scientist: The Role of Undergraduate Research in Students' Cognitive, Personal, and Professional Development." *Science Education* 91(1): 36–74. doi: 10.1002/sce.20173
- Hyönä, Jukka. 2010. "The Use of Eye Movements in the Study of Multimedia Learning." *Learning and Instruction* 20: 172–176. doi: 10.1016/j.learninstruc.2009.02.013
- Jarodzka, Halszka, Katharina Scheiter, Peter Gerjets, and Tamara van Gog. 2010. "In the Eyes of the Beholder: How Experts and Novices Interpret Dynamic Stimuli." *Learning and Instruction* 20: 146–154. doi: 10.1016/j.learninstruc.2009.02.019
- Leinhardt, Gaea, Orit Zaslavsky, and Mary Kay Stein. 1990. "Functions, Graphs, and Graphing: Tasks, Learning, and Teaching." *Review of Educational Research* 60(1): 1–64. doi: 10.3102/00346543060001001
- Linn, Marcia C., Erin Palmer, Anne Baranger, Elizabeth Gerard, and Elisa Stone. 2015. "Undergraduate Research Experiences: Impacts and Opportunities." *Science* 347: 12617571–12617576. doi: 10.1126/science.1261757
- Malamitsa, Katerina, Panagiotis Kokkotas, and Michael Kasoutas. 2008. "Graph/Chart Interpretation and Reading Comprehension as Critical Thinking Skills." *Science Education International* 19: 371–384.
- Mason, Lucia, Patrik Pluchino, Maria Caterina Tornatora, and Nicola Ariasi. 2013. "An Eye-Tracking Study of Learning from Science Text with Concrete and Abstract Illustrations." *Journal of Experimental Education* 81: 356–384. doi: 10.1080/00220973.2012.727885
- Maudlin, Lindsay C., Karen S. McNeal, Heather Dinon-Aldridge, Corey Davis, Ryan Boyles, and Rachel M. Atkins. 2020. "Website Usability Differences between Males and Females: An Eye-Tracking Evaluation of a Climate Decision Support System." *Weather, Climate, and Society* 12(1): 183–192. doi: 10.1175/wcas-d-18-0127.1
- Mayer, Richard E. 2001. *Multimedia Learning*, New York: Cambridge University Press.
- Mayer, Richard E. 2005. *The Cambridge Handbook of Multimedia Learning*. New York: Cambridge University Press.
- Mayer, Richard E. 2010. "Unique Contributions of Eye-Tracking Research to the Study of Learning with Graphics." *Learning and Instruction* 20: 167–171. doi: 10.1016/j.learninstruc.2009.02.012
- McNamara, Danielle S., Eileen Kintsch, Nancy Butler Songer, and Walter Kintsch. 1996. "Are Good Texts Always Better? Interactions of Text Coherence, Background Knowledge, and Levels of Understanding in Learning from Text." *Cognition and Instruction* 14(1): 1–43. doi: 10.1207/s1532690xci1401_1
- Mitra, Ritayan, Karen S. McNeal, and Howard D. Bondell. 2017. "Pupillary Response to Complex Interdependent Tasks: A Cognitive-Load Theory Perspective." *Behavior Research Methods* 49: 1905–1919. doi: 10.3758/s13428-016-0833-y
- Moreno, Roxana, and Richard E. Mayer. 1999. "Cognitive Principles of Multimedia Learning: The Role of Modality and Contiguity." *Journal of Educational Psychology* 91: 358–368. doi: 10.1037/0022-0663.91.2.358
- National Academies of Sciences, Engineering, and Medicine. 2018. *Indicators for Monitoring Undergraduate STEM*

Education. Consensus Study Report. Washington, DC: National Academies Press.

National Science Foundation (NSF). 2019. "Research Experiences for Undergraduates (REU) Sites and Supplements. Program Solicitation 13-542." Accessed February 1, 2021. <https://www.nsf.gov/pubs/2019/nsf19582/nsf19582.htm>

Ooms, Kristien, Gennady Andrienko, Natalia Andrienko, Philippe De Maeyer, and Veerle Fack. 2012. "Analysing the Spatial Dimension of Eye Movement Data Using a Visual Analytic Approach." *Expert Systems with Applications* 39(1): 1324–1332. doi: 10.1016/j.eswa.2011.08.013

Ooms, Kristien, Philippe De Maeyer, and Veerle Fack. 2014. "Study of the Attentive Behavior of Novice and Expert Map Users Using Eye Tracking." *Cartography and Geographic Information Science* 41(1): 37–54. doi: 10.1080/15230406.2013.860255

Paivio, Allan. 1990. *Mental Representations: A Dual Coding Approach*. Vol. 9. New York: Oxford University Press.

Peterman, Karen, Kayla A. Cranston, Marie Pryor, and Ruth Kermish-Allen. 2015. "Measuring Primary Students' Graph Interpretation Skills via a Performance Assessment: A Case Study in Instrument Development." *International Journal of Science Education* 37: 2787–2808. doi: 10.1080/09500693.2015.1105399

Rayner, Keith. 1998. "Eye Movements in Reading and Information Processing: 20 Years of Research." *Psychological Bulletin* 124: 372–422. doi: 10.1037/0033-2909.124.3.372

Rayner, Keith. 2009. "Eye Movements and Attention in Reading, Scene Perception, and Visual Search." *Quarterly Journal of Experimental Psychology* 62: 1457–1506.

Russell, Susan H., Mary P. Hancock, and James McCullough. 2007. "Benefits of Undergraduate Research Experiences." *Science* 316: 548–549. doi: 10.1126/science.1140384

Seymour, Elaine, Anne-Barrie Hunter, Sandra L. Laursen, and Tracee DeAntoni. 2004. "Establishing the Benefits of Research Experiences for Undergraduates in the Sciences: First Findings from a Three-Year Study." *Science Education* 88: 493–534. doi: 10.1002/sce.10131

Shah, Priti, and James Hoeffner. 2002. "Review of Graph Comprehension Research: Implications for Instruction." *Educational Psychology Review* 14(1): 47–69.

Shah, Priti, and Patricia A. Carpenter. 1995. "Conceptual Limitations in Comprehending Line Graphs." *Journal of Experimental Psychology: General* 124(1): 43–61. doi: 10.1037/0096-3445.124.1.43

Stone, Arthur A., Christine A. Bachrach, Jared B. Jobe, Howard S. Kurtzman, and Virginia S. Cain. 1999. *The Science of Self-Report: Implications for Research and Practice*. Mahwah, NJ: Lawrence Erlbaum Associates.

Thiry, Heather, Sandra L. Laursen, and Anne-Barrie Hunter. 2011. "What Experiences Help Students Become Scientists? A Comparative Study of Research and Other Sources of Personal and Professional Gains for STEM Undergraduates." *Journal of Higher Education* 82: 357–388. doi: 10.1080/00221546.2011.11777209

US Environmental Protection Agency. 2014. "Climate Change

Indicators in the United States." Accessed February 1, 2021. <https://www3.epa.gov/climatechange/science/>

Van den Broek, Paul. 2010. "Using Texts in Science Education: Cognitive Processes and Knowledge Representation." *Science* 328: 453–456. doi: 10.1126/science.1182594

Van Gog, Tamara, and Katharina Scheiter. 2010. "Eye Tracking as a Tool to Study and Enhance Multimedia Learning." *Learning and Instruction* 20: 95–99. doi: 10.1016/j.learninstruc.2009.02.009

Wang, Zu Hao, Silin Wei, Wei Ding, Xiuyun Chen, Xiaonan Wang, and Kaiyan Hu. 2012. "Students' Cognitive Reasoning of Graphs: Characteristics and Progression." *International Journal of Science Education* 34: 2015–2041. doi: 10.1080/09500693.2012.709333

Weston, Timothy J., and Sandra L. Laursen. 2015. "The Undergraduate Research Student Self-Assessment (URSSA): Validation for Use in Program Evaluation." *CBE—Life Sciences Education* 14(3): ar33.1-10. doi: 10.1187/cbe.14-11-0206

Willmott, Christopher J. R., Richard P. Clark, and Timothy M. Harrison. 2003. "Introducing Undergraduate Students to Scientific Reports." *Bioscience Education* 1(1): 1–8. doi: 10.3108/beej.2003.01010010

Wyatt, David, Michael Pressley, Pamela B. El-Dinary, Shelly Stein, Peter Evans, and Rachel Brown. 1993. "Comprehension Strategies, Worth and Credibility Monitoring, and Evaluations: Cold and Hot Cognition When Experts Read Professional Articles That Are Important to Them." *Learning and Individual Differences* 5(1): 49–72. doi: 10.1016/1041-6080(93)90026-o

Yeoman, Kay H., and Barbara Zamorski. 2008. "Investigating the Impact on Skill Development of an Undergraduate Scientific Research Skills Course." *Bioscience Education* 11(1): 1–14. doi: 10.3108/beej.11.5

Anne U. Gold

University of Colorado Boulder,
anne.u.gold@colorado.edu

Anne U. Gold is the director of the Education & Outreach Program for the Cooperative Institute for Research in Environmental Sciences (CIRES) at the University of Colorado Boulder. She has been leading a Research Experiences for Undergraduates (REU) program for seven years. Her research is centered on improving geoscience and climate science education.

Rachel Atkins is a PhD candidate in the Marine, Earth, and Atmospheric Sciences Department at North Carolina State University. Her research has centered on student learning in the geosciences, investigating spatial skill evolution, climate communication, and the application of augmented reality in the undergraduate classroom.

Karen S. McNeal is an associate professor of geoscience education and geocognition in the Department of Geosciences at Auburn University. Her research is focused on affect and cognition in learning about the geosciences through qualitative and quantitative methods.