Faculty believe that science students should learn a range of critical thinking skills, including interpreting data, designing an experiment, communicating results, and reading and evaluating published research (Coil et al., 2010). Students learn these skills from classroom, lab, and extracurricular experiences as well as from undergraduate research, which allows them to work in collaborative environments towards common objectives (Hunter et al., 2007). Optimally, this research enables students to work as “scientists in training” (Gonyo & Cantwell, 2014), collaborating with their peers and mentors on an authentic project of situated learning, and it is important for this learning to include writing and reviewing research (Hunter et al., 2007). While critical thinking can mean different things across disciplines in terms of writing (e.g., Rademaekers, 2018), in the sciences, it includes evaluating the claims of other scientific literature, which occurs when scientists read or peer review (Rademaekers, 2018). Peer review is, in itself, a critical thinking activity in the sciences, as it has the potential to help students learn about the process of scientific writing and publishing while evaluating the literature (Trautmann, 2009). To this end, lab mentors can include students in the writing and publication process to give them a space to think critically about research.

While undergraduate research experiences can provide opportunities to read, write, and peer review in the sciences, research positions may not be accessible to or even desired by all students. Other activities have the potential to drive similar skills to supplement or replace undergraduate research, such as student journals. These journals can immerse students in the process of writing, reviewing, and publishing. However, the pedagogical research on what students learn through these journals is lacking, especially with respect to peer review. In this study, we characterize the peer review comments from undergraduate and graduate students for a student-run scientific journal. We also explore why these students donate their time to peer review as an extracurricular activity. In the following introduction, we review STEM student journals as well as research on peer review relevant to our goals. Our overarching aim is to report on how students peer review in a professional context. Therefore, we contextualize our Peer Reviewers in terms of student and expert peer review, particularly in STEM, and the themes in our data set. On a spectrum of expertise, we speculate that our students might align themselves more with expert than novice peer reviewers because they volunteer for this authentic role.

The Rise in Number of Student Journals
The number of student journals has risen exponentially over the years (Ng et al., 2017). The Council of Undergraduate Research (n.d.) lists hundreds of undergraduate journals across disciplines, many of which are student-run. Many STEM journals require authors to be students and publish a range of opinion articles, reviews, original research, and science news. Some journals are indexed in Google Scholar and other databases, expanding the accessibility of student-authored work (Ng et al., 2017).
Support for student journals has been mixed. Some have argued against the value of undergraduate publications beyond student opinions or news. As research journals, they increase the pressure for undergraduates to publish, as well as the pressure for faculty to publish results in an atypical venue (Gilbert, 2004; Siegel, 2004). Yet others have argued that these journals give students opportunities they wouldn’t otherwise have, inviting them into the publishing process, rather than making this goal unattainable, in a short amount of time (Jungck et al., 2004). By giving students the opportunity to publish, these journals send a message of support and acknowledge students’ research and writing contributions.

Encouraging students to participate in the publication process completes the research cycle, bridging research with communication (Spronken-Smith et al., 2013). Students are generally in favor of these journals because they provide so many opportunities. At one institution, students in the health sciences, especially those with plans to enter medical school, supported the creation of a student-run, peer-reviewed scientific journal and indicated interest in publishing (Deonandan et al., 2012). Journals can also be tied to research courses and labs, helping students see the value in their experiments beyond lab reports and instructor audiences (Jones et al., 2011).

Student journals can be edited by faculty or an outside editorial board, or they can ask students to participate in the publishing process as peer reviewers or editors (Kanel, 2008; Stone et al., 2016; Sun et al., 2020; Tatalovic, 2008; Ware & Burns, 2008). Students on an editorial board have commented on the unique experience of becoming active members of a journal and have considered these roles for a future career (Kanel, 2008). In addition, faculty have reported that students need to work like scientists by having a hands-on experience to practice their critical thinking skills (Hunter et al., 2007). These journals offer this hands-on experience outside or in addition to a typical research experience. The act of reviewing submissions may help students understand how to give and apply constructive criticism (Walkington, 2012), promoting critical thinking in the same way that evaluating published research can (Spronken-Smith et al, 2013).

The Role of Student Peer Review in the Classroom

While some student journals include students in peer review, we do not know how students peer review in this context. Most of what we know about student peer review, writing, and critical thinking comes from coursework in STEM. In the sciences, these critical thinking skills can include the ability to understand results and analyze research. These skills can be refined through the writing process as students integrate data while evaluating existing sources (Dowd et al., 2018; Quitadamo & Kurtz, 2007; Stephenson & Sadler-Mcknight, 2016). When this writing process is coupled with peer review over time, students can increase their analysis and inquiry skills as a measure of critical thinking (Weaver et al., 2016). Research with calibrated, online peer-review systems has also connected scientific reasoning and critical thinking gains (Gunersel et al., 2008; Timmerman & Strickland, 2009).

In addition, students often report that providing feedback is in itself a beneficial practice, sometimes more useful than receiving feedback (Gaynor, 2020; Kaufman & Schunn, 2011; Nicol et al., 2014; Pearce et al., 2009). In STEM, providing feedback to peers also helps students reflect on content in writing to better understand assignments (Finkenstaedt-Quinn et al., 2021). Together, these results suggest that the benefits of peer review go beyond providing feedback to the original author.

Research on peer review has categorized the types of comment that students provide, as well as how authors perceive these comments. Two prevalent comment types...
are content and style (Ahmed, 2020; Colthorpe et al., 2014; Mulder et al., 2013). Content comments focus on the topic, while style comments address items like grammar and flow. These types can be further distinguished by affect (positive, negative, or neutral). In one study of oral presentations, peer feedback contained more style than content comments, and the comments were more likely to be positive than negative or neutral (Colthorpe et al., 2014). Positive comments were also more common in other peer feedback research (Cho et al., 2006; Patchan et al., 2009), which may be problematic. In one study, authors were more likely to revise lower quality drafts when receiving feedback that included more praise (Cho & Cho, 2011), suggesting that the affect of comments may influence revision.

To authors, the most helpful comments are ones that include suggestions (Finkenstaedt-Quinn et al., 2019; Walker, 2009), indicating the need for peer reviewers to offer solutions in their feedback, with or without the mention of a problem. Novice reviewers would, however, be unable to provide this feedback if they can’t identify specific content or style problems in the first place.

Affect can also be categorized by the presence of hedging (Finkenstaedt-Quinn et al., 2019; Yallop et al.). Hedging mitigates comments with language of uncertainty, making them appear less assertive. In addition to its use in peer feedback, it is a common technique in scientific research because it allows authors to strategically propose reasoning with their data (Hyland, 1996).

The Influence of Expertise on Feedback

In theory, more experienced writers will provide the feedback necessary to help students who struggle with writing (Mulder et al., 2013). Instructors tend to be more experienced writers and therefore in a better position than students to comment on content. For example, it is common for native and non-native English students to prioritize comments related to language and style, while instructors are more likely to comment on content, organization, relevance, and evidence (Ahmed, 2020). Novice peer reviewers frequently use terms like grammar and good in their feedback, and instructors more frequently use terms such as audience, organization, focus, and purpose (Anson & Anson, 2017). In summary, instructors’ priority on content and organization reflects their expertise in the material and their aim for accuracy.

Student comments also tend to be more positive than those from instructors (Cho et al., 2006; Patchan et al., 2009) and therefore sometimes overate material. Positive feedback is reassuring to authors, but comments that identify a specific problem and solution are more likely to be addressed (Finkenstaedt-Quinn et al., 2019). Instructor comments have this level of specific, directive language more often than student comments (Cho et al., 2006).

Presumably, student peer-review skills will evolve with experience. One study found that students in higher-level biology courses were more likely to use terms related to organization (Donahue & Foster-Johnson, 2019). This focus on organization could suggest that experienced students comment on content more like instructors (Anson & Anson, 2017). Upper-level students also identify more with disciplinary authority than with lower-level students during peer review (Kramer et al., 2022). Lower-level students feel they have authority to give feedback on style and organization rather than on technical content (Kramer et al., 2022). Yet a meta-analysis of peer-review research (Falchikov & Goldfinch, 2000) found no significant difference in the validity of feedback between peers in lower- and upper-level courses. Clearly, more research on peer review

Falchikov & Goldfinch, 2000) found no significant difference in the validity of feedback between peers in lower- and upper-level courses. Clearly, more research on peer review
in the sciences is needed to explore student growth, which may vary with experience and situation.

**Aims of the Current Study**
To foster students’ growth as scientists, faculty can provide students opportunities to peer review beyond classroom assignments. Although some may have reservations about including students in the peer-review process for academic or student journals, a comparison of faculty and graduate students found that both provided equally stringent feedback on a research article submitted for publication (Navalta & Lyons, 2010). Including students in the peer review process even earlier—as undergraduates—can teach them about the publishing process in a situated-learning experience. Importantly, students recognize that being part of publishing contributes to their professional development (Kanel, 2008) and see authentic peer review as an opportunity to enhance critical thinking and apply previous knowledge (Klucevsek, 2016).

When students participate in publishing, they transition from writing that will be seen only by an instructor to writing that will be appreciated by a wider audience. While the cited studies have examined students’ motivations to publish in student journals and students’ ability to peer review in the classroom, we are unaware of any study that combines these themes to explore students’ peer review in an academic journal. Therefore, this study expands our current knowledge of peer review and student-run journals by asking the following questions:

1. What type of feedback do student Peer Reviewers give for a student-run journal?
2. Why do Peer Reviewers volunteer their time to give feedback for a student-run journal?

Students at our institution run a scientific journal that peer reviews and publishes research articles and literature reviews by other students. We wondered how our students reviewed in a professional context, without instructor prompts or grades. Would they review similarly to what is understood in the existing research on student peer review, or would they show evidence of more expert peer review in this situation? Two co-authors of this study are also members of the journal and provided a unique perspective on these results.

**Methodology**

*Participants*
At any given time, there are 10–15 undergraduate and graduate Peer Reviewers for the student journal. Participants for this study were recruited from this pool of Peer Reviewers over a period of two semesters. Nine Peer Reviewers consented to this study (Table 1) and answered open-ended questions about why they peer review and what they have learned from the experience. Table 1 also displays the level and major of each participant at the time of consent, though the peer reviews used in this study could have been completed before or after this time. Both graduate Peer Reviewers were enrolled in doctoral programs. This study was approved by our Institutional Review Board.

*Peer Review Process and Peer Review Samples*
The journal employs a double-blind peer-review process with 2–3 Peer Reviewers per submission. Over the course of an academic year, 10–15 articles will be reviewed and
processed. A Peer Review Coordinator orchestrates this process in conjunction with the Editor-in-Chief, both students on the editorial board of the journal. Peer Reviewers are not formally trained through regular workshops, but the Peer Review Coordinator provides information on the review process, maintains deadlines, gives examples, and offers support via email, similar to academic peer review. The Peer Review Coordinator provides Peer Reviewers with an article, peer-review rubric (Appendix), and due date for completing the review. The Peer Reviewers are instructed to use the rubric, which covers the main sections or needs of most papers, and encouraged to provide in-text comments for greater specificity. However, as noted in Table 1, we found that during this study not all Peer Reviewers provided both rubric and in-text comments to the Peer Review Coordinator.

Because articles are peer reviewed by 2–3 Peer Reviewers, it was possible to have more than one review of the same article. In total, there were 28 articles reviewed in 36 peer-review samples collected for this study. However, we could not compare these reviews because there was not enough overlap across participants. In addition, we do not compare Peer Reviewers directly in this study due to our sample size. Instead, we analyzed comments as a whole rather than as a function of each Peer Reviewer.

Table 1. Number and Type of Peer Reviews by Participant

<table>
<thead>
<tr>
<th>Reviewer and Major</th>
<th>Year of Study</th>
<th>Total No. of Reviews</th>
<th>Rubric Comments</th>
<th>Article Comments</th>
<th>Rubric and Article Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Biology</td>
<td>Undergrad</td>
<td>7</td>
<td>4</td>
<td>429.5</td>
<td>0</td>
</tr>
<tr>
<td>2. Pharmacy</td>
<td>Undergrad</td>
<td>2</td>
<td>0</td>
<td>275.5</td>
<td>0</td>
</tr>
<tr>
<td>3. Forensics</td>
<td>Undergrad</td>
<td>1</td>
<td>1</td>
<td>30</td>
<td>0</td>
</tr>
<tr>
<td>4. Forensics</td>
<td>Undergrad</td>
<td>4</td>
<td>4</td>
<td>458</td>
<td>0</td>
</tr>
<tr>
<td>5. Pharmacy</td>
<td>Graduate</td>
<td>5</td>
<td>2</td>
<td>204.8</td>
<td>0</td>
</tr>
<tr>
<td>6. Biochemistry</td>
<td>Undergrad</td>
<td>4</td>
<td>1</td>
<td>361.33</td>
<td>1</td>
</tr>
<tr>
<td>7. Biology</td>
<td>Undergrad</td>
<td>5</td>
<td>3</td>
<td>546.2</td>
<td>0</td>
</tr>
<tr>
<td>8. Biology</td>
<td>Graduate</td>
<td>4</td>
<td>1</td>
<td>245.5</td>
<td>0</td>
</tr>
<tr>
<td>9. Forensics</td>
<td>Undergrad</td>
<td>4</td>
<td>4</td>
<td>202.75</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>36</td>
<td>20</td>
<td>4</td>
<td>12</td>
</tr>
</tbody>
</table>

*Note. N/A: no comments provided.*

The majority of these articles were literature reviews, though there was one primary research article and one news article. Often, Peer Reviewers request revisions and resubmission, necessitating a second peer review after the author completes revisions. For this study, we coded comments from only the first round of peer review.

**Coding comments**

We coded all available comments on rubrics and in text. As we coded, we noted the location of each comment. For comments on the peer-review rubrics, we used the headings from the rubric itself, which are the most common components or needs of scientific articles: title, abstract, introduction, main text (middle), conclusion or discussion, references, language, and overall impression. For in-text comments, we used locations relative to comment position: title, abstract, introduction, main text (middle), conclusion or discussion, references, and visual (figure or table).

The coding categories described in this article were influenced by work in STEM peer-review feedback: Colthorpe et al. (2014), which assessed type (topic) and affect, and Finkenstaedt-Quinn et al. (2019), which assessed the presence of a problem/solution,
scope, affect, and hedging. However, Table 2 describes how we defined these categories in our work after establishing the coding process. To establish our categories, we first individually read and identified themes across three rubrics and then compared our codes. Based on the initial coding, we agreed on the coding scheme described in Table 2 and used this reference to code the data set. We coded comments independently, followed by a comparison and discussion of the codes until at least 2 of the 3 of us agreed on the coding. Preliminary analysis revealed an interrater agreement of over 97% for the content code and 100% for praise versus neutral/negative.

We coded 881 comments across 36 reviews from nine reviewers. All comments were categorized by type (content or style), affect (positive, negative, or neutral), and scope (global general, global mid-level, or specific), as described in Table 2. A single sentence could contain more than one comment if the sentence joined several ideas in a list or with a conjunction.

Table 2 Category Codes for Comments from Peer Reviewers

<table>
<thead>
<tr>
<th>Category</th>
<th>Code</th>
<th>Notes</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Style</td>
<td>Style</td>
<td>Includes grammar, organization, flow, and paragraph structure</td>
<td>You used issue twice in this sentence.</td>
</tr>
<tr>
<td>Content</td>
<td>Content</td>
<td>Includes language, research, topic, and terminology of the paper</td>
<td>The whole time reading the paper I was curious about data regarding what specific metals and pH changes occurred from each method and it’s great that you compiled this table.</td>
</tr>
<tr>
<td>Affect</td>
<td>Positive</td>
<td>Uses positive language or notes a positive quality of the paper</td>
<td>Expertly organized. This was concise and easy to read.</td>
</tr>
<tr>
<td>Affect</td>
<td>Neutral</td>
<td>Provides an observation or suggestion</td>
<td>I think if you are not going to talk about prevention techniques, then you shouldn’t mention prevention at all.</td>
</tr>
<tr>
<td>Affect</td>
<td>Negative</td>
<td>Uses negative language or criticizes the work</td>
<td>It’ll sound a lot better if you pick a different word here or reword the sentence so you don’t repeat words.</td>
</tr>
<tr>
<td>Scope</td>
<td>General</td>
<td>Describes the paper holistically</td>
<td>All references are correctly cited and used properly.</td>
</tr>
<tr>
<td>Scope</td>
<td>Midlevel</td>
<td>Identifies the section of the paper</td>
<td>The main text also does a good job of analyzing the results of each study and talking about their importance.</td>
</tr>
</tbody>
</table>
Table 2 (continued)

<table>
<thead>
<tr>
<th>Category</th>
<th>Code</th>
<th>Notes</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Specific</td>
<td></td>
<td>Identifies the location of the paper or names terms from the paper</td>
<td>The use of the table is also effective in comparing the treatments.</td>
</tr>
</tbody>
</table>

Beyond these basic categories of type, affect, and scope, comments could also be coded for several additional layers: hedging, problem and/or solution, clarification, and evidence (Table 3). Any comment could be coded for hedging in addition to affect. Comments were coded as problem only if we believed the author could clearly identify the issue from the comment. Broad comments, such as “The style could be better,” would not be coded as a problem because the author would not know how or where to address that problem. Comments could be coded as both problem and solution when the Peer Reviewer clearly identified an addressable issue and suggested a constructive solution. Content comments could be coded for both evidence and clarification layers at the same time.

Table 3 Layer Codes for Comments from Peer Reviewers

<table>
<thead>
<tr>
<th>Layer</th>
<th>Notes</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hedging</td>
<td>Includes uncertain language, regardless of affect</td>
<td>Maybe you could...</td>
</tr>
<tr>
<td>Problem</td>
<td>Indicates an issue that the author could identify and fix</td>
<td>The future research is stated; however, it is general.</td>
</tr>
<tr>
<td>Solution</td>
<td>Describes a suggestion without naming the problem directly</td>
<td>I would try to see if there is any information that can be trimmed from the main texts.</td>
</tr>
<tr>
<td>Clarification</td>
<td>Asks questions or marks unclear content; Only applies to content comments</td>
<td>Does this damage the infrastructure?</td>
</tr>
<tr>
<td>Evidence</td>
<td>Asks for additional supporting data or citations; Comments on the quality of evidence provided by the author, regardless of affect; Only applies to content comments</td>
<td>Establish why this is significant, as it is mentioned in the article but there is no previous research cited.</td>
</tr>
</tbody>
</table>

Analysis of Coding
Comments were coded for the presence (1) or absence (0) of each category. Coding data were compiled using MS Excel 2020 and exported to MS SQL Server 2019 for descriptive analysis. Comments were summed for each review section and further broken down by proportion for type, scope, affect, and problem/solution. Comment sums for hedging, evidence, and/or clarification were stratified by type, scope, affect, and problem/solution. Comments were further categorized by either their location on the peer review rubric or by their position as an in-text comment.
Results

Peer Reviewers Provided Rubric and In-text Comments
To understand the process of peer review in our academic journal, we first analyzed the way in which Peer Reviewers provided comments. Nine Peer Reviewers consented to this study, completing a total of 36 reviews of 28 papers. Each Peer Reviewer provided feedback on 1–7 reviews (Table 1). The majority of reviews included the rubric, while 12 included both rubric and in-text comments. Four reviews included in-text comments but no rubric. Of the 16 total papers with in-text comments, 8 also included in-text line edits, which were not further quantified here.

For further analysis, we separated the rubric and in-text comments to compare the specificity and type. We found that Peer Reviewers provided a range of 30–695 words of feedback on each rubric and a range of 38–787 words on each article. While comment length does not necessarily equate to time spent or quality of a review, there was only one review that contained less than a total of 50 words of feedback. For the purpose of this study, we did not analyze comments as a function of Peer Reviewer because each Peer Reviewer contributed a varying number of reviews. However, on average, most Peer Reviewers provided over 200 words on either rubric, article, or combined rubric and article feedback (Table 1).

Comments Ranged in Type, Scope, and Affect.
To assess the type of comments provided by Peer Reviewers on both rubrics and articles, we first identified a comment as a phrase that addressed either style or content (Table 2). Of 36 reviews, we coded 881 comments as content (695 comments) or style (186 comments) (Figure 1). Across rubric and in-text comments, most comments focused on the main text (22.8% of total comments). This is likely because the main text of an article contains the majority of the research and novel analysis. Contrary to other studies of students' feedback (Ahmed, 2020; Colthorpe et al., 2014), we found that our Peer Reviewers commented more frequently on content than style in all categories (Figure 1). On rubric feedback, 81% of comments were about content, while 72.5% of in-text comments on articles were about content.

Figure 1. Comments categorized by type. Comments (n=881) were coded for style (n= 186) or content (695). A) Comments sorted by rubric category. B) In-text comments sorted by location on the article.
Figure 2 displays comments coded by scope (general, midlevel, or specific). While most of the comments on the rubric were coded as midlevel or general, the in-text comments were almost always specific (Figure 2b). This finding is not surprising, as in-text comments can easily localize the feedback to a sentence or word. However, when Peer Reviewers used the rubric, we still found that 17% of comments were specific and 53% were midlevel, especially when commenting on the title, introduction, or main text. This consistency suggests that students might use language to help make connections, even on a rubric.

![Figure 2](image-url)

*Figure 2. Comments categorized by scope. Comments (n=881) were coded for scope as either General (n = 196), midlevel (n=359), or specific (n=326). A) Comments sorted by rubric category. B) In-text comments sorted by location on the article.*

While praise is helpful, excessive praise could leave writers feeling like the review wasn’t critical or honest. Therefore, to identify affect in our Peer Reviewer’s comments, we coded comments as negative, neutral, or positive. Many comments were suggestions or advice, which we coded as neutral in the absence of negative language. Neutral comments also included affirming comments, such as affirming that the author’s introduction overviewed the paper. In contrast to other studies (Cho et al., 2006; Colthorpe et al., 2014; Patchan et al., 2009), our Peer Review comments were more often negative or neutral than positive. In fact, 70% of rubric comments and 95% of in-text comments were neutral or negative (Figure 3). The section with the largest proportion of positive comments was the Overall Impression section of the rubric (Figure 3a), where 42% of comments were positive. Here, the Peer Reviewers often buffered critical feedback with something positive about the topic or content, as a summary of their feedback.

*Comments Contained Additional Layers*

After coding each comment for type, specificity, and affect, we assessed comments for additional layers (Table 3) in order to further examine the variety of comments that student Peer Reviewers made in this context. For example, during our initial coding, we noticed that some comments clearly stated the problem before offering a solution, while others did not. In some cases, comments identified a problem but did not suggest a solution. Presumably, describing both a problem and a constructive solution would help a writer understand and respond to a comment. Therefore, we coded each of the comments for the presence or absence of a problem and/or a solution (Figure 4). Only
30% of rubric comments noted a problem and/or a solution. There were no obvious trends in the ratio of problem and/or solution comments within each section of the rubric (Figure 4), suggesting that Peer Reviewers employed problem, solution, or problem and solution comments at a similar rate. Overall, we found that 71% of in-text comments contained a problem and/or solution. Most of the problem and/or solution in-text comments were located in the main text of the article, which contains the bulk of citations and evidence from the author. Peer Reviewers often used the in-text comments to point out specific issues that needed to be addressed in revisions.

![Figure 3](image)

**Figure 3.** Comments categorized by affect. Comments (n=881) were coded for affect as either negative (n=183), neutral (n=487), or positive (n=211). A) Comments sorted by rubric category. B) In-text comments sorted by location on the article.

![Figure 4](image)

**Figure 4.** Comments categorized by problem or solution. Comments (n=881) were coded for describing a problem (n=114), solution (n=139), or a problem and a solution (n=108). Five hundred and twenty comments had no problem or solution. A) Comments sorted by rubric category. B) In-text comments sorted by location on the article.

While hedging can be described as a type of affect, we chose to code it as an independent layer to examine how it corresponded with affect categories. Figure 5 displays the rate of hedging within comments on both the rubric and article combined.
We found that 13% of style comments and 17% of content comments contained hedging language (Figure 5a) and that 19% of all negative comments and 22% of all neutral comments contained hedging (Figure 5c). With regard to scope, specific comments contained hedging more often than general or midlevel comments (Figure 5c). We also found that most hedging comments (86.4%) referenced a problem and/or solution (Figure 5d).

Our last coding scheme determined whether a comment was related to either clarification and/or evidence. Together, 12% of comments were coded for clarification and 7% of comments were coded for evidence (Figure 5a). It was more common for clarification comments to be specific (Figure 5b), which likely reflects the fact that many of these were in-text comments too. Evidence comments demonstrated a more modest trend toward specific scope (Figure 5b). Not surprisingly, all clarification comments were also neutral or negative (Figure 5c), while it was possible, though rare, for positive comments to compliment the evidence within an article. Of 108 comments with clarification, 63% also contained a problem and/or solution (Figure 5d). Of 67 comments with evidence, 71% contained a problem and/or solution (Figure 5d).

**Figure 5.** Comments categorized by layers: Hedging, Evidence, and Clarification. These graphs represent the percent of type, scope, affect, or problem/solution comments that contained an additional layer (hedging, evidence, clarification). Each comment could code for more than one layer. Rubric and in-text comments were combined. In total, 140 comments contained hedging, 67 discussed the need or presence of evidence, and 108 comments asked for clarification. A) Comments co-coded by type (style or content) and additional layers (hedging, evidence and clarification). Note that only content comments were coded for evidence and clarification. B) Comments were co-coded by scope and layer. C) Comments were co-coded by affect and layer. D) Comments were co-coded by the presence of a problem/solution and layer.
Peer Reviewers Provided Varied Feedback

From coding comments, we found that students collectively provided an array of comment combinations. Because each student performed a different number of peer reviews, we could not quantitatively compare the comments for individual Peer Reviewers. Instead, to explore the variety of our Peer Reviewers, we present our observations of their comments.

We noticed that several Peer Reviewers consistently made more complex commenting moves in their feedback. These comments were classified by several of our coding layers, such as specific content comments with both a problem and a solution that asked for clarification or more evidence. For example, one student’s comments pointed out that the author mis-analyzed a reference in their article, stating:

*The paper is not saying that the DNA disappeared. It means the small fragments were no longer on the gel in Figure 3... Even where there is no apparent function, there is methylation occurring.*

In another peer review, a different student found an error in analysis as well:

*This is not entirely true- while antibiotics may not be as selective (they may kill normal flora in addition to the infection), there are... different antibiotics with different mechanisms and spectrums of action. For example, metronidazole is selective for anaerobic bacteria, while vancomycin is selective for gram positive bacteria.*

These comments were particularly interesting because we could tell the Peer Reviewers were connecting their expertise and prior knowledge to their reviews, suggesting that the Peer Reviewers were reviewing more like experts. Not surprisingly, the Peer Reviewers who most often made these types of comments were upper-class or graduate students at the time of review.

We also noticed that students frequently combined praise with criticism in their comments or overall feedback as a way to soften criticisms. One Peer Reviewer used this method even when advising the Editors to reject the papers for publication, offering constructive criticism and positive feedback so that the original author could still benefit from the peer review process.

Peer Reviewers Reflected on Their Challenges and Motivations.

Overall, the comments made by Peer Reviewers suggest that they put effort into giving specific and constructive feedback. To understand their motivation for volunteering, we asked the Peer Reviewers in an open-ended question why they became Peer Reviewers for our academic journal (Table 4). Peer Reviewers reported a wide variety of motivations. They wanted to gain experience in critically evaluating research, improve their scientific writing and peer review skills, or build experience for their resume. Reasons also included learning about research outside their field.

In another open-ended question, we asked Peer Reviewers how they perceived their growth while volunteering for the journal (Table 4). Here, several students recognized that peer reviewing papers had improved their ability to find weaknesses in their own writing too. Some students also noted that their peer review confidence increased over time, as they were able to give more constructive suggestions.
<table>
<thead>
<tr>
<th>Reviewer</th>
<th>What motivates you to be a Peer Reviewer?</th>
<th>Briefly describe your growth as a peer reviewer.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>To improve my skills in peer review and writing</td>
<td>I have improved my writing and I better understand research. When I first started reading scientific articles it was very difficult. After performing peer review, I understood the structure of these types of writing better.</td>
</tr>
<tr>
<td>2</td>
<td>Getting experience for a future career in research</td>
<td>I am still a relatively new peer reviewer... but I have noticed that I am paying more attention to things as a scientist rather than a student. Rather than reading these articles from the perspective of someone who will be writing these and trying to learn techniques from them, I have become more critical of the content and putting them in the context of the science I know.</td>
</tr>
<tr>
<td>3</td>
<td>I enjoy reading scientific articles and I have a passion for English and editing. I wanted to be an English major.</td>
<td>I feel that I have been able to continually improve my ability to edit and review. It has helped me when I write my own scientific papers for class.</td>
</tr>
<tr>
<td>4</td>
<td>I think the research that others do is very interesting. I review to learn more about the different ways in which people write their research.</td>
<td>It has improved my ability to be conscious of things while writing (both good and bad).</td>
</tr>
<tr>
<td>5</td>
<td>This opportunity has provided great practice for critically evaluating scientific literature; it has also helped me improve my own scientific writing skills; it also looks good on a CV :)</td>
<td>After gaining experience, I became more comfortable with the peer review process and felt that my comments became more targeted and helpful for the author. This experience provided great exercise to practice evaluating scientific literature, as this is a very important skill for a researcher to have.</td>
</tr>
</tbody>
</table>
We also asked Peer Reviewers what challenges they encountered, but 7 of the 9 Peer Reviewers reported no significant challenges. One student who did mention that content was sometimes a challenge reported taking the time to read background sources before completing the review. Although the Peer Review Coordinator offers help to Peer Reviewers and provides them with examples and a rubric, there is no formal training or instructor input. In our survey, Peer Reviewers reported drawing on previous experiences, including instructor feedback, classroom peer reviewing activities in a scientific writing course or lab, and professional development courses in order to conduct their reviews.

Because two of the co-authors on this study were students involved in the journal, their own reflections helped us contextualize our results.

From co-author, former Editor-in-Chief:
These findings align with my personal experiences. Overall, the Peer Reviewers in this journal offer professional, well-balanced feedback to authors. It is more common for upper-class and graduate students to produce higher quality reviews. This is likely because higher level students have more experience with scientific communication from conducting their own research and taking classes. This does not mean younger students should not become Peer Reviewers. Peer reviewing for the journal introduces them to scientific writing and helps them transition into undergraduate research, internships, etc.

From co-author, former Peer Reviewer and Peer Review Coordinator:

This study has given me a better appreciation for the role peer review can play in student development. During our analysis, we observed variability in reviews between different Peer Reviewers, as well as from the same Peer Reviewer over time. For some Peer Reviewers, there was a clear progression as they gradually honed their analysis and communication skills, offering more thoughtful, specific, and actionable feedback.

These statements suggest that peer review can play a role in professional development and provide an opportunity for students to build essential writing and reviewing skills. Combined with the survey responses in Table 4, these statements also suggest that peer review comments vary, but that Peer Reviewers may improve with experience.

Discussion

Analyzing Peer Reviewers’ Comments and Their Role in the Journal
In the sciences, student-run academic journals offer opportunities for writing, publishing, and peer review. Peer review in the classroom has been studied extensively, but there is limited research on student peer review in authentic contexts, such as student journals. In this study, we examined the comments made by students who volunteered to peer review for a student-run, scientific journal, as well as their motivations. We found that our Peer Reviewers provided complex and varied comments.

For several codes, we found that our Peer Reviewers’ comments were different from those observed in existing research on classroom peer review. Some studies have found that students in the classroom prioritize style comments over content (Ahmed, 2020; Colthorpe et al., 2014) and provide more praise (Cho et al., 2006; Colthorpe et al., 2014; Patchan et al., 2009), which sometimes correlates with lower quality revisions (Cho & Cho, 2011). In contrast, we found that our Peer Reviewers’ comments overwhelmingly prioritized content (Figure 1) and provided neutral or negative comments rather than praise (Figure 3). The scope of Peer Reviewers’ comments was more often midlevel or specific, especially when paired with in-text comments (Figure 2). These results suggest that our students may have aligned themselves with the mindset of professionals when situated in a professional role, focusing on constructive content comments rather than grammar and praise.

Authors find suggestions helpful in their feedback (Finkenstaedt-Quinn et al., 2019; Walker, 2009). Our Peer Reviewers sometimes provided solutions with hedging (Figure 5d). Overall, our Peer Reviewers hedged 17% of content comments and 13% of style comments (Figure 5), which is a rate similar to that found by other studies of
students’ comments in the sciences (Donahue & Foster-Johnson, 2019; Finkenstädt-Quinn et al., 2019). It is unclear if our students hedged because they were unsure of their suggestions or simply wanted to soften criticism of their peers.

Several of our Peer Reviewers reported that they used this experience to learn more about research overall and improve their own writing (Table 4). This motivation aligns with outcomes of classroom peer review, as research has found that the act of peer review has benefits to the peer reviewer (Gaynor, 2020; Kaufman & Schunn, 2011; Nicol et al., 2014; Pearce et al., 2009). We know that writing processes can enhance and promote critical thinking in the sciences, including students’ ability to understand results and research (Dowd et al., 2018; Quitadamo & Kurtz, 2007; Stephenson & Sadler-McKnight, 2016). It’s reasonable to assume that peer review of writing, which should require analysis, would also be a practice that promotes critical thinking about research. Based on previous research (Donahue & Foster-Johnson, 2019), we expect that students improve their peer reviews with experience and education. Our Peer Reviewers and Editors both self-reported improvement over time. In addition, research supports that giving students these authentic, professional experiences helps them learn writing processes in STEM as they work with peers and mentors (Huiling, 2008). This and other research (Klucevsek, 2016) suggests that involving students in out-of-classroom peer review experiences can have benefits beyond peer review because it offers an authentic opportunity for critical thinking about research and professional growth.

Reflecting on the Role of Student Peer Review

Some student academic journals are reviewed or monitored by faculty or experts rather than student peers (Stone et al., 2016; Sun et al., 2020). This policy may be motivated by the desire to ensure a certain quality control over the review process. In support of this reasoning, one study found that instructors indeed made more high-level comments related to data than students (Volz & Saterbak, 2009). However, our journal is run by students to give them ownership over the publishing process. We know that inviting students to participate in publishing and editing gives them the chance to critically evaluate work and may motivate them to publish again in the future (Jungck et al., 2004; Mariani et al., 2013; Walkington, 2012; Weiner & Watkinson, 2014).

Our study observed how Peer Reviewers work in this professional role without faculty influence. We found that Peer Reviewers made constructive content comments, but it was still less common for them to comment on data, by critiquing clarity or evidence (Figure 5). Comments on references are relatively rare in peer review (Walker, 2009), presumably because they require more expertise and time. We did find, however, that one of our Peer Reviewers—a graduate student—read a referenced article and checked that author’s analysis. This type of scientific reasoning and evaluation is the type of activity that could promote critical thinking in the sciences (Dowd et al., 2018) and should be encouraged more.

Though we were happy to see more content comments made by our Peer Reviewers, we also recognize that we could refine this professional development opportunity for future peer review. The journal does not formally train Peer Reviewers, similar to peer review in academic journals. Interestingly, a survey of early career researchers revealed that many of these scientists informally learned how to peer review from their mentors, as well as from participating in journal clubs or receiving feedback themselves (Mcdowell et al., 2019). Half of early career researchers receive no training at all (Mcdowell et al., 2019). One could argue that all journals and mentors should train peer reviewers, especially novice reviewers. Our Editors recognized that guiding novice
Peer Reviewers into this experience might also help them transition from classroom activities to this professional context and bolster institutional support for a student-run journal.

From co-author, former Peer Reviewer and Peer Review Coordinator:

*I would like to see the journal take a more active role in developing its Peer Reviewers. Such an initiative could benefit all parties: Peer Reviewers could learn and grow, authors could receive better constructive feedback for revisions, the journal could publish stronger material, and readers could have access to higher quality articles.*

From our results, we see areas that could be enhanced. For example, comments that identify problems and solutions are more likely to be associated with revisions (Finkenstaedt-Quinn et al., 2019). Some research indicates that students comment on less problems than instructors (Patchan et al., 2009). Surprisingly, we found that in-text comments offered a problem and/or a solution 71% of the time, often with the need for clarification (Figure 5 c). However, only 19.6% of total in-text comments contained both a problem and a solution. Previous research found that instructors also infrequently note both simultaneously (Patchan et al., 2009), suggesting that this strategy may be less common even with expertise. Still, encouraging Peer Reviewers to note both problem and solution could help authors address revisions.

In addition, the current method of peer review for our journal involves giving Peer Reviewers a rubric and encouraging in-text comments without requiring them. While students should be directed to annotate documents because authors find this helpful, some students opt to use only a rubric (Gaynor, 2020). In our study, only 12 of 36 peer reviews contained both rubric and in-text comments, and 4 did not fill out the rubric, providing only in-text comments. Based on results in Figure 2, in-text comments are more specific, and we could require these to increase specificity and help authors locate places for revision.

From co-author, former Editor-in-Chief:

*Based on our experiences, we believe the peer review process can be made more effective by holding workshops that encourage reviewers to use the critical thinking skills they have gained in previous classes and research, even if they are not experts in a particular article topic. In addition, student journals may benefit from collaboration with organizations. For example, writing center consultants often report improved writing skills and the ability to give constructive criticism (Dinitz & Kiedaisch, 2009)*

In fact, several students reported that they used previous peer-review experiences and feedback to help shape their comments for the journal. Future training could collaborate with existing resources on campus, or reinforce peer review experiences from relevant STEM courses. This would be especially helpful for students who want to peer review but are uncertain about their ability.

### Limitations and Future Research

We were unable to compare the comments quantitatively among Peer Reviewers due to the limited number of participants, varying number of reviews, and length of comments. We were also unable to track individual growth quantitatively. Despite these limitations,
this study has allowed us to explore the collective feedback from Peer Reviewers for a student-run journal. Our initial observations suggest that upper-level or graduate students provided more specific reviews, but a larger and more long-term study would be needed to address if feedback changed with experience. Future studies could also explicitly examine the evaluation and critical thinking skills that students employ in authentic peer review experiences outside the classroom. We also recognize a need to explore authors in the peer-review process of academic journals, as well as how they respond to these comments and make revisions.

**Conclusion**
This study examined how Peer Reviewers provide feedback in the context of a student-run scientific journal. Through constructive feedback on content, our Peer Reviewers aimed to review scientific writing beyond grammar and praise. In addition, Peer Reviewers’ responses to our surveys showed that they are motivated to learn about new research, gain experience, and improve their own writing. Offering students the opportunity to peer review journal submissions has the potential to promote their professional development as scientists and writers by allowing them to practice these skills in an authentic context.

**Acknowledgment**
We would like to recognize and thank the students who volunteer to peer review or serve as Editors of the journal. We also thank the students who participated in this study for sharing their work.

**Funding**
This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

**References**


Mariani, M., Buckley, F., Reidy, T., & Witmer, R. (2013). Promoting student learning and scholarship through undergraduate research journals. *PS: Political Science & Politics, 46*(04), 830–835. [https://doi.org/10.1017/s1049096513001133](https://doi.org/10.1017/s1049096513001133)


Appendix

Peer Review Rubric

Peer Reviewer:
Title of Article:
Score (1-5):
Title: Does the title reflect the papers content and is it brief? Does it attract potential audiences and is it able to stand alone?
Abstract: Does it state the purpose/problem of the article, summarize principle findings and point out any major conclusions?
Introduction: Does it give substantial background of the issue, its significance, scope and the limits of the research?
Main Text: Is the structure of the paragraphs clear and concise? Does the writer provide evidence to back up any statements made? Is all of the information mentioned relevant to the article?
Conclusion: Does it interpret the main points of the article and relate it back to the originally stated problem/purpose? Summarize all discussion points without repeating irrelevant material?
Language: Is the terminology used in the article easy to interpret? Are there any grammatical mistakes made throughout the article? What type of audience is the article geared towards?
References: Does the author correctly cite all references? Are there in-text citations and a references page?
Overall impression:
Acceptance to Publication: (Yes (with/without) revisions or No)