How to SOAR through Research with the Crowd: Strategic Reading with Crowdsourced Literature Reviews

Amy Fox
UC San Diego
Department of Cognitive Science
La Jolla, CA, USA

Kandarp Khandwala
UC San Diego
Computer Science & Engineering
La Jolla, CA, USA

Tricia J. Ngoon
UC San Diego
Department of Cognitive Science
La Jolla, CA, USA

ABSTRACT

Reading and reviewing research papers for literature reviews is an important task for academic researchers. Yet these tasks are complex and time-consuming and require more efficient strategic reading methods. We examine the use of the SOAR (Select, Organize, Associate, Regulate) active reading method for literature review of research papers. We then examine the viability of leveraging the crowd in this process. We found that learners found the SOAR method useful and felt more focused in specific tasks when collaboratively reading with a researcher. We plan to explore the value of utilizing the SOAR method within community-sourcing for research communities. However, the crowd produced mixed quality results, potentially requiring more guidance and context. From these findings, we plan to explore the use of the SOAR method within a community-sourcing framework and with a learner-crowd collaboration framework.

Author Keywords
crowdsourcing; literature reviews; active reading

ACM Classification Keywords
H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous

INTRODUCTION

Reading academic research articles is an important but time-consuming task. Students need to read a wide range of articles to build a foundation of knowledge in a domain: identifying relevant theories and methods and integrating those with their own work. More novice researchers typically employ less effective linear reading methods in which the paper is read from start to finish in order. Instead, more experienced researchers utilize "channel surfing" strategies to quickly assess and exploit the main ideas of a variety of research papers [9]. In this work, we focus on a specific active reading strategy studied and developed by educational psychologists known as the SOAR (Select, Organize, Associate, and Regulate) method. The SOAR method has been shown to be effective for improving conceptual understanding and retention of information in memory [3]. We examine the effectiveness of this method for reading and integrating research papers for academic researchers.

In a preliminary study, we asked individuals to read a paper of their choice using the SOAR method first individually and then collaboratively with a researcher of this paper acting as the "crowd." We then explored the viability of using the crowd to aid in the SOAR reading process. Though reading and understanding research papers is a complex task, previous crowdsourcing research has demonstrated that crowd workers could provide novel insights and high quality work [12]. The crowd has also been utilized in similar tasks such as story writing [7], paper editing [1], and sensemaking [4]. We ran a formative crowd study on Amazon Mechanical Turk to determine whether crowd workers can reliably produce high-quality summaries utilizing the SOAR method workflow.

Figure 1. An example of the Selection task (top) and Organize task (bottom) completed by a researcher in the collaborative reading condition of the pilot study.

We found that without context, crowd workers tend to rely on keywords and brevity of sections rather than actual important points. We instead found that our pilot learners

enjoyed the collaborative reading component and found the SOAR method useful. We believe the value of our research may be the use of the collaborative SOAR method to be in community-sourcing, where all members of a research community would benefit from shared artifacts of knowledge from research literature.

RELATED WORK

Making Sense of Academic Research Papers

Scientific literacy is a crucial skill in research. The complex nature of reading academic research requires more flexible, strategic approaches to reading rather than linear reading methods. [9] describe "channel surfing" methods of literature review, quickly assessing and exploiting information from various sources of literature. While reading paper abstracts is useful, abstracts often cannot substitute reading full papers due to generally poor legibility, lack of sufficient evidence of results, and potential errors and inconsistencies within the text body [8]. Making sense of research literature has also become a topic of interest in human-computer interaction research. [14] developed CiteSense, a tool that aids in making sense of research by supporting information search strategies such as searching, filtering, browsing, and reference chasing. The Scholarly Ontologies project also aimed to use claim network models to create sensible mappings between research literature to aid in information foraging and scholarly sensemaking activities [11]. These technologies and research work demonstrate the importance of simplifying the process of research literature reviews. Our work aims to aid the research reading process using the SOAR reading method.

The SOAR Method

The SOAR (Select, Organize, Associate, and Regulate) method is an active reading strategy developed and studied in educational psychology. The SOAR strategy has been shown to be beneficial for improving conceptual understanding and information retention [3]. It has been compared to an older active reading method, the SQ3R (Survey, Question, Read, Recite, and Review) system [10]. However, the SQ3R method is highly difficult to learn and master. In an examination of the efficacy of the two reading strategies, [5] found that students using the SOAR strategy learned more relationships, facts, and concepts than those who used the SQ3R method. Because of the SOAR method's effectiveness in increasing conceptual understanding and retention of information, we will explore using the method for reading and reviewing research papers. We will then examine the method in the context of crowd work to determine if the crowd can be leveraged to aid in the process.

Crowdsourcing of Complex Tasks

Recent work has set a precedent of engaging novice crowd workers in completing complex tasks. Complex creative work is an example area of this research. Mechanical Novel [7] and Ensemble [6]. The crowd has also been recruited to complete complex tasks such as social data analysis through creating annotations and explanations to make sense of social data graphs [12]. In an example of citizen science crowdsourcing, FoldIt allowed ordinary web users to play a protein-folding game, which lead to real scientific discoveries [2]. The commonality between these various research is that the tasks were broken down in a way to encourage higher quality output from crowd workers. Similarly, the Find-Fix-Verify workflow utilized in the crowdsourced word processor Soylent leveraged the crowd to proofread written work [1]. We take inspiration from this workflow within our SOAR method workflow by microtasking the SOAR method for the crowd to determine the viability of using the crowd in reading research papers.

METHODS

In-Person Pilot Study

We recruited two graduate students in Computer Science and Engineering from UC San Diego to take part in our pilot study as "learners." Each learner selected two papers of comparable length and difficulty within the domain of human-computer interaction research. In individual sessions, we began by introducing the student to the concept of active reading and the SOAR method. We posed several open-ended questions to gain insight into the strategies and artifacts the students normally produce when reading academic literature. The student then chose one paper to read independently and one paper to read in collaboration with a researcher from this paper, who played the role of the crowd.

SOAR Method Tasks

In our SOAR method adapted for reading academic papers, we described the following four tasks to our learners:

- Select -read the paper and highlight key points (Figure 1 (top))
- Organize -organize and arrange the highlighted key points into a sensical structure (i.e. an outline or mind map) (Figure 1 (bottom))
- Associate -associate the ideas across the paper by creating a summary in the reader's own words
- Regulate -describe how the ideas relate to prior reading, knowledge, and the reader's own work

In the collaborative condition, learners chose which tasks they wished to do on their own and which they wished for the researcher to do for them except for the Regulate step, which the learner always did due to the adaptable and personal nature of the task. By allowing the learners to choose their tasks, we wanted to see if there was consensus between the two learners to inform our formative crowd study. Both learners chose to complete the Select task and chose to have the researcher complete the Organize and Associate tasks.

Formative Crowd Study

In this formative study, we first wanted to see if the crowd would be able to perform parts of the active reading process and produce artifacts of comparable quality as learners. We utilized the same papers used by the pilot learners and additional papers annotated by the researchers to evaluate the

quality of the outputs, recruiting 14 crowd workers through Amazon Mechanical Turk: seven for the Selection task and seven for Association task (two participants were excluded due to not following instructions). Workers who completed the Selection task were explicitly instructed not to complete the Associate task and vice versa.

SOAR Method for the Crowd

From the feedback provided by our pilot learners, we removed the Organize step to simplify the task. We utilized two Human Intelligence Tasks (HITs) to separate the SOAR method into two microtasks:

- Select –the crowd was instructed to read through a section of a paper and highlight important points
- Associate –using the highlighted annotations, create a coherent summary of a section of a paper

RESULTS

In-Person Pilot Study

Quantitative Statistics

We measured the word count of the summaries produced by individual learners and by the researchers in the pilot study. The average word count of learner summaries was 171.5 words (SD=4.95), and the average word count of the researcher summaries was 295.5 words (SD=61.52). We did not find a significant difference in word count between the individual learners and the researchers (t=-2.64, df=1, p=.23). We also measured the time spent on each of the SOAR tasks in both the individual and collaborative reading conditions (Figure 4). The total amount of time spent on the reading process was longer in the collaborative condition, but the time spent by the individual learner was shorter.

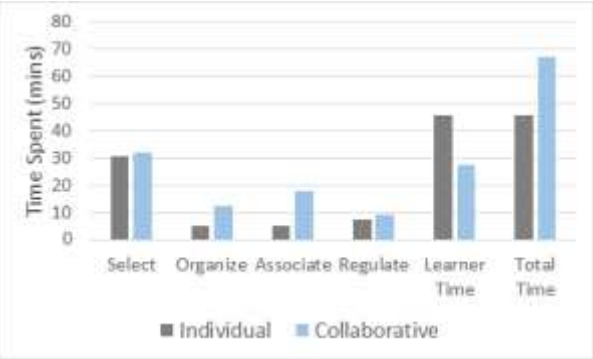

Figure 2. Time spent (in minutes) on each task in both the Individual and Collaborative SOAR reading condition.

Learner Perceptions of the SOAR Method

Learners reported that they normally read research papers in a linear manner, highlighting points and taking notes along the way. At the individual level, they found the SOAR method useful for structuring their thought process while reading. Both learners also found the Regulate step especially useful because they thought about how the paper related to their own research. One learner noted that the summary "helped [him] focus on the main points of the paper." However, both learners thought the Organize step to be repetitive and unnecessary. Learners also both felt that the collaborative reading process was useful because they could focus on one task at a time. One learner reported that she "felt less distracted while reading." The off-loading of tasks to a researcher allowed the learner to focus their attention on other tasks while still gleaning the main points from the paper from the researcher's summary.

Formative Crowd Study

Selection (Annotation) Task

We measured the number of annotations in each section by the crowd and compared it to the number of annotations selected by researchers in the pilot study. On average, the crowd produced 5.11 (SD=2.46) highlighted annotations per section of the paper. For researchers and learners, the average amount of annotations per section was 4.67 (SD=3.06). There was no significant difference between the count of annotations per section (x^2=6, df=4, p=.19). We also measured how many crowd annotations matched those of the learners from the pilot study. About 46% of the annotations produced by the crowd workers matched those of the learners.

Association (Summarization) Task

We examined four measures in the crowd summaries: expert-reviewed quality, number of correct idea units, word count, and time spent. To assess the quality of the crowd summaries, we had each of the researchers of this paper independently provide a Likert-scale rating (1 being of lowest quality, 5 being of highest quality) for the quality of seven summaries. Table 1 displays the descriptive statistics for each of our three measures.

We found a significant correlation between word count of the summaries and the average quality rating (r=0.85, df=5, p=.01) (Figure 4). We also found marginally significant correlation between time spent on the summary and average quality rating (r=0.71, df=5, p=.08). Furthermore, we measured if there was a difference between the length of the crowd summaries and the summaries from the pilot study. There were no significant differences in word count between the individual pilot summaries and the crowd summaries (t=-1.16, df=2, p=.37) or between the collaborative pilot summaries and the crowd summaries (t=-4.53, df=1, p=.14).

	Quality	Correct Idea Units	Word Count	Time Spent (secs)
Mean	2.60	3.78	82.29	436.57
SD	1.34	2.82	47.77	392.14

Table 1. Mean and standard deviation of the four measures of the crowd summaries

DISCUSSION

In this section, we discuss learner perceptions of the SOAR method at both the individual and collaborative levels. We

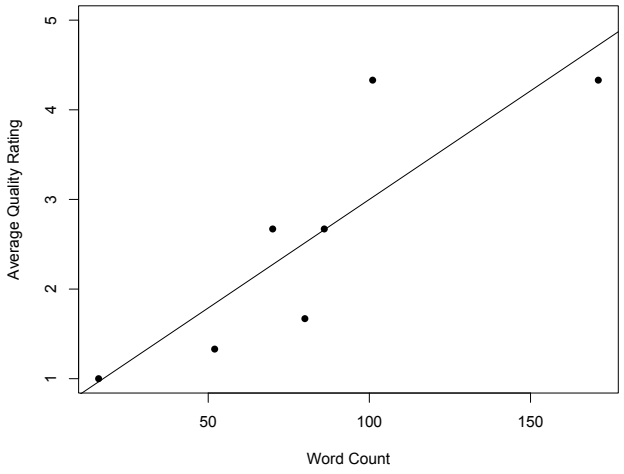

Figure 3. We found a significant correlation between word count and average quality rating of the crowd summaries.

also discuss the reliability of using the crowd to generate high quality summaries of research papers.

Learners Found the SOAR Method Useful and Enjoyed Collaborative Reading

Our pilot learners held generally favorable views towards the SOAR method, both in the individual and collaborative conditions. The method helped them structure their thoughts in reading better while the off-loading of tasks to researchers in the collaborative condition allowed them to focus their time more efficiently. While the total time to complete the paper was greater in the collaborative reading condition, the time spent by the individual learners was decreased. Learners also generally felt they understood the paper's study design and key findings from the researcher's summary even though they did not complete all the reading steps themselves. However, learners both found the Organize step repetitive, meaning the SOAR method may need to be better adapted if used for reading research papers. Importantly, learners found the Regulate step especially helpful because of its relation to integrating the paper with their own work. These self-reports of the usefulness of the SOAR method align with prior work on the SOAR method's primary benefit being the metacognitive component of the Regulate step [3]. These self-reports may point to the potential value of using the SOAR method for literature reviews for researchers.

The Crowd Produced Mixed-Quality Artifacts

The crowd seemed to have difficulty determining the important points of the paper from their selected annotations. They often focused on non-essential keywords or other irrelevant points. In addition, the crowd also produced mixed quality summaries with some being of very low quality and others being of relatively high quality. The tasks may have lacked enough context and guidance for the crowd, particularly for the complex task of reading and understanding research papers. However, those who wrote more in their summaries were also rated higher in terms of quality. This result mirrors that of [13]'s, which showed that longer feedback was correlated with its higher subjective quality ratings.

Domain Expertise is Important in Producing High Quality Annotations and Summaries

Learners in our pilot study found the annotations and summaries produced by the researchers to be useful and informative. However, the crowd-produced annotations and summaries often contained irrelevant points or low quality output. While the crowd has been shown to complete complex tasks, the domain specificity and knowledge required to understand complex research papers may have been the main reason for low crowd performance.

FUTURE WORK

Learner-Crowd Collaboration

The collaboration of learners with the crowd in the SOAR method may aid in improving crowd output quality. The expertise of the learner may provide the context and guidance for the crowd lacking in this current study. Collaboration between crowd workers and a task assigner has been shown to be useful in producing high quality artifacts [6]. We plan to examine this collaboration within the SOAR method and in aiding the literature review process.

Community-Sourcing the SOAR Method

Learners found collaboration in the SOAR method with a researcher helpful. It may be that a collaborative SOAR method may be most useful if conducted by a crowd with greater domain expertise. We believe that this strategy may be beneficial in settings where a shared community would benefit from an overview of literature such as specific research labs. We will explore utilizing the SOAR method within a community of researchers to investigate the effectiveness of collaboration in these settings.

CONCLUSION

This paper examines the use of the SOAR research method for conducting literature reviews, a task that is often difficult and time-consuming. We found that learners found the SOAR method useful for guiding their reading process at the individual level. Learners also found the collaborative reading component to be more focused and time efficient. We then explored leveraging the crowd to aid in the process. The crowd seemed to have difficulty with consistently creating high quality output potentially due to the difficulty and lack of context within the task. We believe that greater guidance for the crowd and collaboration with learners may improve the crowd's output quality. Future work will seek to explore this learner-crowd collaboration as well as community-sourcing of the SOAR method.

ACKNOWLEDGMENTS

We thank Dr. Steven Dow for his feedback and encouragement in this work. We also thank Ailie Fraser and Ariel Weingarten for participating in our pilot study. Finally, we would like to thank our classmates of our Crowdsourcing Research for useful peer feedback and discussion.

REFERENCES

1. Michael S Bernstein, Greg Little, Robert C Miller, et al. 2010. Soylent: a word processor with a crowd inside. *Proceedings of the 23nd annual ACM symposium on User interface software and technology*: 313–322. http://doi.org/10.1145/1866029.1866078
2. Seth Cooper, Firas Khatib, Adrien Treuille, et al. 2010. Predicting protein structures with a multiplayer online game. *Nature* 466, 7307: 756–760. http://doi.org/10.1038/nature09304
3. Tareq A. Daher and Kenneth A. Kiewra. 2016. An investigation of SOAR study strategies for learning from multiple online resources. *Contemporary Educational Psychology* 46: 10–21. http://doi.org/10.1016/j.cedpsych.2015.12.004
4. Nathan Hahn, Joseph Chang, Ji Eun Kim, and Aniket Kittur. The Knowledge Accelerator: Big Picture Thinking in Small Pieces. http://doi.org/10.1145/2858036.2858364
5. Dharma Jairam, Kenneth A. Kiewra, Sarah Rogers-Kasson, Melissa Patterson-Hazley, and Kim Marxhausen. 2014. SOAR versus SQ3R: a test of two study systems. *Instructional Science* 42, 3: 409–420. http://doi.org/10.1007/s11251-013-9295-0
6. Joy Kim, Justin Cheng, and Michael S Bernstein. Ensemble: Exploring Complementary Strengths of Leaders and Crowds in Creative Collaboration. http://doi.org/10.1145/2531602.2531638
7. Joy Kim, Sarah Sterman, Allegra Argent, Beal Cohen, and Michael S Bernstein. Mechanical Novel: Crowdsourcing Complex Work through Reflection and Revision. http://doi.org/10.1145/2998181.2998196
8. Valerie Matarese. 2006. Introductory Course on Getting to Know Journals and on Browsing a Research Paper: First Steps to Proficiency in Scientific Communication. *Croatian Medical Journal* 47, 5: 767–775.
9. Allen H Renear and Carole L Palmer. Strategic Reading, Ontologies, and the Future of Scientific Publishing.
10. Francis Pleasant Robinson and Francis Pleasant. 1970. *Effective study*. Harper & Row.
11. Victoria Uren, Simon Buckingham Shum, Michelle Bachler, and Gangmin Li. 2006. Sensemaking tools for understanding research literatures: Design, implementation and user evaluation. *Int. J. Human-Computer Studies* 64: 420–445. http://doi.org/10.1016/j.ijhcs.2005.09.004
12. Wesley Willett, Jeffrey Heer, and Maneesh Agrawala. Strategies for Crowdsourcing Social Data Analysis.
13. Alvin Yuan, Kurt Luther, Markus Krause, Sophie Isabel Vennix, Steven P Dow, and Bjorn Hartmann. 2016. Almost an Expert: The Effects of Rubrics and Expertise on Perceived Value of Crowdsourced Design Critiques. *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*: 1005–1017. http://doi.org/10.1145/2818048.2819953
14. Xiaolong Zhang, Yan Qu, C Lee Giles, and Piyou Song. CiteSense: Supporting Sensemaking of Research Literature.