Utilizing Automated Scaffolding Strategies to Improve Students' Reflections Writing Process

Saira Anwar
Department of Multidisciplinary Engineering
Texas A&M University
College Station, USA
sairaanwar@tamu.edu
https://orcid.org/0000-0001-6947-3226

Ahmed Ashraf Butt
School of Engineering Education
Purdue University
West Lafayette, USA
butt5@purdue.edu
https://orcid.org/0000-0003-2047-8493

Muhsin Menekse
School of Engineering Education and Department of Curriculum and Instruction
Purdue University
West Lafayette, USA
menekse@purdue.edu
https://orcid.org/0000-0002-5547-5455

***Abstract*—This study explored the effectiveness of scaffolding in students' reflection writing process. We compared two sections of an introductory computer programming course (N = 188). In Section 1, students did not receive any scaffolding while generating reflections, whereas, in Section 2, students were scaffolded during the reflection writing process. Student reflections were collected using two versions of the CourseMIRROR application (standard version in Section 1 and adaptive version in Section 2). By using Natural Language Processing (NLP) algorithms, the app calculated a reflection specificity score for each reflection. We conducted an independent sample t-test between the students' reflection specificity scores in these two sections. The results indicated that students using adaptive versions wrote more specific reflections than students using the standard version of the app, suggesting that scaffolding helped students write more specific reflections, which may be helpful in their overall learning outcomes in an introductory computer programming course.**

Keywords— Scaffolding, mobile application, reflection specificity, NLP algorithms.

I. Introduction

Reflection is considered an important cognitive process that helps students' learning and academic performance [1], [2]. Prior studies have shown that the reflection process keeps students cognitively engaged [3], [4]. However, it is often challenging to implementing reflection activities in large lecture classes, where the instructor has insufficient resources and time to interact frequently and provide feedback to students [5], [6]. This is especially challenging in large science, technology, engineering, and mathematics (STEM) courses as they generally prefer pedagogies based on the content delivery and assessment rather than reflection activity [7]. Also, students who participate in such activities tend to complete the activity with minimal effort and may write content that is not reflective of actual course content [8].

Considering the concerns of involving students in large STEM courses, we designed a mobile application (i.e., CourseMIRROR) based on the Reflection Informed Learning and Instruction (RILI) model [9]. The RILI model expands on the hypothesis that students' meaningful reflection can improve their learning experiences. The RILI model suggests that when both students and instructors are engaged in the diagnosis process of identifying gaps and difficulties of the lecture and concepts, it helps instructors provide effective feedback, improving students' knowledge and skills. Also, students can utilize various responses to improve their learning, such as seeking resources or help materials, enhancing interaction with peers and instructional team, etc.

In this study, to introduce such a model in large classes and ensure students' participation effectively, we used CourseMIRROR mobile app [8], [10], [11], which prompts students to reflect on their learning experiences after each lecture. The students are prompted to generate reflections based on two perspectives: 1) muddiest point (MP), and 2) point of interest (PI). Furthermore, the application uses Natural Language Processing (NLP) algorithms to summarize students' reflections for each lecture. These summaries are made available to both students and teachers. It is noteworthy that the summary's quality depends on how specific students were writing their reflection on each perspective of reflection writing, i.e., MP and PI. To quantify the reflection specificity, a NLP converts each written reflection into an equivalent specificity score ranging from 1-4 points where score 1 indicates shallow reflection specificity and score 4 indicates excellent reflection specificity.

II. Related Work

In literature, researchers have conceptualized reflection in various ways. However, reflection can be broadly seen as a cognitive process where students take a step back from their current learning experience, reflect and comprehend what they learned, process their new learning experience, and connect this new experience with their previous knowledge or learning experiences [12]. Existing STEM studies have shown that reflection is a vital instructional approach that improves students' learning and academic performance (e.g., [13]). Furthermore, engaging students in the reflection process helps in developing their reflective thinking [14], allows them to resolve misconceptions [15], develop self-regulatory skills, and is correlated with their motivation such as achievement goals (e.g., [16]), and self-efficacy (e.g., [11]).

While studies have discussed the potential of engaging students in reflective practices, some studies also argued for providing guidelines and scaffolding during the reflection process to achieve the desired benefits of the reflective practices [16], [17]. Scaffolding in the literature indicates the support provided to students while performing a task [18], [19]. Providing feedback to the students while reflection writing helps them become aware of their past learning process and thus not only promotes students' self-reflection but, in doing so, has been shown to improve their reflection writing [20]. Various mechanisms have been used to introduce the reflection process in STEM classrooms, including journaling (e.g., [21]), questionnaires (e.g., [22]), survey (e.g., [23]), or exit tickets (e.g., [24]). However, introducing such a mechanism in large STEM classes was found to be challenging due to resources and time constraints.

2023 IEEE Frontiers in Education Conference (FIE) | 979-8-3503-3642-9/23/$31.00 ©2023 IEEE | DOI: 10.1109/FIE58773.2023.10343485

Similarly, different activities (e.g., sketching, games) were used for scaffolding students to facilitate students' reflection process. However, studies haven't embedded a real-time personalized feedback mechanism to scaffold students during their reflection writing process. To this end, this study utilizes an educational technology application and the power of NLP algorithms to facilitate students in the reflection writing process while providing real-time scaffolding.

III. Research Design

This study employed a quasi-experimental study to investigate the effectiveness of real-time scaffolding provided through a CourseMIRROR application for all the students in large STEM classrooms.

A. Mobile Educational application

CourseMIRROR application has been developed in two versions: 1) standard version and 2) adaptive version for both iOS and Android platforms. In the standard version, students are asked to write reflections for two prompts (i.e., MP and PI prompts). The students were prompted to generate reflections based on two perspectives: 1) MP- where students described what they found most confusing in the lecture, and 2) PI – where students described what they found most interesting in the application. The submitted student reflections are then passed to the NLP algorithm [25] to calculate the NLP specificity score. The algorithm gives values in the range of 1-4 (1 being shallow and 4 being highly relevant to the reflection prompt or lecture consent). Moreover, two human coders evaluated the reflections using a rubric mentioned in [16] and subsequently assigned scores indicating their specificity. The agreement between the two coders was calculated, revealing a strong agreement with κ (MP) = .617 and κ (POI) = .652 [26]. Then, we evaluated the agreement between the specificity score generated by NLP algorithm and the human coding. The result of the study has shown a strong agreement with, κ (NLP) = 0.775 and κ (NLP) = 0.773, for each set of reflection data discussing the point of interest and muddiest point of lecture [27], respectively.

On the other hand, the adaptive version varies from the standard version on two accounts. First, the application interface is modified in the adaptive version to provide students with a more meaningful and user-friendly experience. The second and most important logical algorithmic difference is using the reflection specificity score to introduce the real-time scaffolding process. In the adaptive version, students are scaffolded to write detailed reflections to enhance students' metacognitive thinking. The CourseMIRROR application uses NLP algorithms to evaluate the specificity of each reflection during students' writing process. The scaffolding process is based on two components 1) students are guided with meaningful textual instruction based on their real-time NLP-based reflection specificity score in a dynamic manner and 2) students get a colored bar indicating the specificity of a reflection. Fig. 1 shows both components of the scaffolding process with the respective reflection writing process.

The red bar indicates a low specificity score, while the green indicates a higher specificity score. Also, with lower specificity scores, students got the instruction "Please tell us what you found confusing or unclear in today's class," while with higher specificity, students got a message of "Great, thanks!". Although students are scaffolded in the adaptive version, the process does not hinder students' reflection writing. Also, the process does not put any threshold specificity value on reflection submission. Like the standard version, the adaptive version thus allows students to submit final written reflections with any specificity score.

B. Site and participants

We collected the data from two sections of the first-year engineering programming course taught at a large U.S Midwestern University. In the course, students were introduced to computer programming by using MATLAB. Following a non-equivalent comparison group design based on the natural assembly of sections and time of data collection [28], we introduced a separate version of the application in each section of the course. One hundred twenty students in Section1 used the standard version of the application. Sixty-eight students in Section2 used the adaptive version. Students in both sections voluntarily participated and submitted reflections throughout the semester. In this study, we explored the reflection specificity of 11 lectures.

C. Data analysis

For the analyses, we used the reflection specificity score from two sections to conduct an independent sample t-test between subjects of two different sections for all lectures and both aspects of the reflection writing process, i.e., MP and PI. We tested the statistical assumptions before running the relevant analysis. We verified data normality with skewness and kurtosis, and Q-Q plots. For skewness, we considered the moderate to low skewed data. All the variables were moderately or low skewed as values remained greater than -1 and less than 1. For kurtosis, we considered the values between -2 and +2. Further, for homogeneity of variances, we used Levene's statistics. In the case of the violated condition, we used the results by not assuming the equality of variance using adjusted degrees of freedom, reducing the chances of Type-I error.

D. Results

In our comparisons, Levene's test was significant for the muddiest point (MP) values for lectures 1, 2, 6, and 10. Therefore, we used the t-test for equal variance not assumed. Similarly, for lectures 4 and 7, Levene's test for point of interest (PI) was significant, and we used

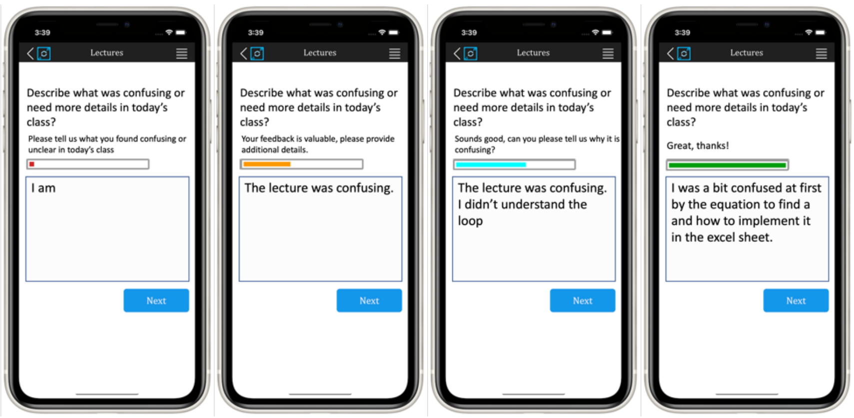

Fig. 1. Scaffolding process in the adaptive version of a figure caption.

a t-test for equal variance not assumed. The independent-sample t-test results are presented in Table 1 and Table 2, describing the results of PI and MP, respectively. In tables, ** indicates that $p < 0.05$.

The results indicated a significant mean difference in the students' reflection specificity score using the app's adaptive and standard versions. The students' reflection specificity score for the PI was significantly higher in the adaptive version for all lectures than the standard version. Also, except for lectures 2, 4, and 11, the mean MP reflection specificity score was significantly higher in the adaptive version than in the standard version. Although the mean MP's reflection specificity for lectures 2, 4, 11 is higher in the adaptive version, it is not statistically significant.

These results indicate that with the scaffolding process, students were able to write more specific reflections for all lectures and from both MP and PI perspectives. The Cohen's d effect size is between the range of 0.64 - 0.76, indicating a large effect size [29]. For MP, the Cohen's d value consistently showed a large effect ranging between .84 – 1.2, with some values greater than 1, indicating that the difference between the two means is larger than one standard deviation.

IV. Discussion & Conclusion

In this quasi-experimental study, using a reflection-informed learning and instruction (RILI) model [9], we investigated the effectiveness of real-time scaffolding during students reflection writing process in a mobile educational app. Our findings showed that providing scaffolding could facilitate students in writing detailed and specific reflections. Prior studies has also showed a relationship between the students' reflection specificity and learning outcomes [9], [30], [31]. Therefore, scaffolding students to construct specific reflections can enhance their learning outcomes.

This finding is consistent with the previous literature [20], [32] as it helped students organize their thoughts in structured ways while reflecting. Additionally, providing clear, and guided reflection prompts to the students could have enabled them to delve deeper into their lecture's learning experiences, leading to more comprehensive and meaningful reflections.

TABLE I. Table Reflection specificity differences between standard and adaptive versions on the point of interest

Lectures	Standard version			Adaptive version			df	*t*	*p*	Cohen's d
	N	M	SD	N	M	SD				
1	106	3.293	.661	44	3.659	.608	148	-3.163	.002**	.646
2	96	2.948	.686	41	3.537	.809	135	-4.353	<.001**	.725
3	68	2.780	.709	38	3.553	.724	104	-5.343	<.001**	.714
4	84	2.941	.628	37	3.287	.740	59.809	-2.555	.013**	.664
5	75	2.813	.692	35	3.429	.698	108	-4.333	<.001**	.694
6	80	2.888	.693	40	3.550	.677	118	-4.971	<.001**	.688
7	63	3.031	.621	44	3.455	.791	77.996	-2.964	.004**	.696
8	78	2/731	.767	44	3.455	.730	120	-5.090	<.001**	.754
9	68	2.882	.783	48	3.500	.684	114	-4.405	<.001**	.744
10	95	2.385	.655	25	3.480	.714	119	-7.306	<.001**	.667
11	39	2.539	.682	42	3.405	.767	79	-5.355	<.001**	.728

TABLE II. Table Reflection specificity differences between standard and adaptive versions on the muddies point

Lectures	Standard version			Adaptive version			df	*t*	*p*	Cohen's d
	N	M	SD	N	M	SD				
1	106	3.311	1.027	44	3.750	.719	113.315	-2.977	.004**	.948
2	96	3.063	.856	41	3.195	1.188	58.503	-0.647	.520	.966
3	68	2.956	.905	38	3.395	.823	104	-2.471	.015	.877
4	84	2.738	1.110	37	3.081	1.090	119	-1.575	.118	1.104
5	75	2.747	.856	35	3.429	.884	108	-3.852	<.001**	.865
6	80	2.538	.941	40	3.550	.711	99.143	-6.561	<.001**	.872
7	63	2.508	1.091	44	3.273	1.042	105	-3.634	<.001**	1.071
8	78	2.628	.941	44	2.955	1.140	120	-1.702	.091	1,017
9	68	2.882	.873	48	3.438	.796	114	-3.496	<.001**	.842
10	95	2.354	.767	25	2.840	1.248	28.892	-1.858	.073	.886
11	39	2.436	1.119	42	2.762	1.246	79	-1.236	.220	1.186

Furthermore, studies have shown that scaffolding reduces the students' cognitive load in complex learning task by breaking them into manageable and meaningful small tasks [33], [34]. In this experiment, we achieved this by providing clear instructions through the color bar and messages to help them focus on one aspect of the lecture (i.e., MP or POI) at a time and the clear instructions to improve their reflection. In doing so, it is possible that students were able to utilize their cognitive resources effectively, leading to comprehensive and in-depth reflection.

Moreover, writing more specific reflections invokes students' critical thinking about their learning process, specifically in terms of what they liked in each lecture and the difficult aspects of each lecture [30]. Our findings are important as they provide insights on using a scaffolding process in more than one way in a mobile application, which may be useful in many emerging educational technology tools, especially in this era of higher technology dependency in education.

These results must be viewed with few limitations. First, there was a small sample size specifically for section 2. Future studies may investigate these comparisons with a larger sample size. Second, in this study, the groups of students were naturally assembled based on the section, and no random assignment was performed. Third, the students were from a single course. In the future studies, such impact may be evaluated with other STEM courses. In addition, the studies may also examine the role of such scaffolding process on students learning and how this process helped instructors revise their lectures and future offerings of the course.

ACKNOWLEDGMENT

The research reported here was supported in part by a grant from the Institute of Education Sciences (R305A180477). The opinions expressed do not represent the views of the U.S. Department of Education.

REFERENCES

[1] J. R. Baird, P. J. Fensham, R. F. Gunstone, and R. T. White, "The importance of reflection in improving science teaching and learning," *J. Res. Sci. Teach.*, vol. 28, no. 2, pp. 163–182, 1991.

[2] J. I. Murphy, "Using focused reflection and articulation to promote clinical reasoning: An evidence-based teaching strategy," *Nurs. Educ. Perspect.*, vol. 25, no. 5, pp. 226–231, 2004.

[3] A. Y. Lee and L. Hutchison, "Improving learning from examples through reflection.," *J. Exp. Psychol. Appl.*, vol. 4, no. 3, p. 187, 1998.

[4] R. Moreno and R. E. Mayer, "Role of guidance, reflection, and interactivity in an agent-based multimedia game.," *J. Educ. Psychol.*, vol. 97, no. 1, p. 117, 2005.

[5] S. Anwar and M. Menekse, "Unique Contributions of Individual Reflections and Teamwork on Engineering Students' Academic Performance and Achievement Goals," *Int. J. Eng. Educ.*, vol. 36, no. 3, pp. 1018–1033, 2020.

[6] R. G. Baldwin, "The climate for undergraduate teaching and learning in STEM fields," *New Dir. Teach. Learn.*, vol. 2009, no. 117, pp. 9–17, 2009.

[7] E. DaMaren, D. Pearlston, and S. Mattucci, "Considerations for development of critical reflection skills in first-year engineering informed through student perspectives," in *Proceedings of the Canadian Engineering Education Association (CEEA)*, 2020.

[8] X. Fan, W. Luo, M. Menekse, D. Litman, and J. Wang, "Scaling reflection prompts in large classrooms via mobile interfaces and natural language processing," 2017, pp. 363–374.

[9] M. Menekse, "The reflection-informed learning and instruction to improve students' academic success in undergraduate classrooms," *J. Exp. Educ.*, vol. 88, no. 2, pp. 183–199, 2020.

[10] X. Fan, W. Luo, M. Menekse, D. Litman, and J. Wang, "CourseMIRROR: Enhancing large classroom instructor-student interactions via mobile interfaces and natural language processing," 2015, pp. 1473–1478.

[11] M. Menekse, S. Anwar, and S. Purzer, "Self-efficacy and mobile learning technologies: A case study of CourseMIRROR," in *Self-Efficacy in Instructional Technology Contexts*, Springer, 2018, pp. 57–74.

[12] A. A. Butt, S. Anwar, and M. Menekse, "WIP: Investigating the relationship between FYE students' reflections and academic performance across gender," in *2022 First-Year Engineering Experience*, 2022.

[13] M. D. N. Lew and H. G. Schmidt, "Self-reflection and academic performance: is there a relationship?," *Adv. Health Sci. Educ.*, vol. 16, no. 4, p. 529, May 2011, doi: 10.1007/s10459-011-9298-z.

[14] N. Sabag, E. Trotskovsky, and S. Waks, "Engineering design projects as a reflection promoter," *Eur. J. Eng. Educ.*, vol. 39, no. 3, pp. 309–324, 2014.

[15] A. S. Halim, S. A. Finkenstaedt-Quinn, L. J. Olsen, A. R. Gere, and G. V. Shultz, "Identifying and Remediating Student Misconceptions in Introductory Biology via Writing-to-Learn Assignments and Peer Review," *CBE Life Sci. Educ.*, vol. 17, no. 2, pp. ar28–ar28, Jun. 2018, doi: 10.1187/cbe.17-10-0212.

[16] D. H. Stratton, S. Anwar, and M. Menekse, "How Do Engineering Students' Achievement Goals Relate to their Reflection Behaviors and Learning Outcomes?," in *2017 ASEE Annual Conference & Exposition*, 2017.

[17] D. Boud and D. Walker, "Promoting reflection in professional courses: The challenge of context," *Studies in Higher Education*, vol. 23, no. 2. Routledge, pp. 191–206, Jan. 01, 1998. doi: 10.1080/03075079812331380384.

[18] B. R. Belland, A. E. Walker, N. J. Kim, and M. Lefler, "Synthesizing Results From Empirical Research on Computer-Based Scaffolding in STEM Education: A Meta-Analysis," *Review of Educational Research*, vol. 87, no. 2. American Educational Research Association, pp. 309–344, Apr. 01, 2017. doi: 10.3102/0034654316670999.

[19] S. Puntambekar and R. Hubscher, "Tools for scaffolding students in a complex learning environment: What have we gained and what have we missed?," *Educ. Psychol.*, vol. 40, no. 1, pp. 1–12, 2005.

[20] J. Novakovich, “Fostering critical thinking and reflection through blog-mediated peer feedback,” *J. Comput. Assist. Learn.*, vol. 32, no. 1, pp. 16–30, Feb. 2016, doi: 10.1111/jcal.12114.

[21] A. M. Humble and E. Sharp, “Shared journaling as peer support in teaching qualitative research methods.,” *Qual. Rep.*, vol. 17, p. 96, 2012.

[22] K. Lethbridge, M.-A. Andrusyszyn, C. Iwasiw, H. K. S. Laschinger, and R. Fernando, “Assessing the psychometric properties of Kember and Leung’s reflection questionnaire,” *Assess. Eval. High. Educ.*, vol. 38, no. 3, pp. 303–325, May 2013, doi: 10.1080/02602938.2011.630977.

[23] K. Wiezbicki-Stevens, *Metacognition: Developing self-knowledge through guided reflection*. University of Massachusetts Amherst, 2009.

[24] C. Amaro-Jiménez, “Preservice teachers’ reflections of their involvement in a home-school connection project in teacher education,” *Biling. Res. J.*, vol. 39, no. 1, pp. 69–85, Jan. 2016, doi: 10.1080/15235882.2016.1139520.

[25] A. Magooda, D. Litman, A. A. Butt, and M. Menekse, “Improving the quality of students’ written reflections using natural language processing: Model design and classroom evaluation,” in *International Conference on Artificial Intelligence in Education*, Springer, 2022, pp. 519–525.

[26] D. G. Altman, *Practical statistics for medical research*. London:UK: Chapman & Hall/CRC, 1990.

[27] A. A. Butt, S. Anwar, A. Magooda, and M. Menekse, “Comparative analysis of the rule-based and machine learning approach for assessing student reflections,” in *Proceeding of International Society of the Learning Sciences (ISLS)*, 2022, pp. 1577–1580.

[28] D. T. Campbell and J. C. Stanley, *Experimental and quasi-experimental designs for research*. Ravenio Books, 2015.

[29] J. Cohen, *Statistical power analysis for the behavioral sciences*. Academic press, 2013.

[30] D. Boud, R. Keogh, and D. Walker, *Reflection: Turning experience into learning*. Routledge, 2013.

[31] D. Heo, S. Anwar, and M. Menekse, “The Relationship Between Engineering Students’ Achievement Goals, Reflection Behaviors, and Learning Outcomes,” *Int. J. Eng. Educ.*, vol. 34, pp. 1634–1643, Sep. 2018.

[32] J. D. Bain, C. Mills, R. Ballantyne, and J. Packer, “Developing Reflection on Practice Through Journal Writing: Impacts of variations in the focus and level of feedback,” *Teach. Teach.*, vol. 8, no. 2, pp. 171–196, May 2002, doi: 10.1080/13540600220127368.

[33] N. R. Aljohani and H. C. Davis, “Learning analytics and formative assessment to provide immediate detailed feedback using a student centered mobile dashboard,” in *Proceedings of the 2013 Seventh International Conference on Next Generation Mobile Apps, Services and Technologies*, in NGMAST ’13. USA: IEEE Computer Society, 2013, pp. 262–267. doi: 10.1109/NGMAST.2013.54.

[34] E. Chew, H. Snee, and T. Price, “Enhancing international postgraduates’ learning experience with online peer assessment and feedback innovation.,” *Innov. Educ. Teach. Int.*, vol. 53, no. 3, pp. 247–259, Jun. 2016.