WIP: SABERR, A Structured Error-Based Assessment in AI Education

Mariana Alvidrez
Curriculum & Instruction
New Mexico State University
Las Cruces, U.S.A
malvidrz@nmsu.edu

Christabel Wayllace
Computer Science
New Mexico State University
Las Cruces, U.S.A
cwayllac@nmsu.edu

Ruth Torres Castillo
Innovative Media Research and Extension
New Mexico State University
Las Cruces, U.S.A
rutorres@nmsu.edu

***Abstract*—This innovative practice WIP paper presents SABERR, a novel formative assessment designed to leverage errors as a learning resource in an artificial intelligence (AI) course. Previous research suggests that utilizing mistakes as learning tools enhances students' problem-solving abilities, metacognitive skills, and deepens their conceptual understanding across STEM fields. The intended outcomes of SABERR are supporting and enhancing students' learning by using their mistakes as resources to improve their awareness about their own knowledge and problem-solving strategies. The application design of the SABERR assessment unfolds in three structured phases which involve a reflective process where students are asked to articulate and explain the reasoning behind their initial errors. Our preliminary findings reveal that students initially struggled with error analysis and articulating knowledge gaps. However, a structured, step-by-step process significantly enhanced their error identification and correction abilities. The SABERR assessment approach significantly shifted student behavior towards prioritizing mastery of material, which was evidenced by increased engagement in class discussions and a reduction in academic dishonesty. Students reported improved ability to detect and correct errors, with 95% acknowledging enhanced understanding of AI concepts. However, this method also increased the workload for instructors and teaching assistants. Overall, 91% of students believed that a productive attitude towards mistakes would benefit their future careers, underscoring the effectiveness of SABERR in fostering a deeper, more analytical learning process among computer science students. The findings from our pilot study emphasize that the SABERR assessment approach not only enhances students' conceptual understanding and problem-solving capabilities but also promotes metacognitive skill development. By framing errors as learning tools, the approach encourages students to become active, reflective participants in their own learning processes. These findings also highlight the need for assessment approaches that enable both instructors and students to reframe errors as learning opportunities, thereby changing their perceptions.**

***Index Terms*—Artificial intelligence education, learning from errors, second attempt assessment, meta cognition skills**

I. BACKGROUND AND SIGNIFICANCE

Research in mathematics and computer science (CS) education, foundational disciplines for artificial intelligence (AI), has focused on how mistakes contribute to advancing students' learning and developing their metacognitive skills from different perspectives [1], [2]. For example, in CS, an implicit practice of error analysis appears commonly in the phase of debugging a program [3], [4]. However, in the context of AI education, students' errors may derive from their limited conceptual understanding, lack of fundamental notions, or improper problem-solving strategies and approaches [5], [6].

This work-in-progress paper presents the Second Attempt: Beyond Errors (SABERR) assessment method. We see errors as initial steps in the learning process instead of conclusive evidence of failure. Aligned with this perspective, SABERR integrates the concept of 'SABER' (meaning 'to learn' in Spanish) with a focus on errors. This assessment uses a constructive metacognitive-based approach that leverages a second-attempt assessment strategy, wherein students actively participate in identifying, analyzing, and justifying their errors to facilitate a deeper understanding of their existing knowledge. The primary goal of this error-based metacognitive analysis is to assist students in reflecting on their assumptions and problem-solving strategies, which led to errors. It emphasizes students' scaffolding process to think about concepts and ideas crucial to solving a problem that students either overlooked or addressed inadequately, as well as incorrect inferences they made about a problem. Additionally, this error-based assessment strategy addresses particular challenges in AI education, such as students' limited foundational knowledge in mathematics and computer science [7], their lack of confidence, and their tendency to prioritize performance over learning [8], [9].

A. *Why a Learning from Mistakes Approach Assessment?*

Supporting students' conceptual thinking across STEM-related fields using mistakes[1] as learning tools is a complex and challenging process due to traditional practices that frame errors as something to avoid and students who err as incapable of learning [10]–[12]. However, research in mathematics [13]–[16], physics [17], and CS education [3], [18], [19] indicates that engaging students in analyzing their mistakes using student-based constructivist approaches enhance their problem-solving skills and deepen their conceptual understanding [20], [21] through trial and error skills [5], [17], [22]–[25]. Further, a common and almost implicit practice in CS is debugging.

Previous research indicates that debugging enables students not only to detect and correct their coding errors but also to

[1]In this paper, we use the terms "errors" and "mistakes" interchangeably

become active agents of their own learning, as they realize errors are not indicative of learning deficits but rather opportunities for enhancing algorithmic solutions [18], [26], [27]. Another strategy documented in CS involves offering students a "second chance" at tests. Previous research states that this method not only enhances students' conceptual understanding, problem-solving skills, and metacognitive abilities but also improves their mastery of the material and reduces their anxiety [28], [29].

B. Learning from Mistakes and the Development of Metacognitive Skills

Enhancing students' learning by using their mistakes as resources involves adopting instructional strategies that encourage students to reflect on their existing knowledge to develop their metacognitive skills [2]. Metacognition is an individual's awareness and understanding of their own cognitive processes involved in problem-solving processes, which encompasses thoughts about thoughts, and knowledge about knowledge or reflections about actions [30], [31].

Previous studies have demonstrated that programming students who engage in discussions to reflect on their previous knowledge, their mistakes, and their successful and unsuccessful strategies to correct them tended to display higher levels of metacognitive awareness [32]. This reflection includes self-explanation—a metacognitive skill to identify which aspects of their prior knowledge were accurate and which were imprecise, vague, unclear, or incorrect [33], [34].

Further, errors may provoke cognitive conflicts that force students to attain a deep understanding of concepts and relations between them [23]. For instance, previous research showed how being aware of their own knowledge facilitates higher levels of abstraction and meta-analysis in CS courses [4], allowing students to review both successful and mistaken steps to plan strategically for solving similar problems in the future [35], [36]. This body of research is significant because even when these studies do not focus on the benefits of using learning from mistakes approaches in AI; they provide a deeper understanding of the salience of the power of active student participation to construct deeper levels of understanding through embracing errors as learning opportunities.

II. The Innovative Practice: The SABERR Assessment

The 'Second Attempt: Beyond Errors' (SABERR) assessment approach asks students to use their prior knowledge and logical reasoning to detect, analyze, explain, and justify the reasoning behind the mistakes made in their initial attempt. Furthermore, students are encouraged to formulate an action plan to refine their understanding and problem-solving strategies. Students do this by comparing their erroneous approaches and solutions to partial standard solution steps or alternatives provided by the instructor.

A. The SABERR Assessment Objectives

This assessment approach fosters the development of students' self-explanation skills allowing them to assess the accuracy of their conclusions and responses by examining and leveraging their prior knowledge. Further, the SABERR assessment was designed to support students in identifying not only their knowledge gaps, thereby enhancing their error awareness but also to evaluate whether their prior knowledge was applied correctly or incorrectly to new concepts in their initial attempt. Furthermore, SABERR was crafted to foster an iterative reflective learning process, encouraging students to critically evaluate their understanding and application of knowledge through an error-analysis reflection to readjust and redefine their problem-solving strategies. It is important to clarify that the first attempt is not graded, encouraging students to prioritize mastery over performance.

B. The SABERR Assessment Procedure

The SABERR assessment unfolds in three structured phases: (1) students complete assignments such as quizzes or homework, (2) after submission, students receive access to intermediate steps and answers for all the problems, (3) the final phase involves a reflective process (described below) where students are asked to articulate and explain the reasoning behind their initial responses by comparing their procedures and answers with those provided in step (2). Included below is a rubric to ensure students' understanding at the level of analysis the instructor is expecting (Figure 1).

Reflective Process

1) ***STEP 1:*** Mistake Identification
   1.1 Review Assignment Responses: Review your assignment procedures and responses based on the answer key provided by the instructor.
   1.2 Identify Errors: For each incorrect answer, you should note the mistake you made. Please highlight them.
2) ***STEP 2:*** Conceptual Analysis
   2.1 List Concepts Involved: For each question, list the (mathematical concepts, formulas, or procedures) involved in solving it.
   2.2 Knowledge Assessment: Categorize these concepts into these groups:
      1) Concepts you knew and applied correctly.
      2) Concepts you did not know but you understand, now.
      3) Concepts you did not know and still do not understand.
3) ***STEP 3:*** Importance of Concepts
   3.1 Concept Significance: Write a brief explanation of why the concepts you didn't know were crucial for solving the problem.
4) ***STEP 4:*** Conceptual Connections
   4.1 Interlinking concepts: Describe how the concepts you knew are connected to the ones you did not, and you understand now. This can include how one concept builds on another or how different concepts interact within a problem.
5) ***STEP 5:*** Reflective Learning

5.1 Learning from Analysis: Write a reflective summary of what you have learned from this analysis. This should include:
   1) Insights gained about your understanding and misconceptions.
   2) Strategies you might employ in the future to avoid similar mistakes.
   3) How has this exercise affected your approach to learning and understanding procedures, mathematical concepts, and formulas?

6) ***STEP 6:*** Action Plan: Bonus Points

6.1 Developing an Improvement Strategy: Based on your reflections, create a personal action plan that outlines:
   1) Specific areas you need to improve.
   2) Resources or study methods you will use.

7) ***STEP 7:*** Use the rubric to ensure the quality of your analysis.

Criteria	**Excellent (20 pts)**	**Good (18 pts)**	**Satisfactory (15 pts)**	**Needs Improvement (10 pts)**
1. Mistake Identification	Identifies all mistakes with accurate descriptions.	Identifies most mistakes with adequate descriptions.	Identifies some mistakes but with incomplete descriptions.	Struggles to identify mistakes or descriptions are unclear.
2. Conceptual Analysis	Accurately lists and categorizes all relevant concepts. Demonstrates deep understanding of known concepts.	Lists and categorizes the most relevant concepts accurately. Good understanding of known concepts.	Lists and categorizes some relevant concepts but misses others. Basic understanding of known concepts.	Incomplete or inaccurate listing of concepts. Poor understanding of known concepts.
3. Importance of Concepts	Provides a detailed and insightful explanation of the importance of unknown concepts in problem-solving.	Provides a clear explanation of the importance of unknown concepts but lacks detail.	Provides a basic explanation of the importance of unknown concepts but may miss key points.	Provides little to no explanation of the importance of unknown concepts.
4. Conceptual Connections	Excellently describes the interconnections between known and unknown concepts, demonstrating deep insight.	Describes connections between known and unknown concepts with a good level of understanding.	Describes some connections, but the explanation lacks depth or clarity.	Struggles to describe connections between concepts or the explanation is very superficial.
5. Reflective Learning	Provides a comprehensive, insightful reflection that demonstrates deep learning and self-awareness.	Provides a clear and honest reflection, showing good learning and self-awareness.	Reflection is somewhat superficial or lacks detail but shows some learning and self-awareness.	Reflection is very basic, lacks insight, and shows limited self-awareness or learning.
6. Action Plan: Bonus	**Excellent** **(5 pts)** Provides a detailed and realistic action plan.	**Good** **(4 pts)** Identifies most areas that need improvement and provides realistic study methods.	**Satisfactory** **(2 pts)** Identifies some areas for improvement and study methods.	**Needs Improvement** **(1 pt.)** Struggles to identify areas to improve or study methods to use.

Fig. 1. Rubric.

III. Methodology

This pilot IRB-approved study occurred at a large public Hispanic Serving Institution (HSI). More than 64% of the students enrolled at the university identify themselves as Hispanic, and more than 50% receive Federal Pell Grant Aid, which indicates low-income status. Within this larger context, the CS department has developed a series of AI courses for CS students aiming to support students in developing and practicing their AI knowledge.

A. Research Context and Participants

Participants in this pilot study were 27 students enrolled in one of those courses. The class met twice a week and students were encouraged to attend their instructor's office hours. To the extent possible, then, each student had the opportunity to talk with the instructor one-on-one about their assessments.

B. Data Collection and Analysis

Since the aim of this pilot study was to investigate students' learning from mistakes, the authors employed research methods that drew on a concurrent mixed methods approach. By employing a concurrent mixed method design and drawing on evidence from both qualitative and quantitative data, the study provided evidence of AI students' learning from mistakes, in particular through their level of analysis in their second-attempt assessment responses [37], [38]. The first author also applied a 12-item survey (four opened-ended questions and eight closed-ended questions) focusing on the students' perspectives and understanding of the learning from errors assessment. This survey instrument was prepared and revised by two of the authors and validated by a faculty member from the psychology department of this institution. Qualitative data sources included student artifacts and participant observations. Researchers drew on this design because it provides triangulation and complementary data [37] that allows for more accurate feedback. Data collected were analyzed through a constant comparative method [39] in which data are coded, sorted, and organized in a structure to emerge into relevant themes.

IV. Preliminary Outcomes

In the initial stage of this pilot study, the research team implemented an error-based assessment that required students to reflect on and explain the reasoning behind their mistakes, without providing further guidance. Analysis of the students' responses revealed that most students were unable to effectively analyze their errors, often providing answers without understanding their mistakes. Based on this analysis, we decided to refine the assessment to encourage students to identify the knowledge gaps that led to their errors. However, many students still struggled to detect, analyze, and correct their mistakes or to connect these mistakes to their lack of knowledge or misapplication of what they knew. Following further refinements, we introduced the SABERR assessment. In this section, we present the findings from this revised assessment strategy.

A. The SABERR Assessment Implementation: Instructor Perspectives

Overall, implementing the SABERR assessment required a shift in student behavior, which could only be effectively reinforced through grading. However, as students acclimated to this approach, they began to prioritize learning over the pursuit of high grades, particularly because many were already achieving satisfactory grades (the average and median grade for the class at the time of writing is 84.9% and 87.5% respectively). This shift in mindset reflects a deeper engagement with the material and a greater emphasis on mastery rather than mere performance.

The SABERR approach was not limited to the second-attempt assessment; instead, it was embraced throughout the course activities. For example, our field notes show that students were consistently motivated to engage actively in class discussions, irrespective of their certainty regarding the correct answer. Further, they felt empowered to identify and openly communicate any type of errors made by them and by their instructor, a practice uncommon in American classrooms where instructors tend to be error-averse [12].

Moreover, our qualitative data indicated that this error-friendly approach significantly reduced academic dishonesty, as students were assessed on their analytical skills rather than the correctness of their answers. For instance, in the first assignment, which involved introductory AI questions with readily available answers online, at least three students initially resorted to copying answers from the Internet. However, their use of incorrect procedures indicated a lack of comprehension. As the course progressed, those students comments indicated a greater inclination to understand the underlying procedures, which facilitated easier identification of errors. Remarkably, one student expressed relief from the pressure to cheat, acknowledging the assurance that their focus on learning would not be penalized for occasional incorrect answers.

We also identified a significant challenge with the SABERR assessment approach: a substantial increase in workload for both instructors and teaching assistants. This increase was due to the detailed feedback, which involved identifying errors and misconceptions not only in technical knowledge but also in the students' analytical approach, which demanded personalized attention. To alleviate this additional burden, a collaborative strategy was implemented wherein students engaged in group analysis, aiding one another in identifying errors and devising solutions.

B. Students' Perceptions of the SABERR Assessment

Near the end of the semester, the research team invited students to participate in a confidential survey. The survey aimed to capture students' understanding, views, and perceptions of the SABERR assessment approach. 24 students responded to the survey. Survey findings include the following:

- Ninety-five percent (95%) of respondents indicated that the SABERR approach has helped them detect and understand the sources of their mistakes to "a great extent" or "a very great extent."
- Sixty-six percent (66%) of respondents agreed that they encounter challenges while engaging with the SABERR rubric.
- Eighty percent (80%) of respondents either agreed or strongly agreed that after analyzing their own errors with the provided intermediate steps, answers, and rubric, they felt confident to correct their errors on a second attempt.
- Ninety-five percent (95%) of respondents either agreed or strongly agreed that the SABERR approach helped them to improve their understanding of the AI concepts covered in the course.
- Ninety-one percent (91%) of respondents indicated that having a productive attitude toward mistakes, which allows them to admit, learn from, and correct their mistakes, will be beneficial to their future career as computer scientists.

In sum, student survey results, along with our qualitative data, illustrate the effectiveness of this approach, with a majority of students reporting enhanced abilities to detect, understand, and correct their mistakes.

V. Conclusions and Implications

Previous research has pointed to learning from mistakes approaches supporting students to develop deeper levels of conceptual understanding, becoming active agents of their own learning in STEM-related fields, more specifically in CS software development courses. We extend prior research by underscoring the transformative potential of error-based learning in the context of AI education. Consistent with prior research, this study has also shown how students can develop metacognitive skills from their participation in approaches that frame errors as learning tools. We extend this finding by showing the impact of the SABERR assessment in an AI course students' learning and metacognitive skills development.

Our pilot study findings illuminate how the constructive metacognitive-based assessment, utilizing a second-attempt strategy, significantly supports students in enhancing their conceptual understanding and refining their problem-solving strategies. By embracing errors as learning opportunities, students are encouraged to engage deeply with their learning processes. Implications of this approach extend beyond mere students' academic achievements. Furthermore, the SABERR assessment fosters a culture where students are motivated to engage with challenging subjects, even when they perceive them as difficult. By framing problem-solving as a low-risk challenge, a learning from mistakes methodology encourages students to perceive mistakes as opportunities for growth and helps prevent academic cheating. While existing literature may not explicitly focus on utilizing mistakes to enhance metacognition, the SABERR assessment approach inherently fosters reflection and critical thinking, two key components known to promote metacognitive development.

Moreover, by framing and using errors as learning opportunities, AI instructors can cultivate educational environments where students feel safe to explore, question, position themselves as capable learners, and ultimately, innovate. Beyond highlighting the affordances and limitations of the SABERR approach, our assessment is a tool that AI educators can use to develop students' adaptive reasoning about the value of error-handling skills in AI education and workforce.

References

[1] R. W. Hollingworth and C. McLoughlin, “Developing science students’ metacognitive problem solving skills online,” *Australasian Journal of Educational Technology*, vol. 17, no. 1, 2001.

[2] S. A. Mathan and K. R. Koedinger, “Fostering the intelligent novice: Learning from errors with metacognitive tutoring,” in *Computers as Metacognitive Tools for Enhancing Learning*. Routledge, 2018, pp. 257–265.

[3] D. Ginat, “The greedy trap and learning from mistakes,” in *Proceedings of the 34th SIGCSE technical symposium on Computer science education*, 2003, pp. 11–15.

[4] D. Ginat and R. Shmalo, “Constructive use of errors in teaching cs1,” in *Proceeding of the 44th ACM technical symposium on Computer science education*, 2013, pp. 353–358.

[5] A. Katz and R. Shmallo, “Learning from errors as a pedagogic approach for reaching a higher conceptual level in database modeling,” in *Advanced Information Systems Engineering Workshops: CAiSE 2016 International Workshops, Ljubljana, Slovenia, June 13-17, 2016, Proceedings 28*. Springer, 2016, pp. 93–102.

[6] P. Zhang, J. White, and D. C. Schmidt, “Holicow: Automatically breaking team-based software projects to motivate student testing,” in *Proceedings of the 38th International Conference on Software Engineering Companion*, 2016, pp. 436–439.

[7] W. Lake, M. Wallin, G. Woolcott, W. Boyd, A. Foster, C. Markopoulos, and W. Boyd, “Applying an alternative mathematics pedagogy for students with weak mathematics: Meta-analysis of alternative pedagogies,” *International Journal of Mathematical Education in Science and Technology*, vol. 48, no. 2, pp. 215–228, 2017.

[8] G. Schraw and D. Moshman, “Metacognitive theories,” *Educational psychology review*, vol. 7, pp. 351–371, 1995.

[9] H. J. Hartman, *Metacognition in learning and instruction: Theory, research and practice*. Springer Science & Business Media, 2001, vol. 19.

[10] M. Alvidrez, N. Louie, and M. Tchoshanov, “From mistakes, we learn? mathematics teachers’ epistemological and positional framing of mistakes,” *Journal of Mathematics Teacher Education*, vol. 27, no. 1, pp. 111–136, 2024.

[11] A. P. Adiredja and N. Louie, “Untangling the web of deficit discourses in mathematics education,” *For the Learning of Mathematics*, vol. 40, no. 1, pp. 42–46, 2020.

[12] J. Metcalfe, J. Xu, M. Vuorre, R. Siegler, D. Wiliam, and R. A. Bjork, “Learning from errors versus explicit instruction in preparation for a test that counts,” *British Journal of Educational Psychology*, 2024.

[13] J. L. Booth, K. E. Lange, K. R. Koedinger, and K. J. Newton, “Using example problems to improve student learning in algebra: Differentiating between correct and incorrect examples,” *Learning and Instruction*, vol. 25, pp. 24–34, 2013.

[14] M. Kapur and K. Bielaczyc, “Designing for productive failure,” *Journal of the Learning Sciences*, vol. 21, no. 1, pp. 45–83, 2012.

[15] K. Loibl and T. Leuders, “How to make failure productive: Fostering learning from errors through elaboration prompts,” *Learning and Instruction*, vol. 62, pp. 1–10, 2019.

[16] D. Tsovaltzi, B. M. McLaren, E. Melis, and A.-K. Meyer, “Erroneous examples: effects on learning fractions in a web-based setting,” *International Journal of Technology Enhanced Learning*, vol. 4, no. 3-4, pp. 191–230, 2012.

[17] E. Yerushalmi and C. Polingher, “Guiding students to learn from mistakes,” *Physics Education*, vol. 41, no. 6, p. 532, 2006.

[18] Y. B. Kafai, D. DeLiema, D. A. Fields, G. Lewandowski, and C. Lewis, “Rethinking debugging as productive failure for cs education,” in *Proceedings of the 50th ACM technical symposium on computer science education*, 2019, pp. 169–170.

[19] Y. Kafai, G. Biswas, N. Hutchins, C. Snyder, K. Brennan, P. Haduong, K. DesPortes, M. Fong, V. J. Flood, O. W.-v. Aalst *et al.*, “Turning bugs into learning opportunities: understanding debugging processes, perspectives, and pedagogies,” 2020.

[20] R. Borasi, “Exploring mathematics through the analysis of errors,” *For the learning of Mathematics*, vol. 7, no. 3, pp. 2–8, 1987.

[21] E. Kazemi and D. Stipek, “Promoting conceptual thinking in four upper-elementary mathematics classrooms,” *Journal of education*, vol. 189, no. 1-2, pp. 123–137, 2009.

[22] J. Kilpatrick, “Problem formulating: Where do good problems come from?” *Cognitive science and mathematics education*, pp. 123–147, 1987.

[23] D. M. Adams, B. M. McLaren, K. Durkin, R. E. Mayer, B. Rittle-Johnson, S. Isotani, and M. Van Velsen, “Using erroneous examples to improve mathematics learning with a web-based tutoring system,” *Computers in Human Behavior*, vol. 36, pp. 401–411, 2014.

[24] E. Melis, “Erroneous examples as a source of learning in mathematics.” *CELDA*, vol. 2004, pp. 311–318, 2004.

[25] E. Melis, A. Sander, and D. Tsovaltzi, “How to support meta-cognitive skills for finding and correcting errors?” in *2010 AAAI Fall Symposium Series*, 2010.

[26] N. H. Ubaidullah, Z. Mohamed, J. Hamid, S. Sulaiman, and R. L. Yussof, “Improving novice students’ computational thinking skills by problem-solving and metacognitive techniques,” *International Journal of Learning, Teaching and Educational Research*, vol. 20, no. 6, pp. 88–108, 2021.

[27] L. Morales-Navarro, D. A. Fields, and Y. B. Kafai, “Growing mindsets: Debugging by design to promote students’ growth mindset practices in computer science class.” in *Proceedings of the 15th International Conference of the Learning Sciences-ICLS 2021*, 2021.

[28] C. Emeka, T. Bretl, G. Herman, M. West, and C. Zilles, “Students’ perceptions and behavior related to second-chance testing,” in *2021 IEEE Frontiers in Education Conference (FIE)*. IEEE, 2021, pp. 1–8.

[29] C. Emeka, G. Herman, and C. Zilles, “Leveraging second-chance testing to improve students’ outcomes,” in *Proceedings of the 2023 ACM Conference on International Computing Education Research-Volume 2*, 2023, pp. 50–51.

[30] J. H. Flavell, “Metacognition and cognitive monitoring: A new area of cognitive–developmental inquiry.” *American psychologist*, vol. 34, no. 10, p. 906, 1979.

[31] A. Shilo and B. Kramarski, “Mathematical-metacognitive discourse: How can it be developed among teachers and their students? empirical evidence from a videotaped lesson and two case studies,” *ZDM*, vol. 51, pp. 625–640, 2019.

[32] J. Lee, A. M. Kazerouni, C. Siu, and T. Migler, “Exploring the impact of cognitive awareness scaffolding for debugging in an introductory programming class,” in *Proceedings of the 54th ACM Technical Symposium on Computer Science Education V. 1*, 2023, pp. 1007–1013.

[33] C. Conati, “Commentary on:“toward computer-based support of metacognitive skills: A computational framework to coach self explanation”,” *International Journal of Artificial Intelligence in Education*, vol. 26, pp. 183–192, 2016.

[34] M. T. Cox, “Metareasoning, monitoring, and self-explanation,” 2011.

[35] M. Havenga, “The role of metacognitive skills in solving object-oriented programming problems: a case study,” *TD: The Journal for Transdisciplinary Research in Southern Africa*, vol. 11, no. 1, pp. 133–147, 2015.

[36] L. Margulieux, P. Denny, K. Cunningham, M. Deutsch, and B. R. Shapiro, “When wrong is right: The instructional power of multiple conceptions,” in *Proceedings of the 17th ACM Conference on International Computing Education Research*, 2021, pp. 184–197.

[37] J. W. Creswell, V. L. P. Clark, M. L. Gutmann, and W. E. Hanson, “Advanced mixed,” *Handbook of mixed methods in social & behavioral research*, vol. 209, 2003.

[38] J. W. Creswell and A. Tashakkori, “How do research manuscripts contribute to the literature on mixed methods?” pp. 115–120, 2008.

[39] J. Saldaña, “The coding manual for qualitative researchers,” 2021.