Shih, JL. et al. (Eds.) (2023). Proceedings of the 31st International Conference on Computers in Education.
Asia-Pacific Society for Computers in Education

# Investigating Cognitive Biases in Self-Explanation Behaviors during Game-based Learning about Mathematics

**J.M. Alexandra ANDRES[a*], Elizabeth B. CLOUDE[a], Ryan S. BAKER[a] & Seiyon LEE[a]**
[a]*Graduate School of Education, University of Pennsylvania, USA*
*aandres@upenn.edu

**Abstract:** The study investigates the impact of cognitive biases on middle-school students' affective experiences while learning about math in a game-based learning environment (GBLE). The study focused on students' confrustion, an affect construct that unifies the several manifestations of confusion and frustration. We studied confrustion in the context of students self-explaining erroneous examples, where they had to find and fix common errors in given math problems and self-explain their problem-solving processes either with or without scaffolding. Text replays were utilized to examine student interactions during game-based learning and identify behaviors that emerged in response to cognitive biases and affect and its impact on learning and performance outcomes. The results revealed that students who demonstrated more pseudo-confidence in their self-explanations had higher self-reported self-efficacy, but were more likely to submit incongruent responses, exhibit confrustion, make errors, and take longer to finish the game. Overall, the findings show that students were vulnerable to cognitive biases and did not always respond in ways that accurately reflected their approach to solving math problems. The insights into how students approach and learn from math games inform the design and implementation of GBLEs by addressing cognitive biases.

**Keywords:** game-based learning, confrustion, cognitive bias, confidence, self-explanation, self-efficacy

## 1. Introduction

Game elements designed to scaffold instructional support are essential for promoting behaviors that support mathematical learning (e.g., prompting students to generate self-explanations about their solutions; Ke et al., 2019; Wouters & van Oostendorp, 2013). A study by McLaren and others (2022) showed that students who generated open-ended and focused self-explanations (i.e., open-ended responses answering a focused prompt), compared to menu-based self-explanations (i.e., select an explanation from a menu), had significantly higher learning. However, in further studies, Johnson and Mayer (2010) found no differences in learning outcomes between open-ended and menu-based self-explanation types.

To better understand the impact of each type of prompt, it is important to explore which learning behaviors, processes, and strategies that emerge during open-ended and focused self-explanations contribute to actively constructing knowledge. One factor impacting successful math learning and achievement is the student’s confidence or self-efficacy (Pan et al., 2022). However, students’ judgments of their abilities are vulnerable to cognitive biases, possibly skewing their perceptions.

Cognitive biases stem from heuristic strategies, or decisional short cuts (Blumenthal-Barby & Krieger, 2015), to reduce the amount of information considered when making a decision (Schwenk, 1986). The emergence of biases is mediated by contextual factors such as attributed risk (Halpern, 1989), age and expertise (Forbes, 2005), and time constraints (Lehner et al., 1997), and these biases can impact a student’s confidence, interest, and achievement in math. For example, when students have high confidence but low knowledge (Pseudo-Confidence), they tend to experience a higher occurrence of confusion and frustration during learning (Di Leo et al., 2019).

Previous studies have also suggested that low-knowledge learners may attempt to alleviate confusion by avoiding activities that require deep thought (Rodrigo et al., 2007); for instance, gaming the system, where a student exploits a platform's help/feedback systems to succeed instead of attempting to learn (Baker et al., 2006), and Without Thinking Fastidiously (WTF) behaviors, where a student engages in activities that are unrelated to the learning goal (Wixon et al., 2012).

In this paper, we extend previous findings by examining pseudo-confident refusal, a cognitive bias where a student refuses to self-explain and instead asserts their correctness despite system feedback indicating otherwise. Specifically, the current study seeks to identify and examine the role of pseudo-confidence during self-explanations on students' affect, learning, and performance within a game-based learning environment (GBLE) called Decimal Point. Focused and open-ended prompts were used to elicit students' problem-solving processes in math and identify learning behaviors that emerged during knowledge construction. Through the use of qualitatively coded text logs of student interactions, the current study investigated the influence of pseudo-confidence, on students' affect, performance, self-efficacy, as well as interaction with the platform.

## 2. Methods

### *Decimal Point*

The platform used in this study was Decimal Point, a GBLE that is based on an amusement park metaphor. Players were guided through different sets of mini games built within the Cognitive Tutor Authoring Tools (CTAT; Aleven et al., 2009) to learn about decimals. The students played a total of 24 mini games during regular class periods (across an average of 3-5 days) which lasted around 45 to 55 minutes each.

In the current study, players used the erroneous examples play mode, where they were required to find and fix common errors in given math problems and then self-explain their problem-solving process either with or without scaffolding ("Think about how you came up with this answer, and drag the correct option to the following blank(s)." versus "How do you know?", respectively). The self-explanation prompts were designed to actively engage a student to develop a deeper understanding of critical concepts in mathematics. Each student was randomly assigned to one of three self-explanation conditions: menu-based, scaffolded, or focused format. In this paper, we study the focused, open-ended, self-explanations, which prompted students to explain their response with minimal guidance.

### *Sample and Student Level Metrics*

Data were collected from a middle school in Pennsylvania between March and November 2021. Three hundred and fifty-eight (*n* = 358; 44% females) students participated in the study; however, only 85 were retained for the current set of analyses due to data completeness across variables of interest. Students participated in a survey before and after they played Decimal Point. The survey measured Decimal Efficacy, which indicates a student's level of self-efficacy or their confidence in using decimal operations. The survey was adapted from literature (Pintrich et al. 1993) to better align with the context of the learning session, as used in Hou et al. (2020). Students were asked to rate statements on a 5-point Likert scale. We averaged three items to compute the Decimal Efficacy score: 1) I can do an excellent job on decimal number math assignments, 2) I can understand the most difficult material presented in decimal number lessons, and 3) I can master the skills being taught in decimal number lessons.

Students' interaction data were examined using text replays (Baker et al., 2006), which display interaction logs from student platform usage in an easy-to-read form. Text replays were then investigated to identify potential behaviors that emerged from cognitive biases and experiences of confrustion (when a student is either confused or frustrated; Mogessie et al., 2020). This method has been used in previous studies to label a variety of student variables (Baker et al. 2006; DiCerbo and Kidwai, 2013; Zhang et al., 2022) quickly and with high inter-rater reliability. For this study, analyses were focused on self-explanation behaviors that occurred within the problem-solving process. As such, the log data were divided into clips at the level of entire problems.

*Construct Operationalization and Qualitative Coding*

The team first analyzed student responses for indicators of self-explanation behaviors using qualitative categories. This approach followed the recursive, iterative process used in Weston et al. (2001), including reviewing literature on previous work using the Decimal Point platform (Forlizzi et al., 2014) and student engagement behaviors within GBLEs (Zhang et al., 2022). Using grounded theory (Charmaz, 1983), we identified common behaviors that were indicative of 1) struggles involving self-explanation from the current mathematical content and 2) were indicative of difficulty generating accurate responses for the given problem.

Following the process used in (Zhang et al., 2022), two coders (1st and 4th authors) coded a set of 900 clips together, identified four behaviors occurring during self-explanations (see details in Table 1), outlined the criteria for each indicator, and created a rubric. The coding manual was reviewed and discussed by the research team to ensure a shared understanding of the criteria and constructs being studied. This process was repeated until the whole team had a mutual understanding of the codebook's criteria and constructs.

Table 1. *Behaviors coded through Text Replays*

| Behavior | Definition |
|---|---|
| Confrustion<br>k = .68 | The learner explicitly mentions they do not know what to do or how to answer the problem and makes repeated answer submissions that are incorrect. |
| Without Thinking Fastidiously (WTF) k = .89 | The learner does not acknowledge the prompts for finding a solution and instead inputs text unrelated to the task at hand.<br>(e.g., “F*** off,” “Stop asking me,” “Ohmygahd” |
| Incongruent Response<br>k = .62 | The learner inputs correct information, but it is phrased differently from responses the system can recognize. These responses often involve minimal changes between response submissions where changes do not semantically change the answer; the student is trying to find a way to phrase the answer so that the system will accept it. |
| Pseudo-Confident Refusal<br>k = .79 | The learner refuses to explain their answers or mathematical process and instead refers to their inherent capacity to understand the mathematical problems by citing themselves as an authority to know the final answer.<br>(e.g., “I know because it’s the right answer,” “Because I’m smart”) |

## 3. Results

Spearman correlations with Benjamini-Hochberg post-hoc controls were computed to find associations between pseudo-confidence, performance metrics, and decimal efficacy survey responses. Students who demonstrated more pseudo-confidence in their self-explanations were more likely to submit incongruent responses ($r$ = 0.29, adj. α = 0.01, $p$ = 0.008) and exhibit confrustion ($r$ = 0.33, adj. α = 0.007, $p$ = 0.002). In addition, pseudo-confidence in self-explanations were moderately and positively associated with making errors ($r$ = 0.25, adj. α = 0.02, $p$ = 0.02) and taking longer to finish the game ($r$ = 0.37, adj. α = 0.003, $p$ = 0.004). Students who demonstrated more pseudo-confidence in their self-explanations tended to score higher on Post-test Decimal Efficacy surveys ($r$ = 0.25, adj. α = 0.025, $p$ = 0.023) and had increases in the normalized change in self-efficacy from before to after learning with the game ($r$ = 0.24, adj. α = 0.028, $p$ = 0.028). These results implyed that pseudo-confident responses may emerge in response to challenges responding to self-explanation prompts. Experiences of confrustion were related to increased errors and time spent problem solving, and in response, students may turn to pseudo-confident responses to avoid further engaging with the material. What is interesting to note is that pseudo-confident students were also more likely to positively estimate their decimal self-efficacy. This relationship may point to a tendency of pseudo-confident students to overestimate their abilities despite generating multiple errors and needing more time to solve math problems.

Table 2. *Correlations for behaviors, achievement metrics, and survey level responses*

| | Pseudo-Conf | Confru. | Incon. Resp. | WTF | Errors | Time | Dec. Eff. (pre) | Dec. Effi. (post) | Change in Dec. Eff. | Pre-test | Post-test | Del. Post | Change in Del. Post | Change in Learning |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Pseudo-Confident | 1 | 0.33 | 0.29 | | 0.25 | 0.37 | | 0.25 | 0.24 | | | | | |
| Confrustion | | 1 | | | 0.59 | 0.43 | 0.25 | 0.35 | | -0.53 | -0.47 | -0.54 | -0.25 | |
| Incongruent Response | | | 1 | | | | 0.06 | | | -0.26 | -0.26 | | | |
| WTF | | | | 1 | | | | | | | | | | |
| Errors | | | | | 1 | 0.67 | 0.31 | 0.28 | | -0.68 | -0.72 | -0.72 | | -0.29 |
| Time problem-solving | | | | | | 1 | 0.32 | 0.27 | | -0.55 | -0.52 | -0.49 | | |
| Decimal Efficiency (pre) | | | | | | | 1 | 0.47 | | -0.25 | -0.26 | -0.31 | | |
| Decimal Efficiency (post) | | | | | | | | 1 | 0.75 | -0.26 | | | | |
| Change in Decimal Efficacy | | | | | | | | | 1 | | | | | |
| Pre-test | | | | | | | | | | 1 | 0.83 | 0.76 | | |
| Post-test | | | | | | | | | | | 1 | 0.84 | 0.41 | 0.58 |
| Delayed Post-test | | | | | | | | | | | | 1 | 0.69 | 0.43 |
| Change in Delayed Post-test | | | | | | | | | | | | | 1 | 0.59 |
| Change in Learning | | | | | | | | | | | | | | 1 |

A linear regression predicting the rate of pseudo-confident responses using the interaction between time spent problem-solving and pre-test scores of achievement revealed that increased pre-test scores positively predicted pseudo-confidence ($\beta$ = .02, $p$ = .022), suggesting that students who were more familiar with the content were more likely to assert their perceived correctness and refuse to explain. A significant interaction was also observed between pre-test and time: students with higher pre-test scores who spent a longer amount of time problem solving also applied more pseudo-confident responses ($\beta$ = .28, $p$ = .018).

Table 3. *Results for linear regressions predicting achievement and Decimal Efficacy (N = 85, p < .05). Significant results shown in bold.*

| | Post-test | | Learning Gains | | Decimal Efficacy (post-test) | | Decimal Efficacy (Gains) | |
|---|---|---|---|---|---|---|---|---|
| | β | P | β | p | β | p | β | P |
| (Intercept) | **-0.07** | **<0.001** | **-0.03** | **0.001** | **0.01** | **<0.001** | -0.01 | 0.22 |
| Pseudo-confidence | **-0.13** | **0.02** | -0.05 | 0.43 | 0.15 | 0.56 | 0.09 | 0.95 |
| Time | **-0.48** | **<0.001** | **-0.25** | 0.03 | 0.20 | 0.11 | 0.08 | 0.65 |
| Interaction (Pseudo-confidence * Time) | **0.21** | **0.02** | 0.08 | 0.43 | -0.02 | 0.87 | 0.02 | 0.81 |
| $R^2$ | 0.29 | | 0.07 | | 0.08 | | 0.02 | |

Table 3 shows that across both achievement metrics (post-test and learning gains), a student who spent more time problem solving was related to decreases in their math performance. Students who gave more pseudo-confident responses were less likely to perform well on the post-test math assessment. A significant interaction was observed between time and the rate of pseudo-confident behaviors for predicting post-test. There is a complicated relationship here -- a simple slopes analysis revealed a significant relationship where pseudo-confident students who spent less time were less likely to have increases in their post-test scores. However, pseudo-confident students who spent more time did better.

## 4. Discussion

This study examined pseudo-confidence, a cognitive bias impacting math learning, and its relation to affect (e.g., confrustion) and other learning behaviors (e.g., WTF, incongruent responses) during open-ended and focused self-explanations within a GBLE. We found positive associations between pseudo-confident responses with time spent problem solving and errors made during game-based learning about math. The positive correlation between these variables indicated a potential struggle self-explaining problem solving, where students may defer to heuristics to complete the activity. Additionally, students may demonstrate pseudo-confident responses as a means of self-presentation and impression management around math difficulties (Schwenk, 1986).

We also found that higher pseudo-confidence was associated with more confrustion and incongruent responses during game-based learning. These relationships can be explained by the student failing to actively construct their knowledge when self-explaining, contributing to more cognitive incongruencies (i.e., confrustion) or more incongruent responses (Di Leo et al., 2020). If a student made multiple errors during problem solving, but continued to refuse to explain their problem-solving approach (pseudo-confident), this may result in more confrustion in the long term. The results also showed that the more pseudo-confident a student was, the higher their self-reported Decimal Efficacy was after game-based learning, indicating that responding in this fashion may increase student efficacy despite lower knowledge.

Next, we examined whether pre-test scores and time spent in the game impacted the rate of pseudo-confidence in self-explanations. The results showed that students with higher pre-test scores also had higher rates of pseudo-confidence in their self-explanations. This finding suggested that students who had more prior knowledge of math topics demonstrated more pseudo-confidence. Students with higher prior knowledge may be vulnerable to the Dunning-Kruger effect; they may be knowledgeable enough to be confident in their responses, but unable to detect blind spots in their knowledge.

Lastly, linear regression analyses revealed that the more time spent in the game and pseudo-confident responses made, the worse students performed on the post-test assessment. A simple slopes analysis revealed that pseudo-confident students who spend less time problem solving performed worse on the post-test assessment. These findings may align with previous research which suggest that students tend to make less effective decisions when required to solve problems or make decisions with less time (Lehner et al., 1997). However, pseudo-confident students who spent more time performed better on the post-test, possibly suggesting that this subset of students was still learning.

This study used a mixed-methods approach to investigate pseudo-confidence while students explained their approach to solving math problems during game-based learning. While self-report data outside the context of the activity was valuable for measuring global confidence in math, this approach presents challenges for studying the role of confidence during math learning. Confidence is a dynamic construct, one that changes during a learning session based on a student's goals and in response to various types of feedback they may receive (e.g., time left or confrustion; Pavlas, 2010). Thus, real-time measurement of confidence is important for advancing our understanding on if, when, and how confidence evolves during game-based learning about math. Future research should evaluate the role of other common cognitive biases that impact math learning, including examining the presence of stereotype threat or the hard-easy effect (overconfidence on harder problems) during math learning. For example, to what extent does the rate of pseudo-confidence change depending on how difficult the math problems are? This research has implications for identifying key learning behaviors and building adaptive self-explanation prompts and interventions that maximize students actively constructing knowledge about math with GBLEs.

## References

Aleven, V., Mclaren, B. M., Sewall, J., & Koedinger, K. R. (2009). A new paradigm for intelligent tutoring systems: Example-tracing tutors. *Int'l Journal of AI in Education, 19*(2), 105-154.

Baker, R.S.J.d., Corbett, A.T., Wagner, A.Z. (2006) Human Classification of Low-Fidelity Replays of Student Actions. *Proc. of the Educational Data Mining Workshop at ITS2006,* 29-36.

Baker, R.S.J.d., D'Mello, S.K., Rodrigo, M.M.T., Graesser, A.C. (2010) Better to be frustrated than bored: The incidence, persistence, and impact of learners' cognitive-affective states during interactions with three different computer-based learning environments. *Int'l Journal of Human-Computer Studies*, 68 (4), 223-241.

Blumenthal-Barby, J. S., & Krieger, H. (2015). Cognitive biases and heuristics in medical decision making: A crit. rev. using a systematic search strategy. *Med. Decision Making, 35*(4), 539-557.

Charmaz, K. (1983). The grounded theory method: An explication and interpretation. *Contemporary Field Research*, R. Emerson, Eds. Boston, MA: Little, Brown and Company. 109–126.

Di Leo, I., Muis, K. R., Singh, C. A., & Psaradellis, C. (2019). Curiosity… confusion? Frustration! The role and sequencing of emotions during mathematics problem solving. *Contemporary Educational Psychology*, *58*, 121-137.

Dunning, D. (2011). The Dunning–Kruger effect: On being ignorant of one's own ignorance. In *Advances in Experimental Social Psychology* (Vol. 44, pp. 247-296). Academic Press.

Forbes, D. P. (2005). Are some entrepreneurs more overconfident than others?. *Journal of Business Venturing, 20*(5), 623-640.

Forlizzi, J., McLaren, B. M., Ganoe, C., McLaren, P. B., Kihumba, G., & Lister, K. (2014,). Decimal point: Designing and developing an educational game to teach decimals to middle school students. In *8th Euro. Conf. on Games-Based Learning:* (pp. 128-135).

Halpern, J. J. (1989). Cognitive factors influencing decision making in a highly reliable organization. *Industrial Crisis Quarterly, 3*(2), 143-158.

Hou, X., Nguyen, H. A., Richey, J. E., & McLaren, B. M. (2020). Exploring how gender and enjoyment impact learning in a digital learning game. *Proc. Artificial Intelligence and Education,* 255-268.

Kahneman, D., & Tversky, A. (2000). *Choices, Values, and Frames*. NY; Cambridge

Karumbaiah, S., Baker, R.S., Tao, Y., Liu, Z. (2022). How does students' affect in virtual learning relate to their outcomes? A systematic rev. challenging the positive-negative dichotomy. In *Proc. of the 12th International Conference on Learning Analytics & Knowledge*, 24-33.

Ke, F., M. Clark, K., & Uysal, S. (2019). Architecture game-based mathematical learning by making. *International Journal of Science and Mathematics Education*, 17, 167-184.

Lehner, P., Seyed-Solorforough, M. M., O'Connor, M. F., Sak, S., & Mullin, T. (1997). Cognitive biases and time stress in team decision making. *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, *27*(5), 698-703.

McLaren, B. M., Nguyen, H. A., Richey, J. E., & Mogessie, M. (2022). Focused self-explanations lead to the best learning outcomes in a digital learning game. In *Proc. of the 16th International Conference on the Learning Sciences* (pp. 1229-1232).

McNamara, D. S., Jackson, G. T., & Graesser, A. (2010). Intelligent tutoring and games (ITaG). In *Gaming for Classroom-Based Learning: Digital Role Playing as A Motivator of Study*, 44-65.

Mogessie, M., Elizabeth R., J., McLaren, B. M., Andres-Bray, J. M. L., & Baker, R. S. (2020). Confrustion and gaming while learning with erroneous examples in a decimals game. In *Proceedings of the International Conference on Artificial Intelligence in Education,* 208-213.

Pan, Y., Ke, F., & Xu, X. (2022). A systematic review of the role of learning games in fostering mathematics education in K-12 settings. *Educational Research Review*, 100448.

Pekrun, R., & Stephens, E. J. (2012). Academic emotions, *APA Educ. Psych. Handbook: Individual Differences & Cultural and Contextual Factors*, Vol. 2, 3–31.

Richey, J. E., McLaren, B. M., Andres-Bray, M., Mogessie, M., Scruggs, R., Baker, R., & Star, J. (2019). Confrustion in learning from erroneous examples: does type of prompted self-explanation make a difference?. In *Proc. of the 20th Int'l Conf. on Artificial Intelligence in Education*, pp. 445-457.

Rodrigo, M.M.T., Baker, R.S.J.d., Lagud, M.C.V., Lim, S.A.L., Macapanpan, A.F., Pascua, S.A.M.S., Santillano, J.Q., …& Viehland, N.J.B. (2007). Affect and usage choices in simulation problem solving environments, Proc. of the *13th Int'l Conf. on AI in Education*, pp. 145–152.

Sanchez, C., & Dunning, D. (2018). Overconfidence among beginners: Is a little learning a dangerous thing?. *Journal of Personality and Social Psychology, 114*(1), 10.

Schwenk, C. H. (1986). Information, cognitive biases, and commitment to a course of action. *Academy of Management Review, 11*(2), 298-310.

Weston, C., Gandell, T., Beauchamp, J., McAlpine, L., Wiseman, C., & Beauchamp, C. 2001. Analyzing interview data: the development and evolution of a coding system. *Qual. Soc., 24*(3), 381–400.

Zhang, J., Andres, J. M. A. L., Hutt, S., Baker, R. S., Ocumpaugh, J., Nasiar, N., ... & Young, T. (2022). Using machine learning to detect SMART model cognitive operations in mathematical problem-solving process. *Journal of Educational Data Mining*, *14*(3), 76-108.