Jiang, B. et al. (Eds.) (2025). Proceedings of the 33rd International Conference on Computers in Education. Asia-Pacific Society for Computers in Education

SRL profiles in math problem-solving: The Essential Role of Monitoring and Translating for Outcomes and Beliefs

Nidhi NASIAR[a*], Jaclyn OCUMPAUGH[b] , Ryan S. BAKER[c], Caitlin MILLS[d], Samuel RHODES[e], Sheela SETHURAMAN[e]
[a]*University of Pennsylvania, USA*
[b]*University of Houston, USA*
[c]*Adelaide University, Australia*
[d]*University of Minnesota, USA*
[e]*CueThink Inc.,USA*
*nasiar@upenn.edu

Abstract: Self-regulation positively impacts learning, which has prompted efforts to detect and foster SRL strategies. We used Gaussian Mixture Modeling to cluster students on five SRL strategies in problem solving (PS)—measured via SRL detectors grounded in Winne's SMART model (2017). Seven distinct SRL profiles emerged from the data, which were examined for their relation to math beliefs, anxiety, and PS measures. Findings show that SRL profiles with high proficiency across multiple SRL skills achieved higher accuracy, spent more time on pre-tests, and reported more opportunities to share their math thinking. These students also frequently engaged in monitoring and translation. Notably, profiles with strong assembling skills underperformed peers who balanced SRL strategies with relatively higher monitoring and translating. Overall, these results highlight the dynamic relationship between SRL and math beliefs in PS, and suggest emerging profiles to design tailored interventions.

Keywords: Self-regulated learning, clustering, mindset, math beliefs

1. Introduction

Self-regulation is important for effective learning and academic achievement. Engaging in self-regulated learning (SRL) improves outcomes; yet, fostering SRL skills in students remains challenging and requires targeted, personalized support (Schunk, 2001). Recent work on automated SRL detection has leveraged trace data to capture fine-grained behaviors (Azevedo et al., 2012), with an increasing shift in focus from measuring to fostering SRL skills (Bouchet et al., 2013; Ting-Chia, 2024). Automated detectors can track individual SRL behaviors in real time and inform adaptive scaffolds. However, designing separate scaffolds for each SRL skill detector is resource-intensive and may overlook how strategies naturally co-occur in problem-solving (PS). Instead, grouping learners based on similar SRL behavior profiles offers a scalable approach to personalized support.

One way to identify learners with common SRL behaviors, as explored in multiple projects, is through clustering techniques (Gasevic et al., 2017), which has provided insights that advanced both theoretical and empirical understanding of SRL. A growing body of work employs various clustering methods to form groups for targeted support (Bouchet et al., 2013; Ting-Chia, 2024). These clustering approaches have shown connections to learner engagement and achievement, with profiles characterized by SRL difficulties correlating with poorer outcomes (Farhana et al., 2021). Grounding clustering analyses in established SRL theories enhances the effectiveness and relevance of the research and informs scaffold design. Winne's SMART model (2017) with cognitive behaviors of searching, monitoring, assembling, rehearsing, and translating, is widely applied in empirical SRL studies (Azevedo et al., 2012; Farhana et al., 2021). Our study uniquely integrates a suite of automated SRL detectors based on Winne's (2017) model to capture students' SRL behaviors, and then clusters detectors' outputs to identify distinct learner subgroups for targeted support.

The relationship between math beliefs, anxiety, and SRL is complex and cyclic (Maloney & Beilock, 2012). Effective PS involves more than cognitive strategies; it is heavily influenced by learners' beliefs and affect (Pekrun et al., 2017). Students' epistemic beliefs, such as viewing math merely as rule memorization, shape their SRL behaviors, where positive beliefs improve persistence and strategizing while negative ones reduce them (Dweck, 2014). Math anxiety further disrupts strategy selection and progress monitoring (Maloney & Beilock, 2012). However, limited work has examined combinations of SRL skills necessary for effective PS and math beliefs. This study seeks to address that gap by investigating how SRL behaviors co-occur in math PS and their interaction with students' math mindset and beliefs within CueThink, an online math platform. The study is driven by two main questions: i) What distinct SRL profiles emerge in math PS (based on distinct SRL strategies' usage)?, ii) How are these SRL profiles associated with measures of PS, math beliefs, and anxiety?

2. Methods

2.1. Learning Platform: CueThink

This study is based on CueThink, an online platform that scaffolds math PS and promotes regulation. In CueThink, students complete "Thinklets" math problems that follow 4 phases based on Winne (2017): in the Understand phase, they identify what's given and asked, and note observations; in Plan, they select strategies and outline steps for solving; in Solve, they record screencast explanations using mathematical tools; and in Review, they refine their solutions via a guided checklist. Students move freely between phases as needed. Data was collected from 230 7^{th}-8^{th} grade students at a suburban Eastern U.S. middle school in 2022–2023 for 768 problems. Pre- and post-tests were administered at the start and end of the year.

2.2. SRL Detectors used for Clustering

SRL strategies were detected from Zhang et al.'s (2022a) detectors developed using qualitatively coded SRL behaviors from trace data, aligned with Winne's SMART model (2017). They were validated for generalizability using 10-fold student-level cross-validation, using Area Under the ROC Curve (AUC) as primary metric (chance AUC = 0.5; perfect = 1.0), and checked for algorithmic bias (see Zhang et al., 2022). The SRL detectors were developed for i) numerical representation (NR; AUC = 0.894), showing learners' understanding of numerical components, and ii) contextual representation (CR; AUC = 0.813) for noting contextual details-- both classified as Assembling. iii) Data transformation (DT; AUC =0.815), as students manipulate data to find a solution (Translating). iv) Following plans (FP; AUC = 0.808) as correctly including chosen strategies into the plan (Monitoring), and v) incorporating information (II; AUC = 0.803) as correctly including previously assembled information into the plan (Rehearsing). These SRL detectors identified students' SRL processes in the Understand and Plan phase within each problem they solved, using their interaction data. The usage of the 5 SRL strategies was calculated for every student across all their solved problems based on the probability outputs of the 5 SRL detectors, which was then used as input to clustering.

2.3. Research Instruments

Three PS measures were used: i) *Accuracy Score*, with one PS item given per trimester, with items from a standard curriculum. ii) *Normalized learning gains* (LG), calculated as (post-pre)/(1-pre) when post>pre and (post-pre)/pre when post<=pre. iii) *Duration of Pre & Post test*, seconds spent from the start to the end time on test. We measured *math beliefs* using three SRL subscales (0–100) from the abbreviated Indiana Math Beliefs Scale (Cipora et al., 2015): Speed Orientation (4-items), the degree to which a student values thinking deeply (100) over the speed of solving math problems (0); Nature of Math (7-item), whether students see math from a constructionist lens (100) or as rigid and non-creative (0); Math Share, a single-item asking about the opportunities that students have to talk about math. *Math anxiety* was assessed with the 9-item Modified Abbreviated Math Anxiety Scale (Carey et al., 2017), slightly modified for American English, scored 0–100, with two factors—Learning and Evaluation.

2.4. Gaussian Mixture Modeling (GMM) clustering

We used Gaussian Mixture Modeling (GMM) clustering, with oval-shaped covariance and an outlier threshold of 0.7, to cluster the proportion of strategy usage based on each student's 5 SRL detector outputs. The function returned probabilities of a data point belonging to each cluster. We evaluated different N of clusters using the Bayesian Information Criterion (BIC), Akaike Information Criterion (AIC), and Silhouette score. Overall, 7 clusters performed best.

Differences in PS and survey measures across SRL clusters were tested using the Kruskal–Wallis (KW) H test, a nonparametric test for non-normal continuous data. A significant result indicated that at least one cluster differed. For those measures, pairwise Mann–Whitney U tests identified the cluster pairs that differed, limiting comparisons to reduce Type I error. Benjamini–Hochberg correction was applied to control the false discovery rate and mean ranks were used to determine which cluster exhibited higher/lower values for the measures.

3. Results

3.1. RQ1: Examining SRL Profiles from GMM Clustering

Students were clustered by their SRL strategy use; Table 1 lists each cluster's median of variables (NR, CR, DT, DP, II). Each median was classified relative to its overall detector distribution percentiles (Table 2): low(<25th), moderately low(25th to <50th), intermediate(50th), moderately high(>50th to <75th), and high(>=75th). Cluster names reflect these labels—"advanced" for high or moderately high levels, "intermediate" for median levels, and "beginner" for moderately low levels—omitting skills at low levels except in clusters with uniformly minimal use. E.g., A's NR median(0.87) is between 50th(0.77) and 75th(0.95) percentiles, and is labeled moderately high. Clusters are ordered from the most advanced to emerging SRL strategists.

Table 1. *Cluster Medians for 5 SRL strategies, shaded as per detector distribution of each strategy (see Table 2), with darker shading indicating higher values*

		Assemble		Translate	Monitor	Rehearse
	N (%)	NR	CR	DT	FP	II
Cluster A	82 (36.1)	0.87	0.85	0.84	0.99	1
Cluster B	45 (19.8)	0.73	0.69	0.75	0.95	0.99
Cluster C	52 (22.9)	0.77	0.61	0.64	0.78	0.92
Cluster D	9 (4.0)	0.92	0.91	0.2	0.44	0.93
Cluster E	6 (2.6)	0.93	0.2	0.31	0.52	0.85
Cluster F	27 (11.9)	0.63	0.57	0.37	0.54	0.91
Cluster G	6 (2.6)	0.22	0.02	0.06	0.7	0.11

Table 2. *Detector output distribution*

	Assemble		Translate	Monitor	Rehearse
Level	**NR**	**CR**	**DT**	**FP**	**II**
High.	x ≥ 0.95	x ≥ 0.87	x ≥ 0.85	x ≥ 0.99	x=1
Mod high	0.77 ≤ x <0.95	.69 ≤ x<0.87	0.7 ≤ x<0.85	0.93 ≤ x <0.99	0.99 ≤ x <1
Mod low	0.63 ≤ x <0.77	0.51 ≤x <0.69	0.47 ≤ x <0.7	0.74 ≤ x <0.93	0.9 ≤ x < 0.99
low	x < 0.63	x < 0.51	x < 0.47	x < 0.74	x < 0.9

Cluster A (Advanced SRL strategists) is characterized by high levels of CR, FP, and II, and moderately high usage of NR, and DT. Learners have high usage of SRL, including rehearsing (using information from their plan), monitoring (applying selected strategies), and assembling contextual information. While overall SRL engagement is high, some variability was found in NR, CR, and DT. All students consistently show high levels of FP and II.
Cluster B (Advanced Monitors and Translators with intermediate contextual assembling and rehearsing) is characterized by moderately high DT and FP, medium levels of CR and II, and moderately low NR. These students engage in data manipulation (translation) and follow their plan. Strengthening their assembling skills for NR could create a more balanced SRL profile for these students. Low variability is observed across students.
Cluster C (Intermediate Numerical Assemblers, beginner for other strategies) is characterized by medium values of NR with moderately low CR, DT, FP, and II. These students engage in average amounts of NR math word problems (assembling), but less monitoring, rehearsing, and translating. Students had low variability across strategies.
Cluster D (Advanced Assemblers with beginner rehearsing) is characterized by high CR, moderately high NR, moderately low II, and low DT and FP. These students frequently engage in problem representation and assembling, and sometimes use rehearsing strategies.
Cluster E (Advanced Numerical Assemblers only) is characterized by moderately high NR but low levels of all other strategies. Students in this group often assemble numerical information from math problems, but they do not independently show contextual understanding, DT, FP, and II. Low to moderate variability is observed across SRL strategies, other than for DT.
Cluster F (Beginner Numerical and Contextual Assemblers) is characterized by moderately low NR and CR, and low DT, FP, and II. Low variability is seen across strategies, except II.
Cluster G (Emerging SRL Strategists) is characterized by the lowest usage of all SRL strategies. This group represents students who are failing to use these SRL strategies in math PS. Low variability is observed across all strategies except NR.

Overall, Clusters A and B show well-balanced and comprehensive SRL profiles, with a range of key processes of assembling, monitoring, rehearsing, and translating. Clusters C, D, and E exhibit a mixed pattern of strategy utilization, with learners exhibiting strengths in some strategies but struggling to use others (or not realizing they are useful). Clusters F and G show minimal usage of SRL strategies, and are likely require support to develop SRL skills.

3.2. RQ2: Association of PS, beliefs, and anxiety measures to SRL profiles

The KW H statistic (Table 3) is used to test differences in the survey measures among clusters.

Table 3. *Results from Kruskal-Wallis (KW) H test for measures. Significant values are shaded*

Type	Measures	Pretest		Posttest	
		H	p	H	p
Problem Solving (PS)	Duration (Dur)	20.91	0.002	7.38	0.288
	Accuracy (Acc)	9.21	0.162	15.35	0.010
	Norm. Learning Gains (LG)	NA	NA	6.48	0.372
Anxiety	mAMAS Anxiety Learning (AxL) mAMAS	5.80	0.446	4.07	0.67
	Anxiety Evaluation (AxEV)	1.41	0.965	2.18	0.90
Beliefs	IMBS Speed Orientation (Speed)	3.04	0.804	18.79	0.005
	IMBS Nature of Math (Nature)	7.04	0.317	12.90	0.045
	IMBS Math Share	6.23	0.398	13.35	0.038

The PS measure of duration of the pre-test was significantly different (p=0.002) across at least one cluster pair. By the post-test, differences in duration between clusters were not observed. There was no difference among clusters in accuracy in the pre-test, however, there was a significant difference between clusters for the post-test (p=0.010).

No cluster differences were observed for learning gains or either form of anxiety. However, all three math belief constructs -- measure of speed orientation, nature of math, and math share -- had significant post-test differences among clusters (p=0.005, 0.045, 0.038) but not the pre-test (p=0.804, 0.317, 0.398), meaning the SRL usage clusters had different beliefs on what math was about, growth mindset in math, sharing math ideas, and on importance of thinking deeply over quickly on a math problem. Differences at post but not pre suggest that these beliefs may have been influenced by students' experience of using SRL strategies to solve problems in CueThink for a year, although a between-conditions comparison would be needed to be certain. Next, pairwise Mann-Whitney U tests were conducted for significant KW test measures (i.e. pre-test duration, and post-test measures of accuracy, speed, and nature).

Table 4. *Pairwise Mann-Whitney U tests only with p <0.05; significant values after B&H correction are shaded.*

	PS (Pre)		**PS (Post)**		**Beliefs (Post)**					
	Dur		**Acc***		**Share**		**Speed**		**Nature**	
	U	p	U	p	U	p	U	p	U	p
A vs. B	-	-	-	-	1042.5	0.03	983.5	0.01	-	-
A vs. C	2555.5	0.002	1354.5	0.036	-	-	-	-	1073	0.01
A vs. D	445	0.04	-	-	-	-	397.5	0.02	-	-
A vs. E	370	0.015	116	0.049	-	-	101	0.03	-	-
A vs. F	1391	0.001	583.5	0.027	-	-	-	-	-	-
B vs. D	-	-	213	0.012	-	-	231.5	0.002	210	0.017
B vs. F	689.5	0.03	-	-	685.5	0.003	625	0.036	-	-
C vs. D	-	-	228	0.016	-	-	240.5	0.008	248	0.004
C vs. F	-	-	-	-	703.5	0.021	-	-	-	-
D vs. E	-	-	4	0.018	-	-	0	0.001	-	-
D vs. F	-	-	111	0.02	-	-	-	-	-	-
E vs. F	-	-	-	-	-	-	34	0.04	-	-

Pre Duration. Significant differences were found in the time taken to finish pre-test between clusters A and C, and A and F (see Table 4). Students in cluster A spent more time on the pre-test than those in C and F, suggesting that advanced SRL strategists invested more time during initial assessment. In contrast, students with lower and intermediate SRL strategy usage spent less time, possibly indicating less engagement or cognitive effort . These effects were not found for the duration of the post-test, however, perhaps due to familiarity with the test, or due to practice with similar math problems in CueThink for a year.

Accuracy at Post. There were significant post-test accuracy differences between clusters B-D, C-D, D-E, and D-F. Students in C, B, and F achieved significantly higher post-test accuracy than in D, indicating that students who balance assembling with other SRL strategies like translation and monitoring (Clusters C and B), or even beginners (Cluster F), have higher accuracy than those focusing solely on assembling skills (Cluster D). Conversely, Cluster D had higher post-test accuracy than E, suggesting that contextual understanding of the problem may support better performance.

Speed Orientation Post. For this measure, a higher value indicated greater preference for deep thinking over speed wile solving problems. Significant differences were found in the post-test between Clusters C-D, B-D, and E-D. Based on their mean ranks (B=23, C=26, D=4.5, E=3.5), students in C and B valued deep thinking over speed more than D, suggesting that C and B cluster students prioritize thoughtful PS while engaging deeply with concepts. In turn, D valued deep thinking more than E. Cluster E students, frequently engaged in NR but had low usage of other SRL strategies and preferred solved problems more quickly.

Math Share Post. Significant differences were found in the post-test Math Share scores between Cluster B and Cluster F (p=0.03). Students in B reported sharing their mathematical thinking in class more than Cluster F, based on their mean ranks (B=23, F=14). This suggests that students with lower levels of SRL strategy usage may perceive fewer opportunities to share their mathematical ideas in class.

Nature of Math Post. Significant differences were observed in post-test perceptions of the nature of mathematics between Clusters C and D (p=0.004), and Clusters B and D (p=0.017). Clusters C and B hold more constructivist views of mathematics (highlighting creative, exploratory, and interconnected nature of math) than Cluster D.

4. Discussion and Conclusion

This study identified 7 student SRL profiles, revealing differences across the use of 5 SRL strategies, math problem solving (PS), and beliefs. Key findings show that students in clusters with a higher proficiency and balance of SRL skills (cluster A and B) have better PS accuracy, spent more time on the pre-tests and reported sharing of their math thinking than clusters with lower usage of fewer SRL skills (cluster E and F). Results further show that these clusters A and B integrate considerable usage of monitoring and translating strategies, processes critical for learning (Winne et al., 2017), whereas other students (in cluster E and F) primarily engaged in NR and sometimes in CR (assembling) but not as much in other SRL strategies. Notably, even students with strong assembling (cluster D) underperformed peers who balanced SRL strategies with a relatively higher monitoring and translating (B, C, and F), illustrating the importance of a well-rounded and diverse SRL profile in general, and specifically these two SRL strategies, for math problem solving. Post-test differences in beliefs (e.g., cluster D showed a lower preference for deep thinking over speed in solving math than clusters B and C, but higher than cluster E; and cluster D exhibited more fixed-mindset, maladaptive math beliefs than B and C, who showed a growth mindset) further highlight the nuances of how SRL experiences shape math beliefs over time.

A limitation of the study is seen in the small sample sizes for Clusters D, E, and G, suggesting these results should be replicated in the future with larger datasets across more types of content and grade levels. Also, the absence of significant links between SRL strategies and anxiety contrasts with prior work (Pekrun et al., 2017), and merits further investigation. Longitudinal analyses could also clarify whether cluster membership shifts as students gain experience. Correlating to additional motivational and affective variables (e.g., self-efficacy) could enrich interpretation. Overall, this study highlights the dynamic interplay between SRL and evolving math beliefs, with potential to inform future studies for personalized learner support based on their SRL gaps and math beliefs. In addition, the ML-detector–based clustering pipeline can be used to automate support for students with distinctive SRL profiles.

References

Azevedo, R., Behnagh, R.F., Duffy, M., Harley, J.M., Trevors, G., 2012. Metacognition and self-regulated learning in student-centered learning environments, in: Theoretical foundations of learning environments. Routledge, pp. 171–197.

Bouchet, F., et al., 2013. Clustering and profiling students according to their interactions with an intelligent tutoring system fostering self-regulated learning. J. Educational Data Mining 5, 104–146.

Carey, E., Hill, F., Devine, A., Szűcs, D., 2017. The modified abbreviated math anxiety scale: A valid and reliable instrument for use with children. Frontiers in Psychology 8, 11.

Cipora, K., Szczygieł, M., Willmes, K., Nuerk, H.C., 2015. Math anxiety assessment with the abbreviated math anxiety scale: Applicability and usefulness. Frontiers in Psychology 6, 1833.

Dweck, C., 2014. Mindsets and math/science achievement .

Farhana, E., Potter, A., Rutherford, T., Lynch, C.F., 2021. Feedback and self-regulated learning in science reading. *International Conference on EDM, 820-826.*

Gasevic, D., Jovanovic, J., Pardo, A., Dawson, S., 2017. Detecting learning strategies with analytics: Links with self-reported measures and academic performance. J. Learning Analytics,113-128.

Maloney, E.A., Beilock, S.L., 2012. Math anxiety: Who has it, why it develops, and how to guard against it. Trends in cognitive sciences 16, 404–406.

Pekrun, R., Lichtenfeld, S., Marsh, H.W., Murayama, K., Goetz, T., 2017. Achievement emotions and academic performance: Longitudinal models of reciprocal effects. Child development 88, 1653–1670.

Schunk, D.H. 2001. Self-regulation through goal setting. ERIC Document Service, 462671.

Ting-Chia, H. S. U. (2024). Effects of the Self-Regulated Learning and Motivation on Learning Achievements of the Programming Courses. In *International Conference on Computers in Education*.

Winne, P.H., 2017. Learning analytics for self-regulated learning. Handbook of LA, 754, 241–249.

Zhang, J., Andres, J.M.A.L., et al., 2022a. Using machine learning to detect smart model cognitive operations in mathematical problem-solving process. J. Educational Data Mining 14, 76–108.