Metacognition, Formative Assessment, and Student Perspective: Learning About Metacognition Through In-Class Comparison of Response Systems

Melanie Brady, EdD; Christopher P. Forest, MSHS, PA-C

Purpose Educators influence learning context through, among other things, triadic reciprocal causation—a behavioral model used in social cognitive theory. Eliciting responses from learners during instruction may contribute to learning by engaging their metacognition. In this study, metacognition was examined based on learners' experience with 2 different polling methods: student response systems (SRS), commonly known as "clickers," and a low-technology response system (LTRS), where students raise labeled signs, each with a different response. Scale item results are reported.

Methods Scales measuring the influence of metacognition from polling methods were administered, following treatment (SRS) and comparison method (LTRS), to first-year physician assistant (PA) students (n = 54). Statistical tests of significance and effect size for each item were compared.

Results Performance outcomes suggest that the SRS enhances learning experiences more than low-technology polling. Self-reported surveys indicate that self-monitoring, note-taking, and understanding during lectures are significantly improved with the use of clickers. Peer and social comparison items did not demonstrate significance collectively, but survey results indicate that candidates compared themselves to their peers significantly more with the LTRS than with the SRS.

Conclusion Findings support the practice of using an electronic SRS to poll PA students and enhance learning.

Feature Editor's Note:

Physician assistant (PA) education relies on a broad range of techniques and methods of varying sophistication and practical utility. One technique, student response systems (SRS), is now commonly used as a stimulus to increase student engagement and a means to improve student performance. The technique's impact on higher-order, multifactorial, educational constructs has been less thoroughly investigated. Brady and Forest examine metacognitive self-regulation, comparing the SRS approach with a less technically sophisticated, but similar and familiar, training method. They embed their discussion of metacognitive self-regulation in social cognitive theory and the personality construct of self-monitoring. This research advances understanding of the full impact of the SRS technology and its incremental improvement over a viable training alternative, along with presenting evidence for the effectiveness of SRS in supporting higher-order learning.

Doug Brock, PhD

INTRODUCTION

Student engagement is an important element in improving learning outcomes in physician assistant (PA) programs.[1] Educators seek instructional strategies and technological tools to actively engage students.[2] It seems clear that use of student response systems (SRS) is associated with some level of learning benefits, based on a recent meta-analytic study.[3] Empirical research on SRS indicates that they provide the closest possible experience to one-on-one tutoring in the classroom setting.[4]

Social cognitive theory (SCT) places the responsibility of learning on the individual for event-related changes in learning outcomes because individual cognition and belief systems are "central determinants"—elements in the control of the individual learner that drive the outcomes in the learning process.[5] Social cognitive theory also introduces the concept of "triadic reciprocality" to contextualize the interplay among the individual (learner, instructor), behavior (instruction, knowledge acquisition), and environment (instructor, learner, peers, social setting, instructional methodology).[5] Instructors select the instructional design and strategies used. Following through with this idea, according to SCT, instructors can be described as the actors who exert the primary influence on identifying concepts that lack clarity to structure lecture content around these concepts. As the primary influence, instructors can create priming-retrieval cues to access prior knowledge, stimulating pathways that improve metacognition.

Graduate health science courses increasingly use SRS in contexts ranging from the classroom to real-time polls during medical internships. This study compares the use of a high-technology

The authors declare no conflict of interest.

J Physician Assist Educ 2018;29(2):104–108

DOI 10.1097/JPA.0000000000000203

SRS that uses clicker technology with that of an easily accessible, low-technology response system (LTRS) that uses labeled signs, roughly equivalent to students raising hands. The SRS is an anonymous feedback system while the LTRS is similar to flashcards. When the SRS is used, students select a response on a handheld device that uses a radiofrequency signal to transmit and record their response. Students' responses can be displayed by the instructor in the form of a histogram on a PowerPoint slide if the instructor chooses. Low-technology response systems are handheld, low-cost alternatives in which responses are public, meaning that responses can be seen by peers before the correct answer response is known. When a question is asked, the learner raises the desired response card, and the instructor takes a visual inventory of student responses. A low-technology response system is similar enough in nature to SRS to provide a valid basis for comparison and provides feedback on student learning commonly believed to be equivalent to SRS.

Research supports using these learning tools and strategies to modify the traditional classroom to accommodate the knowledge explosion because they increase opportunities for productive learning.[6] Although there is much literature about SRS and attendance, performance outcomes, and engagement in the learning process, there is little information on metacognitive self-regulation and SRS use.[6] This is important because SRS have been linked with improved ability by students to identify key concepts as well as improved ability by instructors to gauge students' levels of preparation and understanding.[7] Learning activities can be scaffolded into instructional design to support learning and engage the self-regulation process that occurs in the context of lectures.[8] Self-regulation pertains to one's individual management of learning to reach personal educational goals. Furthermore, the literature suggests that this process can also be used in the clinical education setting to engage learners in self-regulation that supports the development of clinical skills.[8]

Recent research explores the accuracy of self-monitoring that occurs "in the moment" rather than through global self-assessments that occur after the experience and reflect more general feelings about learning.[9] Self-monitoring increases the level of accuracy that is specific to the area of knowledge. According to research, another aspect of learning pertains to the level of supervision and self-regulated learning. Studies find that unsupervised learning interventions are associated with poorer overall outcomes than supervised ones.[8] Focusing on learning as a shared responsibility of student and instructor could also help prepare learners for collaborative clinical practice. Comparing different methods of garnering feedback is relevant to research because SRS provide this support. There is a notion among instructors that forms of garnering feedback that lack the anonymity of SRS also provide this level of support, promoting a group learning effort with individual self-reflection components. The LTRS style reflects the overt, public nature of feedback methods more commonly used in clinical and medical education settings. Polling responses in these contexts are sometimes made by uncertain learners, whose lack of confidence can make them vulnerable to peer influence and groupthink more than individual reflection. Insight gained from anonymous SRS use, followed by reflection on the visualized group results, improves metacognitive self-regulation and self-reflection.

METHODS

Participants and Design

Participants in this study were first-year students attending a PA program at a large university in the Southwestern United States. An online demographic survey administered during the first week of the course revealed that, of the 53 participants (29 female participants; M_{age} = 26.3 years; age range = 22–41 years), 17.0% were Hispanic, 9.4% African American, 18.9% Asian, 47.2% white, and 7.5% other. When queried about their previous experience with SRS, 29.8% of participants reported "no experience," 21.3% "very little," and 40.4% "much experience."

The SRS was integrated into the course design to pretest for general knowledge, evaluate comprehension of prereading assignments, assess recollection of material taught in the previous class session, and preview upcoming lecture topics. The SRS was employed during weeks 1 to 5 and the LTRS during weeks 8 to 12. The course was structured so that polling questions were embedded in the PowerPoint presentations and were easily modified for the 2 different polling systems. The scales used in this study to measure metacognition were the Metacognition in Lecture Scale (MCiL) and the Metacognitive Attribution to Device Scale (MCADS). The MCiL was designed to measure metacognition that learners experience in lecture contexts and the MCADS to measure metacognition that learners may attribute directly to the response device used. An examination of scale alphas demonstrated a strong degree of reliability for both the SRS (α = 0.91) and the LTRS (α = 0.94) scales measuring learner perception of metacognition elicited through lecture, and of the scale that measured learner attribution of metacognition to response device for both SRS (α = 0.72) and LTRS (α = 0.70).

Table 1. Inventory Mean Values, SD, and *t*-Test Results According to Construct Components, Treatment Versus Comparison

Construct	M_C	M_P	SD_C	SD_P	Pearson's *r*	*df*	*t* Statistic
Self-regulation							
Note-taking	3.71	2.92	0.347	0.251	0.96	4	9.644*
Self-monitoring	3.79	3.35	0.101	0.232	0.67	6	3.257†
Peer/social comparisons	4.85	4.76	1.31	2.25	0.83	3	0.2333
Understanding	4.71	3.40	0.219	0.124	0.48	6	8.093*

*Significance at .001.

†Significance at .01.

M_c, clicker item or construct mean; M_p, paddle mean; SD_c, clicker standard deviation; SD_p, paddle (or LTRS) item standard deviation.

Table 2. Construct Descriptive Statistics According to Indices by Item

Item No.	Index	M_C	SD_C	M_P	SD_P
	Self-regulation: note-taking				
MCiL Q1	Clarifies purpose	3.91	1.63	2.86	1.67
MCiL Q4	Decisions about notes to take	3.68	1.94	2.86	1.89
MCiL Q5	Decisions about information to ignore	3.02	1.63	2.35	1.48
MCiL Q14	Helps decide concepts to highlight	4.58	1.96	3.73	1.77
MCiL Q15	Helps with ability to paraphrase	3.37	1.66	2.80	1.54
	Index total	3.70	1.80	2.90	1.70
	Self-regulation: self-monitoring				
MCiL Q6	See how lecture fits with text	3.87	1.62	3.20	1.79
MCiL Q7	Helps to understand my notes	3.43	1.73	2.92	1.64
MCiL Q8	Does not help to see if notes I write are important	3.83	1.72	3.96	1.93
MCiL Q9	Helps identify questions to ask	3.29	1.38	2.84	1.52
MCiL Q10	Helps to refocus	4.13	1.57	3.31	1.73
MCiL Q11	Does not help with questions to ask	4.10	1.73	4.06	1.90
MCiL Q12	Focus on questions to note and ask later	3.87	1.58	3.14	1.43
	Index total	3.20	1.60	3.20	1.70
	Peer/social comparisons				
MCADS Q2	Histogram/class indicates I was wrong (reading), gained understanding	3.47	1.57	3.46	1.66
MCADS Q3	Histogram/class indicates I was wrong (lecture), gained understanding	4.64	1.81	3.46	1.89
MCADS Q5	Compared self to peers	5.06	1.63	5.92	1.37
MCADS Q6	Compared self to histogram/class	6.24	1.16	6.18	1.13
	Index total	4.90	1.50	4.80	1.50
	Conceptual understanding				
MCiL Q2	Understanding of subject	4.61	1.50	2.86	1.67
MCiL Q3	Reading for lecture on track	4.61	1.80	3.14	1.83
MCiL Q13	Helps decide key concepts for notes	4.54	1.69	3.76	1.81
MCADS Q1	Gained understanding when wrong	5.57	1.12	3.82	1.76
MCADS Q4	Impacted more than hand-raising	4.98	1.81	3.37	2.18
MCADS Q7	Understanding of concepts improved	4.63	1.40	3.22	1.61
MCADS Q8	Deeper thoughts about course concepts	4.04	1.57	3.61	1.54
	Index total	4.70	1.60	3.40	1.80

M_c, clicker item or construct mean; MCADS, Metacognitive Attribution to Device Scale; MCiL, Metacognition in Lecture Scale; M_p, paddle mean; SD_c, clicker standard deviation; SD_p, paddle (or LTRS) item standard deviation.

RESULTS

Results were grouped according to the metacognitive component that the questions were designed to inform. Results of SRS were tracked for grading purposes, whereas LTRS results were used as a measure of group understanding. Mean values for SRS and LTRS are detailed in Table 1. Table 1 lists mean values of indices, Pearson's *r*, and *t* statistics. Table 2 lists descriptive statistics for the individual items aligned with the respective index. Table 3 displays changes in learner perception of metacognitive influence and effect size.

DISCUSSION

The results confirm previous findings that demonstrate significance of the SRS over the LTRS for scales.[4,10] Items were individually examined, including statistical analysis for effective size to further validate the results. Although multiple *t*-tests of similar items can falsely indicate statistical significance, the authors postulate that these results are significant for 2 reasons. The first is that previous results comparing scales resulted in significance. The second is that metacognitive self-regulation is a complicated construct involving measurement of several components of the construct.

Significance of Related Items

Significance was found for self-regulation, self-monitoring, and understanding in favor of the SRS (treatment) over the LTRS (comparison). Both methods resulted in high degrees of social comparison. However, the items "comparison of self to

Table 3. Contrast of Treatment (SRS) and Comparison (LTRS) Item-by-Item Indicating Change in Metacognitive Influence

	SRS		LTRS				95% CL		
MCiL	M	SD	M	SD	$t_{(47)}$	*P*	LL	UL	Cohen's *d*
Q1	3.96	1.68	2.85	1.69	3.71	.001*	0.543	1.832	−0.660†
Q2	4.70	1.52	3.71	1.89	2.92	.005‡	0.323	1.760	0.578†
Q3	4.70	1.78	3.25	1.86	4.34	.000*	0.771	2.104	0.797†
Q4	3.80	1.92	3.00	1.96	2.32	.025‖	0.119	1.673	0.528†
Q5	2.96	1.65	2.33	1.44	2.39	.021‖	0.109	1.266	0.411§
Q6	3.94	1.62	3.21	1.76	2.59	.013‖	0.162	1.297	0.431§
Q7	3.44	1.79	2.92	1.66	1.63	.111	−0.133	1.258	0.302§
Q8	4.22	1.73	4.06	1.86	4.02	.000*	0.615	1.844	0.089
Q9	3.30	1.40	2.75	1.45	2.04	.047‖	0.009	1.157	0.385§
Q10	4.12	1.60	3.31	1.72	2.54	.015‖	0.177	1.532	0.489§
Q11	3.94	1.73	3.98	1.93	−.22	.828	−0.850	0.683	−0.022
Q12	3.86	1.60	3.27	1.46	2.00	.051	−0.003	1.295	0.385§
Q13	4.56	1.70	3.85	1.82	2.16	.036‖	0.051	1.490	0.402§
Q14	4.50	1.96	3.81	1.78	−1.06	.297	−1.046	0.326	0.368§
Q15	3.30	1.66	2.85	1.58	1.37	.179	−0.227	1.185	0.278§
MCADS									
Q1	5.54	1.11	3.96	1.77	5.23	.000*	0.973	2.194	1.068#
Q2	4.48	1.54	4.69	1.64	−.66	.516	−0.933	0.475	−0.132
Q3	4.60	1.57	3.40	1.83	3.42	.001*	0.497	1.919	0.705†
Q4	4.98	1.82	3.48	2.21	3.68	.001*	0.662	2.255	0.740†
Q5	5.04	1.64	5.94	1.38	−2.54	.014‖	−1.567	−0.183	−0.595†
Q6	6.22	1.17	6.21	1.15	.18	.860	−0.430	0.513	0.009
Q7	4.66	1.39	3.67	1.64	3.20	.002‡	0.372	1.628	0.650†
Q8	4.08	1.56	3.21	1.65	2.96	.005‡	0.287	1.504	0.541†

*Seven-point Likert scale. Significance and effect size are indicated as follows: *P* = .001.

†Effect size for Cohen's *d*: moderate <0.5.

‡Seven-point Likert scale. Significance and effect size are indicated as follows: *P* = .01.

§Effect size for Cohen's *d*: small effect size is <0.2.

‖Seven-point Likert scale. Significance and effect size are indicated as follows: *P* = .05.

#Effect size for Cohen's *d*: large <0.8.

CL, confidence limit; LL, lower limit; LTRS, low-technology response system; M, mean; MCADS, Metacognitive Attribution to Device Scale; MCiL, Metacognition in Lecture Scale; SRS, student response system; UL, upper limit.

peers" demonstrated significance in favor of the SRS. This suggests that the SRS may provide a better basis for comparison between self and peers. Whether and how these social influences compare requires further study.

Statistically Significant Items

Metacognitive monitoring for level of preparation, knowledge, and level of confidence in course materials improved with SRS as compared to the LTRS (Table 3). Confidence intervals and effect sizes were moderate to strong, indicating statistical significance for the following items: (1) note-taking improvements; (2) gaining understanding in lecture; (3) gauging level of preparation for lecture; (4) identification of important concepts and key words in lecture and for note-taking; (5) refocusing; and (6) calibrations aligning understanding with correct answers to specific questions and understanding lecture concepts, self-comparisons to peers, more beneficial than hand-raising, and encouraging deeper cognition.

Regarding refocusing, most participants felt that attention was reengaged. It is interesting that, in this learning environment, focus was affected because polling questions were not dispersed throughout the lecture.

The item "SRS doesn't help me to know if what I'm writing is important" demonstrated a very high degree of significance, with a minuscule effect size. This result indicates that perception varied among individuals, a key component of metacognition. Additional data are needed to ferret out the individual nature of the effect. This question was asked both as a positively framed and a reverse-coded item, but no significance was found when framed in either way. This identified an inconsistency between results for these corresponding items.

Learners did not perceive either response system as contributing more to their ability to take notes, discern sufficiency of notes, or focus better on difficult topics. Both systems drew students' attention to what is important, through the types of questions asked and topics selected (Table 2, MCiL Q11 and MCiL Q12).

Responses for both the SRS and the LTRS were highly positively skewed, with virtually identical distribution of responses (Table 2, MCiL Q2 and MCiL Q6). These items pertained to understanding that occurred with use of response methods in real-time lectures. This means that both response methods helped learners connect relevancy of prelecture preparation and conceptual understanding. Questions addressing peer comparisons did not demonstrate significance between comparison and treatment.

Students using the LTRS began to compare their responses to others before the correct response was revealed, whereas students using the SRS compared results following the anonymous display of correct responses on-screen, not by directly viewing peer responses. It is natural to engage in comparisons with information at hand. Anonymity associated with use of the SRS attenuated students' sense of vulnerability. Although these items were not significantly different from each other, both ratings were significantly higher on average than other items on the inventories, meaning that while some items indicate a slight influence on metacognition, the social comparison items indicated that response systems had a large degree of influence on metacognition. The presentation of questions was different with each system, raising the possibility that comparisons based on the same type of information were filtered when responding anonymously, as opposed to answering in a visible way to peers.

Strengths and Limitations

Strengths of this study were the rigor of the analysis, the number of participants, interdisciplinary collaboration, and the quasi-experimental design of the study. Limitations of the study were related to how the systems were employed. Results from the SRS were tracked for grading purposes, whereas LTRS results were used as a measure of group understanding. The reduced ability to track LTRS responses may have resulted in some learners experiencing less pressure to prepare for lectures when they know in advance that results will not be graded. This feature will be addressed in subsequent studies.

CONCLUSION

Each year, students present with higher levels of literacy in technology, challenging faculty members to develop increased technological skills. Faculty members may find themselves exploring new teaching methods to meet the evolving learning needs of these digital natives.

Metacognitive judgment can be tested and taught by simply following up an SRS question with a second question that queries learners about their confidence level in answering the question. An examination of the type of question, knowledge recall or deductive reasoning skills used, and individual metacognitive judgment and confidence of the student could provide further insight into this process. Factor analysis of the measurement instruments is recommended because this subjects the instrumentation to rigorous analysis, showing independent latent variables that may be measured. Further studies are needed to examine how these instruments would factor into the measurement of metacognition.

It is important to examine metacognition and the improved ability to gauge level of knowledge, set up a learner-centered environment, and effectively engage learners in deeper cognition. Instructional design with SRS questions may also be improved by expanding its application to support development of self-regulation in clinical learning contexts with less supervision. Findings of this study support the practice of using electronic SRS to poll PA students and enhance learning.

Melanie Brady, EdD, is an adjunct assistant professor of Clinical Education at the University of Southern California, Los Angeles, California.

Christopher P. Forest, MSHS, PA-C, is a professor and founding program director of the Master of Science Physician Assistant Program at California State University, Monterey Bay. He conducted this research while an assistant professor of family medicine at the University of Southern California Keck School of Medicine in Alhambra, California.

Correspondence should be addressed to: Melanie Brady, EdD, Rossier School of Education, University of Southern California, 3740 Trousdale Parkway, Waite Philips Hall (WPH) 404, Los Angeles, CA 90089. Telephone: (714) 658-4477; Email: melanilb@usc.edu

REFERENCES

1. Forest CP. The effect of audience response systems on adult learning: evidence-based rationale and ARS implementation guide. *J Physician Assist Educ.* 2012;23(4):54-59.
2. Grahame JA. Digital note-taking: discussion of evidence and best practices. *J Physician Assist Educ.* 2016;27(1):47-50.
3. Hunsu NJ, Adesope O, Bayly DJ. A meta-analysis of the effects of audience response systems (clicker-based technologies) on cognition and affect. *Comput Edu.* 2016;94:102-119.
4. Brady M, Seli H, Rosenthal J. "Clickers" and metacognition: a quasi-experimental comparative study about metacognitive self-regulation and use of electronic feedback devices. *Comput Edu.* 2013;65:56-63.
5. Zimmerman BJ. A social cognitive view of self-regulated academic learning. *J Educ Psychol.* 1989;81(3):329-339.
6. De Corte E, Verschaffel L, Masui C. The CLIA-model: a framework for designing powerful learning environments for thinking and problem solving. *Eur J Psychol Educ.* 2004;19(4):365-384.
7. Barber M, Njus D. Clicker evolution: seeking intelligent design. *CBE Life Sci Educ.* 2007;6(1):1-8.
8. Brydges R, Butler D. A reflective analysis of medical education research on self-regulation in learning and practice. *Med Educ.* 2012;46(1):71-79.
9. Eva KW, Regehr G. Exploring the divergence between self-assessment and self-monitoring. *Adv Health Sci Educ Theor Pract.* 2011;16(3): 311-329.
10. Brady M, Rosenthal J, Forest CP. The effect of audience response systems on metacognition in graduate students: a two-year mixed method study. Presented at: Innovations in Medical Education Conference; February 21, 2015; San Gabriel, CA.