

REPORT DOCUMENTATION PAGE

*Form Approved
OMB No. 0704-0188*

The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing the burden, to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.

PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.

1. REPORT DATE (DD-MM-YYYY)			2. REPORT TYPE		3. DATES COVERED (From - To)	
4. TITLE AND SUBTITLE				5a. CONTRACT NUMBER		
				5b. GRANT NUMBER		
				5c. PROGRAM ELEMENT NUMBER		
6. AUTHOR(S)				5d. PROJECT NUMBER		
				5e. TASK NUMBER		
				5f. WORK UNIT NUMBER		
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)				8. PERFORMING ORGANIZATION REPORT NUMBER		
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)				10. SPONSOR/MONITOR'S ACRONYM(S)		
				11. SPONSOR/MONITOR'S REPORT NUMBER(S)		
12. DISTRIBUTION/AVAILABILITY STATEMENT						
13. SUPPLEMENTARY NOTES						
14. ABSTRACT						
15. SUBJECT TERMS						
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON	
a. REPORT	b. ABSTRACT	c. THIS PAGE			19b. TELEPHONE NUMBER (Include area code)	

Abstract Words: 282
Narrative Words: 5,274
References: 29
Tables: 3
Figures: 1
Appendices: 1

**Exploring Clinical Reasoning in Novices:
A Self-Regulated Learning Microanalytic Approach**

****Anthony R. Artino, Jr.¹, Timothy J. Cleary²,
Ting Dong³, Paul A. Hemmer⁴, and Steven J. Durning⁵**

Uniformed Services University of the Health Sciences
4301 Jones Bridge Road, Bethesda, Maryland 20814

1. Associate Professor, Department of Preventive Medicine and Biometrics, Uniformed Services University of the Health Sciences (USU)
2. Associate Professor, Graduate School of Applied and Professional Psychology, Rutgers University, New Brunswick, New Jersey
3. Research Associate, Department of Preventive Medicine and Biometrics, USU
4. Professor and Vice Chairman, Educational Programs, Department of Medicine, USU
5. Professor, Department of Medicine, USU

**** Correspondence:** Anthony R. Artino, Jr., PhD, Associate Professor, Department of Preventive Medicine and Biometrics, 4301 Jones Bridge Road, Bethesda, Maryland 20814
Email: anthony.artino@usuhs.edu; Phone: (301) 295-3693

Abstract

Objectives: The primary objectives of this study were to examine the regulatory processes of medical students as they completed a diagnostic reasoning task and to examine whether the strategic quality of these regulatory processes were related to short-term and longer-term medical education outcomes.

Methods: A self-regulated learning (SRL) microanalytic assessment was administered to 71 second-year medical students while they read a clinical case and worked to formulate a most likely diagnosis. Verbal responses to open-ended questions targeting forethought and performance phase processes of a cyclical model of SRL were recorded verbatim and subsequently coded using a framework from prior research. Descriptive statistics and stepwise linear regression models were used to examine the relationships between the SRL strategies and several outcomes.

Results: Most participants (90%) reported focusing on specific diagnostic reasoning strategies during the task (metacognitive monitoring), but only about one-third of students referenced these strategies (e.g., identifying symptoms, integration) in relation to their task goals and plans for completing the task. After accounting for prior undergraduate achievement and verbal reasoning ability, strategic planning explained significant additional variance in course grade ($\Delta R^2=.15, P<.01$), second-year grade point average ($\Delta R^2=.14, P<.01$), United States Medical Licensing Examination Step 1 score ($\Delta R^2=.08, P<.05$), and National Board of Medical Examiner subject exam score in internal medicine ($\Delta R^2=.10, P<.05$).

Conclusions: These findings suggest that most students in the formative stages of learning diagnostic reasoning skills are aware of and think about at least one key

diagnostic reasoning process or strategy while solving a clinical case, but a substantially smaller percentage set goals or develop plans that incorporate such strategies. Given that students who developed more strategic plans achieved better outcomes, the potential importance of forethought regulatory processes is underscored.

Introduction

Producing competent physicians is the goal of medical education. Although most medical trainees succeed in medical school and residency training, some struggle to meet accepted standards. When underperformance occurs, medical educators typically want to gain a better understanding of the reasons underlying the subpar performance. Unfortunately, evaluating these reasons and providing effective forms of feedback and remediation to assist these strugglers has proven to be quite challenging.^{1,2} From our perspective, one way to provide medical educators with better diagnostic information about the *key causal factors* underlying poor performance is to administer assessment tools that reliably target such factors.³ In medical education contexts, students may struggle to meet accepted standards for a plethora of reasons, including knowledge and skill deficits, poor motivation, and even learning or emotional disabilities.

However, there is an emerging literature within medical education, as well as within postsecondary school contexts, that deficits in self-regulated learning (SRL), such as poor planning, inadequate self-monitoring, and insufficient self-reflection, are robust predictors of a range of performance indicators.⁴⁻⁶ In addition, given that SRL is often considered to be a modifiable process and a teachable skill, rather than a fixed trait,⁷ it may behoove medical educators to develop and use assessment tools that not only target self-regulation skills but do so in relation to how students regulate before, during, and after key clinical tasks. In particular, determining whether and how SRL plays a role in medical student learning and performance has implications for intervention and remediation programs that can help to change the beliefs and behaviors of medical trainees. The purpose of this study was to employ a context-specific

assessment procedure, called SRL microanalysis, during a clinical reasoning task in order to identify the strategic quality of SRL processes exhibited by novice learners during clinical reasoning and to investigate whether these regulatory processes were related to important medical school performance outcomes.

Self-Regulated Learning Microanalysis

Social-cognitive researchers have defined SRL as “the degree to which [students] are metacognitively, motivationally, and behaviorally active participants in their own learning processes.”^{8, p. 49} From this perspective, SRL is a contextualized, teachable skill that operates as a three-phase cyclical process. In this model, forethought processes that precede action (such as goal setting and planning) impact learning efforts (such as self-monitoring and strategy use), which then influence how learners react to and judge their performance successes or failures (self-reflection; see Figure 1).⁹ In a general sense, sophisticated self-regulated learners are those who become *strategically engaged* in their own learning and are likely to display strategic thinking and action before, during, and after learning.¹⁰

Investigators who employ SRL microanalytic techniques assume that students’ thoughts, feelings, and actions are contextually bound and thus will often fluctuate across educational tasks.^{11,12} Self-regulated learning microanalysis is a structured interview protocol that attempts to capture students’ cyclical regulation (i.e., their thoughts, feelings, and actions) as they engage in particular tasks. It involves administering questions about students’ regulatory processes as they approach, complete, and reflect on their performance across specific tasks or situations. This

approach contrasts sharply with traditional SRL assessment approaches, such as self-report questionnaires, which tend to measure self-regulation as a global, fixed entity.¹³

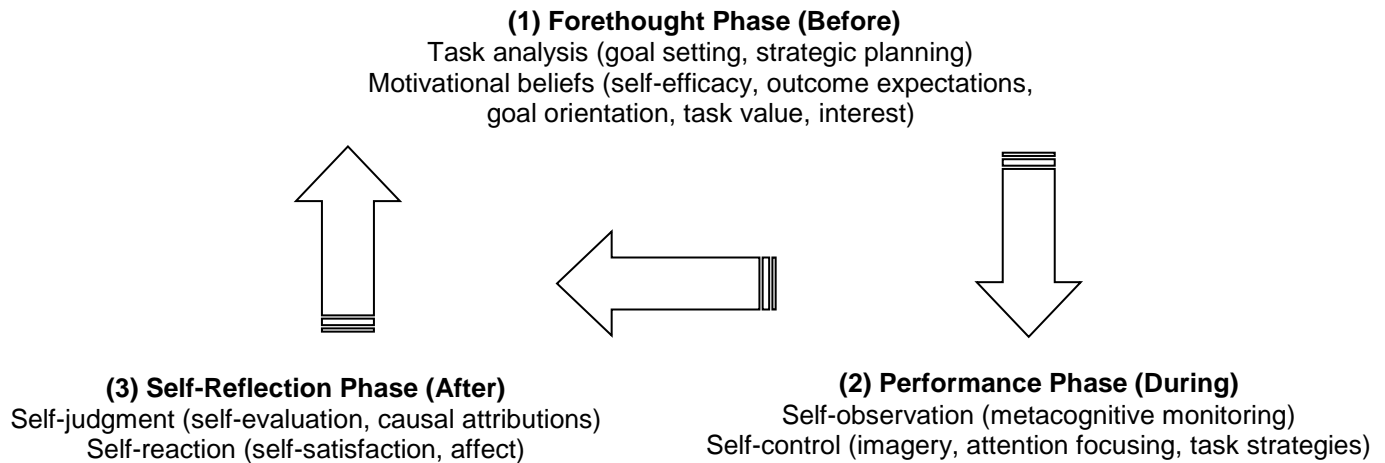


Figure 1. Self-regulated learning conceptualized as a three-phase model containing forethought (before), performance (during), and self-reflection (after) processes. In this model, self-regulation is hypothesized to be a teachable skill that operates in a cyclical manner, as shown in the figure. (Adapted from Zimmerman BJ, Campillo M. Motivating self-regulated problem solvers. In: Davidson JE, Sternberg RJ, eds. *The Nature of Problem Solving*. New York, NY: Cambridge University Press; 2003:239. Used with permission.)

The majority of research utilizing SRL microanalytic protocols has involved non-medical education tasks, such as free-throw shooting,¹⁴ volleyball serving,⁶ and studying.¹⁵ The lone exception is a recent qualitative pilot study.¹⁶ In this latter study, the authors were interested in demonstrating the potential applicability and usefulness of SRL microanalysis to a clinical task (venipuncture). Even though the authors reported that high achievers were very strategic as they approached and performed the venipuncture task, whereas low achievers were more outcome focused, the small sample size precluded the use of inferential statistics. The authors concluded

that although SRL microanalysis has some potential, it is critical for researchers to evaluate the reliability, validity, and generalizability of this approach.

Study Purpose and Hypotheses

The overall goal of the current study was to develop and customize a SRL microanalytic protocol to examine the nature of medical students' regulatory processes (i.e., goal setting, strategic planning, metacognitive monitoring) as they solved a diagnostic reasoning task. We elected to focus on diagnostic reasoning, or more broadly, clinical reasoning, because it lies at the core of what physicians do in practice—they decide on diagnoses and institute treatments.¹⁷ In addition, SRL microanalytic protocols are applicable to tasks that have a clear beginning, middle, and end.¹¹ It is important to note at the outset, however, that our plan was not to use SRL microanalysis to comprehensively assess students' reasoning skills, but rather as a method to measure their regulatory approach and strategic thinking in relation to a clinical reasoning task. There is some obvious conceptual overlap between diagnostic reasoning and SRL processes (e.g., strategic skills, monitoring), but our primary goal was to focus on participants' regulatory processes during the reasoning task. We recognize that the assessment of clinical reasoning is quite challenging and multi-dimensional and that perspectives on clinical reasoning (and its assessment) are quite varied.^{18,19}

We had two specific objectives in this study. First, we wanted to identify the strategic quality of the goals, plans, and metacognitive monitoring of second-year medical students (who were novices in clinical reasoning) as they attempted to solve a specific case. Given that this objective was exploratory in nature, we descriptively examined the frequency and quality of strategic

thinking without making *a priori* hypotheses. The second objective was to evaluate whether these self-regulatory processes were related to short and longer-term medical school performance outcomes. Based on findings from prior microanalytic research, we hypothesized that students who focused on the strategic process of diagnostic reasoning before and during the actual task, as measured by three microanalytic questions, would exhibit better performance on course outcomes and standardized medical education exams, even after controlling for prior undergraduate achievement and verbal reasoning ability.

Methods

This study was conducted at the F. Edward Hébert School of Medicine, Uniformed Services University of the Health Sciences (USU); the study protocol was administered across two academic years (2010-2011 and 2011-2012). At the time of the study, USU offered a traditional four-year curriculum: two years of basic science courses followed by two years of clinical rotations (clerkships).

Participants and Study Context

Second-year medical students were recruited from an Introduction to Clinical Reasoning (ICR) course. Students who volunteered to participate in the study were offered three extra credit points in the ICR course, whereas non-participants could earn the same extra credit points through an alternate means.

At the time of this study, the ICR course was offered in the second year of medical school and represented students' initial exposure to formal instruction in diagnostic reasoning. The ICR course is organized as a series of lectures and small-group activities that expose students to various symptoms, physical exam findings, laboratory test abnormalities, and syndromes. Within the small groups, students are asked to work through paper cases to synthesize presenting symptoms and findings into a problem list, differential diagnosis, and initial management plan. Prior to the ICR course, students had no prior didactic or significant clinical experiences in diagnostic reasoning and thus were viewed as "novice learners" for the task described below.

Procedures

During the last month of the 10-month course, the primary author administered a diagnostic reasoning task to each participant on an individual basis during a 25-30 minute session outside of the normal ICR class time. Participants were asked to read a one-page paper case and then complete a post-encounter form (PEF). While working through the paper case, students were not allowed to use secondary aids, such as books or computers. On the PEF, students were prompted to write a summary statement, prioritize the problem list, identify a differential diagnosis, and record a most likely diagnosis. The feasibility, reliability, and validity of the paper case and PEF are supported by previous research.²⁰

Using guidelines provided by Cleary,¹¹ an SRL microanalytic assessment protocol was developed to examine the strategic quality of participants' regulatory processes before completing the diagnostic reasoning task and during task completion. This assessment methodology uses open and close-ended questions that target forethought, performance, and self-

reflection phase processes of a cyclical model of self-regulation (see Figure 1).⁹ In this manuscript, we discuss forethought (i.e., goal setting, strategic planning) and performance phase processes (i.e., metacognitive monitoring), given our primary objective of evaluating how novice learners approached and monitored their actions during the reasoning task. All sessions were audio recorded for later transcription by a research assistant, and all participants provided written informed consent. The university's Institutional Review Board approved the study protocol.

Measures

Goal setting. A single-item, microanalytic measure designed to assess participant task goals was administered immediately following students' initial reading of the clinical case but prior to beginning the PEF. Participants were asked, "Do you have a goal (or goals) in mind as you prepare to do this activity? If yes, please explain." Participant responses were coded into one of seven categories: *task-specific process*, *task-general process*, *outcome*, *self-control*, *non-task strategy*, *do not know/none*, and *other*. The coding scheme was an adaptation of coding rubrics used in prior research across different psychomotor tasks.^{14,16} The *task-specific process* category involved responses pertaining to five key strategies typically used for diagnostic reasoning tasks: *identifying symptoms*; *identifying contextual factors (e.g., social or environmental factors)*; *prioritizing relevant symptoms*; *integrating/synthesizing symptoms and other case features*; and *comparing/contrasting diagnoses*). These strategies were identified by reviewing the clinical reasoning literature and obtaining expert consensus from three experienced clinicians through their review of preliminary transcripts of student responses.^{18,19,21} An example of a response coded for this category is, "My goal is to figure out how the symptoms connect." The *task-general process* category involved responses pertaining to a general method or procedure to

follow, such as “To do all the right steps to solve the case.” An *outcome* response pertained to getting the correct diagnosis (e.g., “To get the correct diagnosis on my first attempt”), whereas the *self-control* category involved responses pertaining to effort, focus, concentration, or other management tactics designed to enhance performance on the task (e.g., “To make sure I fully concentrate on what I am doing”). The *non-task strategy* category involved responses pertaining to some outcome or process that was irrelevant to the actual task or one which was not possible given the constraints of the given task (e.g., “To find more information on the Internet”). Finally, the *do not know/none* category involved responses that explicitly indicated that the student did not have a task goal (e.g., “Nothing really, just do it”), whereas the *other* response category included any response that did not fit into the above categories. All participant responses were coded independently by two of the authors (ARA and TD), and a percent agreement of 90% was attained by the two coders. Disagreements were resolved through discussion among all authors.

Strategic planning. A single-item, microanalytic measure designed to assess participant plans for approaching the diagnostic reasoning task was administered immediately after the goal question but preceding the student’s attempt to generate an accurate diagnosis. Participants were asked, “What do you think you need to do to perform well on this activity?” Similar coding procedures to those used for the goal setting item were adhered to for the strategic planning measure; however, the responses were coded into one of six categories: *task-specific process*, *task-general process*, *self-control*, *non-task strategies*, *do not know/none*, and *other*. The *outcome* category was removed because it did not relate conceptually to planning. All participant responses were again coded independently by two of the authors (ARA and TD), and a percent agreement of

88% was attained by the two coders. Disagreements were resolved through discussion among all authors.

Metacognitive monitoring. Another single-item, microanalytic measure was developed to examine each participant's thoughts during task completion. To standardize the administration of this measure, we stopped all participants immediately after they recorded their first differential diagnosis on the PEF. The majority of the participants reached this point in the protocol in approximately 10 minutes after beginning the task. Participants were asked, "As you have been going through this process, what has been the primary thing you have been thinking about or focusing on?" If a response was provided, they were probed, "Is there anything else you have been focusing on?" For these questions, student responses were again coded into one of seven categories. Five categories were similar to the other measures, *task-specific process*, *task-general process*, *outcome*, *self-control*, and *other*. Two additional categories were added to this coding scheme, *perceived ability* and *task difficulty*. The *perceived ability* category involved responses pertaining to students' perceived ability to perform the task and/or their knowledge related to the task. Examples of responses coded for this category are "I was never that good at diagnosing" and "I have no idea what these terms mean." The *task difficulty* response category involved responses pertaining to the inherent challenges or difficulty level of the task (e.g., "There is not enough information in this case"). Participant responses were again coded independently by two of the authors (ARA and TD), and a percent agreement of 88% was attained by the two coders. Disagreements were resolved through discussion among all authors.

Given the open-ended nature of these microanalytic questions, many participants provided multiple responses for each question. Thus, it was possible for any given response to receive multiple codes to distinct categories. However, similar statements within a response that were coded to the same category were counted only as a single instance of that category.

Performance outcome measures

To exam the relations between participants' regulatory processes and their performance in medical school, a series of short and longer-term measures of medical knowledge were used: ICR course grade, second-year grade point average (GPA), the United States Medical Licensing Examination (USMLE) Step 1, and the National Board of Medical Examiners (NBME) subject examination in internal medicine.

ICR course grade. Student performance in ICR was calculated as the average score on three course-specific exams administered at the end of each trimester. Exams 1 and 2 were 65-item, multiple-choice tests that employed clinical vignettes. The internal reliabilities of these exams were considered adequate for teacher-made exams, with actual Cronbach's alphas of .60 and .90, respectively. Exam 3 was a cumulative, short-essay test that was scored by the course director. It required that students read case vignettes and then complete several tasks, including posing additional history and physical examination questions, constructing a differential and most likely diagnosis, and proposing next steps in patient management.

Second-year GPA. A second-year medical school GPA was calculated by multiplying each course grade in year two by the number of contact hours for the given course, summing the

weighted grades across courses, and then dividing the sum by the total number of contact hours. The resulting averages were converted to a common 4-point scale ranging from 0.0 to 4.0.

USMLE Step 1. All students at USU are required to pass the USMLE Step 1 examination in order to graduate. Students in the present study completed the Step 1 examination at the end of their second year (i.e., approximately 1 month after completing the ICR course and the diagnostic reasoning task administered in this study). Scores on the USMLE Step 1 examination are given in three digits ranging from 140 to 280.

NBME subject examination in internal medicine. The NBME offers a variety of multiple-choice subject examinations for medical students. In this study, students completed the subject exam in internal medicine during their third year of medical school (i.e., at the end of their internal medical clerkship, approximately 6 to 12 months after completing the ICR course and the diagnostic reasoning task administered in this study). Scores on the NBME examination are reported on a 100-point scale ranging from 0 to 100.

Analysis

Prior to analysis, we screened the data for accuracy and missing values and checked each variable score for normality. Next, to investigate the representativeness of our sample, we compared students who completed the study and those who did not on gender (using a chi-square test), as well as undergraduate GPA, Medical College Admission Test (MCAT) verbal reasoning score, and first-year medical school GPA (using independent samples *t* tests). Following these preliminary analyses, we used both descriptive and inferential statistics to address the research

questions. In terms of descriptive analysis, we were interested in examining the percentage of participants who set strategic goals and plans *prior to* completing the diagnostic reasoning task and those who referenced the strategic process *during* the task (metacognitive monitoring). We defined “strategic” in terms of five key strategies typically used for diagnostic reasoning: *identifying symptoms; identifying contextual factors (e.g., social or environmental factors); prioritizing relevant symptoms; integrating/synthesizing symptoms and other case features; and comparing/contrasting diagnoses.*

To examine whether the strategic quality of novice learner’s approach to the diagnostic reasoning task was related to important medical school outcomes, we conducted a two-step, hierarchical multiple regression analysis across the four performance outcomes (ICR grade, second-year GPA, USMLE Step 1, and NBME). For each hierarchical regression model, we entered undergraduate GPA and MCAT verbal reasoning score in step one and the microanalytic measures that were correlated with the performance outcomes in step two. Because we believed prior achievement and verbal reasoning ability could have important effects on performance when solving a written case, we adjusted for undergraduate GPA and MCAT verbal reasoning score. In doing so, we hoped to control for any preexisting differences in students’ prior achievement and ability to read critically, comprehend, and draw inferences and conclusions from written material.

Given that many participants provided multiple, codeable responses to each question, we elected to transform the categorical responses for all microanalytic questions to a metric scale. The scoring system was designed to capture the strategic quality of the participants’ regulatory

processes during the specific task, with greatest weight given to responses that reflected one or more of the five strategic steps identified for the diagnostic reasoning task. This scoring system was developed before data analysis and was an adaptation of a prior scoring scheme.²² Theory, prior research, and expert consensus were the major determining factors in deciding the direction (i.e., positive or negative) and quantity of points given for each coded response (see the Appendix for an example of how this scoring scheme was applied).

Results

Of the 342 second-year medical students invited to participate, 71 students (21%) completed the entire study. The sample included 46 men (65%) and 25 women, which is similar to the overall medical student population at USU (71% men). Preliminary analysis comparing participants and non-participants were mixed. There were no statistically significant differences between the two groups on gender, $\chi^2(1, N = 336) = 2.76, P = .10$, undergraduate GPA, $t(340) = -1.17, P = .24$, or MCAT verbal reasoning scores, $t(340) = -1.02, P = .31$. However, on average, the study participants were slightly older ($M = 29.07, SD = 3.57$) than non-participants ($M = 28.07, SD = 3.36$), $t(334) = 2.19, P < .05$, and displayed a higher first-year GPA ($M = 3.29, SD = 0.45$) than non-participants ($M = 3.04, SD = 0.50$), $t(349) = 3.70, P < .001$.

Forethought phase processes

Descriptive analysis for the two forethought microanalytic questions (goal setting and strategic planning) revealed that approximately two-thirds of the participants were *not* focused on specific diagnostic reasoning tactics. In terms of goals, while 32% of participants did provide a strategic goal, approximately 50% of participants either conveyed goals that focused on the outcome of

getting the correct diagnosis (18%) or did not report any type of goal (31%). More detailed analysis revealed that of the 23 participants who focused on task-specific processes, most (70%) reported focusing on the most basic strategic step of *identifying symptoms*.

A similar pattern of results emerged for the planning question, as the majority of respondents reported plans that did not pertain to the diagnostic reasoning process, including *self-control* (16%; e.g., attention focusing or concentration), *non-task strategies* (16%; e.g., seeking information or help [neither of which was possible during the task]), or *other* (25%; e.g., getting the right answer). Furthermore, only 24 participants (34%) reported strategic plans before engaging in the task. Similar to the goal-setting question, of these 24 participants who referenced some aspect of the diagnostic reasoning process, the largest category (50%) entailed the most basic step, *identifying symptoms*.

Performance phase processes

Unlike the pattern of results observed for the forethought phase processes, 90% of students reported that they were focused on *task-specific processes* while they completed the diagnostic reasoning task. More detailed analysis revealed that of the 64 participants who focused on task-specific processes, the largest percentage (59%) reported focusing on *integrating/synthesizing symptoms* followed by 52% of students who focused on *identifying symptoms*. Furthermore, 52% of students reported focusing on multiple task-specific processes. Of those who reported multiple task-specific processes, the most frequently reported responses were categorized as *identifying symptoms* and *identifying contextual factors*.

Correlation and regression results

Table 2 presents the descriptive statistics and Pearson correlations for all the variables. As indicated, strategic planning was statistically significantly correlated with goal setting ($r = .33, P < .01$) and with all performance outcomes: course grade ($r = .40, P < .001$), second-year GPA ($r = .39, P < .01$), USMLE Step 1 score ($r = .29, P < .05$), and NBME score ($r = .33, P < .01$). Types of goals and metacognitive monitoring were not significantly correlated with any of the performance outcomes.

Table 3 presents the results of the two-step, hierarchical multiple regression analyses. Even though undergraduate GPA and MCAT verbal reasoning score were not statistically significantly correlated with any of the performance outcomes, we elected to include them in step one of the regression models given that they were included as part of our *a priori* hypotheses. Further, because we were primarily interested in the ΔR^2 after entering the variables in step two, we present those results here. Strategic planning explained significant variance in course grade ($\Delta R^2 = .15, P < .01$), second-year GPA ($\Delta R^2 = .14, P < .01$), USMLE Step 1 score ($\Delta R^2 = .08, P < .05$), and NBME score ($\Delta R^2 = .10, P < .05$); these effects are considered moderate. In general, students who were focused on several task-specific processes as they approached the diagnostic reasoning task (i.e., during strategic planning) achieved better results on both short and longer-term performance outcomes. The other two microanalytic measures were not related to the performance outcomes.

Table 1. Frequency of Process and Non-Process Responses to Microanalytic Questions: Goal Setting, Strategic Planning, and Metacognitive Monitoring for 71 Second-Year Medical Students Engaged in a Clinical Reasoning Task, Uniformed Services University of the Health Sciences, Academic Year 2010-2011 and 2011-2012.

Response category	Goal setting* <i>n</i> (%)	Strategic planning* <i>n</i> (%)	Metacognitive monitoring* <i>n</i> (%)
Task – specific process	23 (32.4%)	24 (33.8%)	64 (90.1%)
Identifying symptoms	16 (69.6%) [†]	12 (50.0%) [†]	33 (51.6%) [†]
Identifying contextual factors	2 (8.7%) [†]	3 (12.5%) [†]	22 (34.4%) [†]
Prioritizing relevant symptoms	3 (13.0%) [†]	2 (8.3%) [†]	9 (14.1%) [†]
Integrating/synthesizing symptoms	13 (56.5%) [†]	11 (45.8%) [†]	38 (59.4%) [†]
Comparing/contrasting diagnoses	2 (8.7%) [†]	4 (16.7%) [†]	11 (17.2%) [†]
Task – general process	16 (22.5%)	19 (26.8%)	14 (19.7%)
Outcome	13 (18.2%)	N/A [‡]	3 (4.2%)
Self-control	2 (2.8%)	11 (15.5%)	6 (8.5%)
Non-task strategies	6 (8.5%)	11 (15.5%)	N/A [‡]
Perceived ability	N/A [‡]	N/A [‡]	2 (2.8%)
Task difficulty	N/A [‡]	N/A [‡]	3 (4.2%)
Teacher skill	N/A [‡]	N/A [‡]	0
Do not know/none [§]	22 (31%)	1 (1.4%)	N/A [‡]
Other	4 (5.6%)	18 (25.4%)	2 (2.8%)

Column numbers represent the number (*n*) and percentage (%) of the total sample of 71 students who provided a particular response category. The total percentage in each column is greater than 100% because students' responses to a given question could be coded to more than one response category to each microanalytic question.

* For goal setting, 5 students provided responses coded into more than one response category; for strategic planning, 8 students provided responses coded into more than one response category; and for metacognitive monitoring, 23 students provided responses coded into more than one response category.

[†] Column numbers represent the number (*n*) of students who provided a response coded as one of the five key strategies within the task-specific process response category. The percentage (%) is calculated by dividing this number by the sub-set of students who provided a task-specific response. Thus, for goal setting, the denominator is 23; for strategic planning, the denominator is 24; and for metacognitive monitoring, the denominator is 64.

[‡] N/A = Response category was not applicable to the particular microanalytic question.

[§] A response of "do not know/none" indicates that a do not know/none responses was provided without reference to any other response category. If a participant's response included both a do not know/none and a distinct response category, the do not know/other response was ignored.

Table 2. Descriptive Statistics and Pearson Correlations Between the Three Microanalytic Variables, Undergraduate Grade Point Average (GPA), Medical College Admission Test (MCAT) Verbal Reasoning Scores, Course Grade, Second-Year GPA, United States Medical Licensing Examination (USMLE) Step 1, and National Board of Medical Examiners (NBME) Subject Exam in Internal Medicine for 71 Second-Year Medical Students, Uniformed Services University of the Health Sciences, Academic Year 2010-2011 and 2011-2012.

Variables	Mean	Standard deviation	Range	1	2	3	4	5	6	7	8	9
1. Goal setting	0.30	2.25	-2 to 8	--								
2. Strategic planning	0.86	1.74	-3 to 7	.33 [†]	--							
3. Metacognitive monitoring	3.08	2.05	-1 to 8	.01	.03	--						
4. Undergraduate GPA	3.49	0.24	2.98 to 4.00	-.09	.11	.11	--					
5. MCAT verbal reasoning	9.80	1.44	7 to 14	-.07	.02	.04	-.25 [*]	--				
6. Course grade	83.39	5.02	71 to 92	.01	.40 [‡]	.01	.14	-.06	--			
7. Second-year GPA	3.32	0.49	2 to 4	.07	.39 [†]	-.01	.21	-.12	.81 [‡]	--		
8. USMLE Step 1 score [‡]	220.80	18.21	188 to 256	.04	.29 [*]	.10	.26	-.15	.61 [‡]	.83 [‡]	--	
9. NBME score	86.99	7.75	67 to 100	.13	.33 [†]	-.01	.06	-.06	.63 [‡]	.76 [‡]	.77 [‡]	--

MCAT verbal reasoning scores were measured on a 15-point scale, course grades and NBME scores were measured on a 100-point scale, undergraduate and second-year GPA were measured on a 4-point scale, and USMLE Step 1 scores were measured on a scale ranging from 140 to 280.

* $P < .05$.

† $P < .01$.

‡ $P < .001$.

‡ Only students from academic year 2010-2011 had USMLE Step 1 scores ($N = 56$).

Table 3. Hierarchical Multiple Regression Models Using Undergraduate Grade Point Average (GPA), Medical College Admission Test (MCAT) Verbal Reasoning Scores and Microanalytic Strategic Planning Scores to Predict Four Performance Outcomes for Second-Year Medical Students, Uniformed Services University of the Health Sciences, Academic Year 2010-2011 and 2011-2012.

	Performance outcomes																			
	<i>Course grade (N = 70)</i>					<i>Second-year GPA (N = 67)</i>					<i>USMLE Step 1 score (N = 56)[‡]</i>					<i>NBME score (N = 68)</i>				
	<i>B</i>	<i>SE B</i>	β	ΔR^2	R^2	<i>B</i>	<i>SE B</i>	<i>B</i>	ΔR^2	R^2	<i>B</i>	<i>SE B</i>	β	ΔR^2	R^2	<i>B</i>	<i>SE B</i>	<i>B</i>	ΔR^2	R^2
Step 1				.02	.02				.05	.05				.08	.08				.01	.01
Undergraduate GPA	1.83	2.42	.09			.31	.24	.15			15.93	9.44	.22			.25	3.83	.01		
MCAT verbal reasoning score	-1.15	.41	-.04			-.03	.04	-.09			-1.57	1.59	-.13			-.34	.64	-.07		
Step 2				.15 [†]	.17 [†]				.14 [†]	.19 [†]				.08*	.16*				.10 [†]	.11 [†]
Strategic planning score	1.13	.33	.39 [†]			.11	.03	.38 [†]			2.97	1.31	.29*			1.39	.51	.32 [†]		

USMLE = United States Medical Licensing Examination; NBME = National Board of Medical Examiners subject exam in internal medicine.

* $P < .05$.

[†] $P < .01$.

[‡] Only students from academic year 2010-2011 had USMLE Step 1 scores ($N = 56$).

Discussion

This study was important because it represents an initial attempt to examine SRL microanalytic processes during a core physician task (clinical reasoning) and to determine whether these processes were related to performance in medical school. More specifically, we attempted to uncover the strategic quality of second-year medical students' regulatory processes *during* a clinical reasoning task and examined whether these processes were associated with both proximal (e.g., course grade) and distal outcomes (e.g., NBME score). The key finding was that most participants, who can be considered novices or non-experts in diagnostic reasoning, were not highly strategic in terms of their approach (goals, plans) to the diagnostic reasoning task. Interestingly, their strategic plans emerged as a relatively robust predictor of all achievement outcomes. Given that these findings parallel the general premise emanating from the expert-novice literature (i.e., that non-experts typically exhibit rudimentary strategic approaches to learning and performance^{14,15,23,24}) this study may prove useful in stimulating additional research that examines SRL in medical students as they first learn and attempt to master important clinical activities.

To address our first research objective, we descriptively examined the self-regulatory processes of novice learners as they approached and completed a diagnostic reasoning task. Our descriptive analyses indicated that although most participants (approximately 90%) reported focusing on at least one of the key diagnostic processes *during* the reasoning activity, approximately two-thirds were largely non-strategic in how they *approached* the task (i.e., their goal setting and strategic planning). A few noteworthy issues need to be considered with regard to these findings. First, as part of the ICR course, all participants received approximately eight months of instruction in the

diagnostic reasoning process (e.g., creating a problem list, generating a differential diagnosis) immediately preceding this study. Despite this didactic training, only one-third of participants reported plans and goals that specifically pertained to the diagnostic activity and of those participants, the majority reported very rudimentary or simple strategic steps, such as *identifying symptoms*. Very few participants reported using higher-level processes (e.g., *comparing/contrasting diagnoses*; see Table 1).

In contrast to the participants' strategic approach to the clinical reasoning task, approximately 90% of the students conveyed that they were monitoring at least one of the strategic steps of the diagnostic reasoning process *during* task completion. Although these results suggest high levels of strategic thinking by the participants, two factors need to be considered. First, in this study all participants were provided with a post-encounter form (PEF), which explicitly included many of the broad diagnostic or strategic steps needed to solve the case. It is highly probable that this methodological feature may have prompted students to become more aware of these strategic steps and thus biased their responses to the open-ended or free-response metacognitive monitoring question. In addition, similar to the types of strategies reported by participants during goal setting and planning, the majority of their responses to the monitoring question also reflected rudimentary or simple strategic steps and not higher-level strategic thinking (see Table 1).

Broadly speaking, the fact that participants in this study received eight months of didactic training prior to the study—and yet these students still used very rudimentary strategies—seems to underscore the notion that getting medical students to exhibit higher levels of strategic

thinking (e.g., integration) during diagnostic reasoning may require more practice than is often provided in traditional coursework. Examining the developmental trajectory of strategic learning in medical students as it pertains to particular clinical tasks is an important area of future research, particularly because mastering some types of clinical skills may be more time intensive and gradual than others.²⁵

As alluded to previously, our findings are consistent with expert-novice research showing that novices often use very simplistic strategies when approaching and solving new problems.^{14,15,23,24} For example, in a non-medical education context, Cleary and Zimmerman¹⁴ found that non-experts (novice and intermediate basketball players) set fewer specific goals and used less technique-related strategies than experts when practicing their free-throw skills. Although results from medical education are much less definitive—due, in part, to the complex nature of clinical reasoning and the different theoretical frameworks employed in medical education—some findings suggest that, under the right conditions, experts employ more strategy-oriented approaches than novices.²⁶ That is, although experts generally use non-analytic, pattern-recognition techniques when faced with typical case presentations,^{18,19} when confronted with difficult or unusual clinical problems that are not amenable to pattern-recognition methods (and when given sufficient time), experts appear to become much more intentional, reflective, and strategic in their reasoning approach.^{21,27}

In terms of our second research objective, we anticipated that all three regulatory processes (i.e., goal setting, strategic planning, and metacognitive monitoring) would be correlated with important medical education outcomes. These results were mixed, however, as only the strategic

planning measured explained a significant amount of variance in both the short and longer-term performance outcomes. Of interest was that the strategic planning measure moderately correlated with goal setting, yet goal setting did not relate to the longer-term outcomes. Although an adequate explanation for the low correlation between goal setting and the performance outcomes is not entirely clear, it is possible that different types of SRL processes, even when measured similarly and in a contextualized manner, may show distinct predictive power for longer-term performance outcomes. Although highly speculative at this point, it is possible that student responses to the microanalytic strategic planning measure may be indicative of the strategic behaviors that medical students typically display in other learning situations; as a result, these students achieved better outcomes. Addressing the issue of the context-independence versus context-dependence nature of microanalytic questions is a viable avenue for future research.

As for the poor correlation between metacognitive monitoring and the performance, it is possible that the previously mentioned methodological limitations contributed to this problem. That is, exposure to the PEF may have prompted the majority of students to focus on the strategic steps, which unfortunately led to a restriction of range for this variable. In future research using microanalytic protocols to predict performance outcomes, it is recommended that these types of procedural facilitators or prompts not be provided to students as they complete the diagnostic reasoning task.

A few additional limitations in this study are worth noting. First, this investigation was restricted to a convenience sample of predominantly male medical students recruited from one university who were, on average, somewhat older and attained a higher first-year GPA than non-

participants. Furthermore, these students participated in a controlled clinical reasoning activity that lacked the authenticity of an actual clinical encounter. Thus, care should be taken not to over-generalize the results and conclusions drawn in this study to other medical school populations or to other, more authentic clinical reasoning activities. Second, the homogeneity of our sample may have restricted the range of responses across measures and thus negatively impacted our correlation and regression analyses.²⁸ Future research that employs SRL microanalytic methodology should include a more diverse group of novice learners and perhaps a wider range of expertise (e.g., comparing SRL processes in novices versus more experienced clinicians).

In this article, we only examined two of the three phases of SRL (forethought and performance); we did not address how the participants reflected on and adapted their performance on the diagnostic reasoning task. Future research should examine this issue so that medical education researchers can more comprehensively explore the regulatory approaches of medical students on authentic clinical activities. Fourth, other than the undergraduate GPA and MCAT verbal reasoning score, our regression models only included the microanalytic regulatory predictors. Future research should include other measures that are related to medical education outcomes so that the unique contribution of SRL microanalytic processes can be further examined. It is also important to note that the four performance outcomes used in this study were highly correlated, in part because they all largely measured medical knowledge. We suggest that researchers explore how SRL measures relate to performance on medical education exams tapping more applied clinical knowledge and skills, at both a proximal and more distal level. Finally, there are inherent limitations to using microanalytic interview techniques in conjunction with diagnostic

reasoning tasks because such techniques fail to capture non-analytic approaches to clinical diagnosis and because physicians may not be consciously aware of the specific cognitive or behavioral activities that lead to a successful diagnosis.^{26,29}

Notwithstanding the limitations described above, our results suggest that the contextualized measures developed in this study were useful in uncovering students' regulatory processes in relation to a diagnostic reasoning task. Although preliminary in nature, the high levels of inter-rater reliability, moderate relations between forethought processes (goal setting and planning), and the predictive validity of the planning measure all suggest that SRL microanalytic protocols have the potential to serve as a useful adjunct to existing assessment frameworks used in medical education. Furthermore, given that a primary goal of SRL microanalysis is to provide a diagnostic lens for exploring students' thinking and action before, during, and after task performance, future research needs to further establish the validity of this assessment approach and to explore how such information can be used by medical educators in a formative fashion to positively influence training or remediation. Finally, since SRL is conceptualized as a modifiable process and a teachable skill, future research would benefit from intervention studies designed to assess the effectiveness of remediation instruction aimed at enhancing the strategic quality of students' regulatory processes.

Disclosures

Funding/Support: This research was supported by funds from the United States Air Force's Clinical Investigations Program, Office of the Air Force Surgeon General, Washington, DC.

Other disclosures: None.

Ethical approval: This study was approved by the Institutional Review Board of the Uniformed Services University of the Health Sciences, Bethesda, MD.

Contributors: ARA co-conceived of the research design, coded the qualitative data, and wrote the first draft and compiled the final draft of the manuscript. TJC co-conceived of the research design and assisted with the coding rubric and data analysis. TD coded the qualitative data and conducted the data analyses. SJD and PAH assisted with the research design and the coding rubric, and SJD provided access to the study sample. All authors revised the manuscript critically for important intellectual content and edited the final draft for submission.

Disclaimers: The views expressed in this article are those of the authors and do not necessarily reflect the official policy or position of the Uniformed Services University of the Health Sciences, Department of Defense, nor the U.S. Government.

Previous presentations: The abstract of an earlier version of this article was presented at the 2012 annual meeting of the Association for Medical Education in Europe, Lyon, France.

References

1. Hauer KE, Ciccone A, Henzel TR, et al. Remediation of the deficiencies of physicians across the continuum from medical school to practice: A thematic review of the literature. *Acad Med* 2009;84:1822-1832.
2. White CB, Ross PT, Gruppen LD. Remediating students' failed OSCE performance at one school: The effects of self-assessment, reflection and feedback. *Acad Med* 2009;84:651-654.
3. Durning SJ, Cleary TJ, Sandars JE, Hemmer P, Kokotailo PK, Artino AR. Viewing "strugglers" through a different lens: How a self-regulated learning perspective can help medical educators with assessment and remediation. *Acad Med* 2011;86:488-495.
4. Brydges R, Butler D. A reflective analysis of medical education research on self-regulation in learning and practice. *Med Educ* 2012;46:71-79.
5. Graham S, Harris KR. Almost 30 years of writing research: Making sense of it all with The Wrath of Khan. *Learning Disabilities Res Prac* 2009;24:58-68.
6. Kitsantas A, Zimmerman BJ. Comparing self-regulatory processes among novice, non-expert and expert volleyball players: a microanalytic study. *J App Sport Psych* 2002;14:91-105.
7. Zimmerman BJ. Attaining self-regulation: a social cognitive perspective. In: Boekaerts M, Pintrich PR, Zeidner M, eds. *Handbook of Self-Regulation*. San Diego, CA: Academic Press;2000:13-39.
8. Zimmerman BJ. Motivational sources and outcomes of self-regulated learning and performance. In: Zimmerman B, Schunk D, eds. *Handbook of Self-Regulation of Learning and Performance*. New York: Routledge 2011.

9. Zimmerman BJ, Campillo M. Motivating self-regulated problem solvers. In: Davidson JE, Sternberg RJ, eds. *The Nature of Problem Solving*. New York, NY: Cambridge University Press; 2003:239
10. Schunk DH. Social cognitive theory and self-regulated learning. In: Zimmerman BJ, Schunk DH, editors. *Self-Regulated Learning and Academic Achievement*. 2nd edn. Mahwah, NJ: Lawrence Erlbaum Associates Publishers 2001:125–151.
11. Cleary TJ. Emergence of self-regulated learning microanalysis: Historical overview, essential features, and implications for research and practice. In: Zimmerman B, Schunk D, eds. *Handbook of Self-Regulation of Learning and Performance*. New York: Routledge 2011.
12. Cleary TJ, Callan GL, Zimmerman BJ. Assessing self-regulation as a cyclical, context-specific phenomenon: Overview and analysis of SRL microanalysis protocol. *Ed Res International* 2012;doi:10.1155/2012/428639.
13. Winne PH, Perry, NE. Measuring self-regulated learning. In: Boekaerts M, Pintrich PR, Zeidner M, eds. *Handbook of Self-Regulation*. San Diego, CA: Academic Press;2000:532-568.
14. Cleary TJ, Zimmerman BJ. Self-regulation differences during athletic performance by experts, non-experts and novices. *J App Sport Psych* 2001;13:61-82.
15. DiBenedetto MK, Zimmerman BJ. Differences in self-regulatory processes among students studying science: A microanalytic investigation. *Intl J Educ Psych Assess* 2010;5:2-24.
16. Cleary TJ, Sandars J. Self-regulatory skills and clinical performance: A pilot study. *Med Teach* 2011;33:e368-e374.
17. Elstein AS. Thinking about diagnostic thinking: A 30-year perspective. *Adv in Health Sci Educ Theory Pract* 2009;14:S7-S18.

18. Eva KW. What every teacher needs to know about clinical reasoning. *Med Educ* 2004;39:98–106.
19. Norman G. Research in clinical reasoning: Past history and current trends. *Med Educ* 2005;39:418–427.
20. Durning SJ, Artino AR, Boulet JR, La Rochelle JS, van der Vleuten CPM, Arze B, Schuwirth LWT. The feasibility, reliability, and validity of a post-encounter form for evaluating clinical reasoning. *Med Teach* 2012;34:30-37.
21. Durning SJ, Artino AR, Boulet JR, Dorrance K, van der Vleuten CPM, Schuwirth LWT. The impact of selected contextual factors on internal medicine experts' diagnostic and therapeutic reasoning: Does context impact clinical reasoning performance in experts? *Adv in Health Sci Educ Theory Pract* 2012;17:65-79.
22. Cleary TJ, Callan GL, Peterson J, Adams TA. Using SRL Microanalysis in an academic context: Conceptual and empirical advantages. Paper presented at the annual conference for the American Educational Research Association, New Orleans, LA 2011.
23. Bransford JD, Brown AL, Cocking RR, eds. *How People Learn: Brain, Mind, Experience, and School*. Washington, DC: National Academy Press 2000;51–78.
24. Ericsson KA. Deliberate practice and acquisition of expert performance: A general overview. *Acad Emerg Med* 2008;15:988-994.
25. Schuwirth LWT. Is assessment of clinical reasoning still the Holy Grail? *Med Educ* 2009;43:298-299.
26. Byrne A. Mental workload as a key factor in clinical decision making. *Adv in Health Sci Educ Theory Pract*. 2012;doi:10.1007/s10459-012-9360-5.

27. Mamede S, Schmidt HG, Rikers RMJP, Penaforte JC, Coelho-Filho JM. Breaking down automaticity: Case ambiguity and the shift to reflective approaches in clinical reasoning. *Med Educ* 2007;41:1185–1192.
28. Cohen J, Cohen P, West SG, Aiken LS. *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences*, 3rd edn. Mahwah, NJ: Lawrence Erlbaum Associates 2003.
29. Eva KW, Brooks LR, Norman GR. Forward Reasoning as a Hallmark of Expertise in Medicine: Logical, Psychological, and Phenomenological Inconsistencies. In: Shohov, SP, eds. *Advances in Psychological Research*, Vol. 8. New York: Nova Scotia 2002:42.

Appendix (Supplemental Digital Content)

Scoring Scheme

This scoring system was developed before data analysis and was an adaptation of a prior scoring scheme.²² Using the goal setting measure as an example, positive points were awarded for responses pertaining to specific task process responses or other adaptive responses such as *self-control tactics* used to complete the task. Negative points were applied for responses deemed to be maladaptive, such as *do not know/none* or *non-task strategies*. In addition to the direction of points (positive or negative), there was variation in the amount of points awarded for a given response. To illustrate, one point was awarded for a relatively simplistic diagnostic reasoning step of *identifying symptoms*, two points were awarded for *prioritization of relevant symptoms*, and three points were awarded for *integrating/synthesizing symptoms* and *comparing/contrasting diagnoses*. The rationale for developing a scoring scheme in this fashion was to award points based on the skill needed to successfully perform a given step of the process

Scoring Template

Goal setting

a) Process

1. Identify symptoms – add 1 point
2. Identify contextual/environmental factors – add 1 point
3. Identify or prioritize relevant/pertinent symptoms or features – add 2 points
4. Integration – add 3 points
5. Comparing/contrasting diagnoses – add 3 points
6. General/Other – add 0 points

b) **Volition (self-control)** – add 1 point

c) **Outcome** – add 0 points

d) **Non-task strategies** – minus 1 point (-1)

e) **Confidence/ability** – add 0 points

f) **Do not know/none** – minus 2 points (-2) ****Cannot have this response plus another category****

g) **Other** – not scored

Strategic planning

a) Process

1. Identify symptoms – add 1 point
2. Identify contextual/environmental factors – add 1 point
3. Identify or prioritize relevant/pertinent symptoms or features – add 2 point
4. Integration – add 3 points
5. Comparing/contrasting diagnoses – add 3 points
6. General/Other – add 0 points

b) **Volition (self-control)** – 1 point

c) **Non-task strategies** – minus 1 point (-1)

d) **Do not know/none** – minus 2 points (-2)

e) **Other** – not scored

Metacognitive monitoring**a) Process**

1. Identify symptoms – add 1 point
2. Identify contextual/environmental factors – add 1 point
3. Identify or prioritize relevant/pertinent symptoms or features – add 2 point
4. Integration – add 3 points
5. Comparing/contrasting diagnoses – add 3 points
6. General/Other – add 0 points

b) Volition (self-control) – add 1 point

c) Task difficulty – minus 1 point (-1)

d) Outcome – add 0 points

e) Ability skills – minus 2 points (-2)

f) Perceived knowledge – minus 2 points (-2)

g) Teacher skill – minus 1 point (-1)

h) Do not know/none – minus 2 points (-2)

i) Other – not scored