

Metacognitive strategies and test performance: an experience sampling analysis of students' learning behavior

Ulrike E. Nett, Thomas Goetz, Nathan C. Hall, Anne C. Frenzel

Angaben zur Veröffentlichung / Publication details:

Nett, Ulrike E., Thomas Goetz, Nathan C. Hall, and Anne C. Frenzel. 2012. "Metacognitive strategies and test performance: an experience sampling analysis of students' learning behavior." *Education Research International* 2012: 958319.
<https://doi.org/10.1155/2012/958319>.

Research Article

Metacognitive Strategies and Test Performance: An Experience Sampling Analysis of Students' Learning Behavior

Ulrike E. Nett,¹ Thomas Goetz,^{2,3} Nathan C. Hall,⁴ and Anne C. Frenzel⁵

¹Ulm University, Ulm, Germany

²University of Konstanz, Konstanz, Germany

³Thurgau University of Teacher Education, Kreuzlingen, Switzerland

⁴McGill University, Montreal, QC, Canada

⁵University of Munich, Munich, Germany

Correspondence should be addressed to Ulrike E. Nett, ulrike.nett@uni-ulm.de

Received 11 November 2011; Revised 16 April 2012; Accepted 2 May 2012

Academic Editor: Mariel F. Musso

Copyright © 2012 Ulrike E. Nett et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The aim of the present study was to explore students' learning-related cognitions prior to an in-class achievement test, with a focus on metacognitive strategy use. A sample of 70 students in grade 11 (58.6% female, $M_{\text{age}} = 17.09$ years) completed a series of structured, state-based measures over a two-week period via the experience sampling method until the day before a class test. Results illustrated students' self-regulatory ability to preserve their motivational and cognitive resources, with test-related cognitions evidenced significantly more often in learning-related as opposed to leisure settings. Metacognitive strategy use was also found to significantly increase as the test date approached underscoring the goal-oriented nature of situated learning behaviors. Higher intercepts and increases in frequency of test-related cognitions over time positively corresponded to test performance. Of the three metacognitive strategies assessed, monitoring was found to positively correspond with test performance. Implications for future practice as well as implications for future research employing the experience sampling method are discussed.

1. Introduction

Metacognitive competencies have over the past three decades developed from a largely neglected issue to one of the most elaborated areas of theory and research in the educational sciences [1, 2]. Since the pioneering work of Flavell [3], the concept of metacognition has become inextricably linked with learning theories [4] and nearly equated with the construct of self-regulated learning in terms of planning, monitoring, and evaluation of learning and problem solving [5, 6]. This immense growth in interest is largely due to an emerging consensus among policy makers, teachers, educators, parents, and researchers concerning the importance of fostering students' ability to autonomously direct their learning processes [5]. In light of the present societal emphasis on lifelong learning and economic climate requiring individuals to rapidly acquire new employment skills, this increase in theoretical and empirical interest in individuals' metacognitive ability to independently and efficiently regulate their learning is likely to continue.

Despite considerable research on metacognitive strategies, several questions warrant further investigation. More specifically, how often and when are metacognitive strategies applied in a learning process (e.g., when is it best to start to prepare for a test)? Which metacognitive learning strategies are most commonly employed in actual learning situations and most effective with respect to academic performance? Whereas previous research has consistently evaluated the relative utility of metacognitive strategies, it is also important to investigate the frequency and effectiveness of metacognitive strategy use as evaluated during the actual learning process [7]. Further, despite clear theoretical assumptions concerning the achievement benefits of the ability to regulate one's own learning, empirical findings are more mixed in showing strategy-achievement relations [8]. In attempting to explain inconsistent findings, one explanation is that strategy-achievement relations may vary as a function of situational factors [8]. In some academic settings, for instance, it may not always be necessary to use metacognitive strategies (e.g., if the teacher provides strict instructions for task

completion). Further, the lack of objective, reliable, and valid measures of metacognitive strategy use is often cited as a possible confound [9]. Until recently, various measures have been employed to assess metacognitive strategy use in the self-regulation research literature, many of which are limited by their inability to account for situational contingencies or do so in an intrusive, inefficient, or short-term manner [9].

The objective of the present study was to evaluate the relation between students' use of metacognitive strategies as assessed during the actual learning process preceding an achievement test. To this end, the experience sampling method (ESM; [10, 11]) was employed to assess metacognitive strategy use in a highly valid and objective manner. Students were provided with handheld devices that signaled several times throughout the day, after which students were presented a short questionnaire about their learning behavior to be immediately completed. In addition to allowing for an in-depth assessment of the most effective metacognitive strategies, this longitudinal, *in vivo* method also permits an analysis of the developmental nature of the self-regulated learning process as well as the extent to which students' use of specific strategies is indeed contingent upon real-life situational factors. The present research thus aims to contribute to existing research on metacognition and self-regulated learning through the use of experience sampling methods that allow for critical research questions concerning the frequency and utility of metacognitive strategies to be evaluated, and further, that address a critical limitation of global measures by measuring situation-specific strategy use.

1.1. Metacognitive Strategies. Put simply, the term *metacognition* refers to "thinking about thinking" [12]. In addition to cognitive and motivational components, metacognition is consistently represented in established theoretical models as an equally important element of successful self-regulated learning (e.g., [7, 13–15]). In the three-layered model of self-regulated learning proposed by Boekaerts [5], for example, the regulation of the learning process through the use of metacognitive strategies is presented as the middle layer of self-regulatory focus juxtaposed between the more basic layer involving the regulation of cognitive strategies during specific learning tasks (i.e., processing modes) on the one hand, and the higher-order layer pertaining more globally to the regulation of one's self with respect to one's goals and motivational resources.

1.1.1. Components of Metacognitive Strategies. Although a number of metacognitive regulatory strategies has been examined [16], there exists a general consensus among researchers that three metacognitive strategies are most critical for regulating the learning process, namely, *planning*, *monitoring*, and *evaluation* [5, 17, 18]. The strategy of planning includes the setting of goals, selecting adequate cognitive strategies to achieve this goal, as well as the allocation of personal resources such as effort or time. Monitoring refers to being aware of one's comprehension and task performance—monitoring one's learning and continually comparing one's current state of learning to one's learning

or achievement goals. In response to monitoring feedback, evaluation involves the correction of learning problems and adjustments in learning plans [5, 16]. For example, an evaluation of unsatisfactory progress occurs when, as a result of monitoring one's learning strategy use and comprehension, an individual concludes that they will not achieve their learning goals by persisting in the use of the present learning behaviors. Concerning the functional overlap between these core metacognitive strategies, questionnaire findings consistently show these three strategies to be statistically interdependent and not hierarchically or linearly related [19, 20]. Nevertheless, the strategies of planning, monitoring, and evaluation can be clearly differentiated from a theoretical perspective in that the observed empirical overlap between them may be due to temporal as opposed to functional similarities. More specifically, as suggested in a model of self-regulated learning proposed by Schmitz [21], these strategies are assumed to be utilized in a cyclical manner, and thus, are likely to co-occur during the learning process due to earlier strategies (e.g., planning) prompting later strategies (e.g., monitoring) that recursively influence the preceding strategy (e.g., modified plans) through feedback loops.

Given the above findings concerning the co-occurrence of these three strategies, particularly with respect to monitoring, this metacognitive strategy has been a focus of particular empirical interest in the context of self-regulated learning. For example, incorporated into the theoretical model of Winne and Hadwin [7] is an omnipresent metacognitive monitoring process responsible for generating internal feedback during any phase of the self-regulated learning process. As further evidenced by monitoring having been explored in considerable experimental learning research (e.g., [22–24]), this metacognitive strategy is often regarded as a key self-regulated learning process critical to developing an understanding of complex information [25].

1.1.2. Situational Nature of Self-Regulated Learning. According to Winne and Hadwin [7], self-regulated learning incorporates both students' dispositions (aptitude) as well as situational factors (learning events). The term "aptitude" describes relatively stable dispositional features that have an impact on self-regulated learning, such as differences in goals or the ability to use metacognitive strategies [9]. In contrast, the term "event" refers to aspects of one's learning environment that influence the learning process, such as the time remaining prior to an achievement test or optimal opportunities to use specific metacognitive strategies [26]. Considering that the learning process can be viewed as sequence of self-regulatory states [21], it is thus important to not only examine the frequency and effectiveness of specific metacognitive strategies, but also when these strategies are used and are most effective [13].

1.2. Metacognitive Strategies and Academic Performance. Regulating one's own learning through the use of metacognitive strategies is widely considered to be the most elaborated form of learning (e.g., [6, 27]) and is often assumed to be one of the strongest predictors of optimal learning [28]. In educational research, it is typically hypothesized that

students' self-regulatory efforts toward their learning goals should have a direct impact on subsequent achievement, and further, mediate linkages between dispositional factors such as cognitive abilities or contextual characteristics such as the classroom environment and achievement [29]. In contrast to theoretical assumptions, empirical findings concerning the relationship between the use of metacognitive strategies and academic achievement are inconsistent [30].

Some correlation studies show the use of metacognitive strategies to positively correspond with academic achievement (e.g., [20, 31, 32]). Meta-analytic reviews of intervention studies also suggest that the promotion of self-related learning strategies typically has a positive impact on learning behavior and academic achievement, particularly when metacognitive awareness is encouraged [33–35]. Conversely, other correlational empirical studies reveal notably weak or no significant relationships between the use of metacognitive strategies and academic achievement [36, 37] or show the relation between metacognition and performance to be entirely mediated by other factors (e.g., self-efficacy; [38]). Veenman and van Hout-Wolters [39] (as cited in [8]) showed in a review of 21 questionnaire studies and more than 7000 participants that, on average, no more than 3% of the variance in students' achievement was explained by the use of metacognitive strategies. Additional research further suggests that students tend to not demonstrate or prematurely terminate metacognitive strategy use in learning settings without sufficient instructional support or prompting [40, 41]. Such heterogeneous findings concerning the relationship between metacognitive strategies and academic achievement are likely due to multiple intervening factors such as situational factors (e.g., lack of opportunity afforded by task characteristics) or measurement error (i.e., retrospective biases in self-report measures versus indicators of real-life strategy use).

Concerning the moderating effect of academic environments on self-regulated learning, it is assumed that certain academic settings afford more limited opportunities for engaging in self-regulation than do others [26]. For example, whereas higher education affords various opportunities to choose subject domains, modify or supplement learning tasks, and pursue independent study, fewer options for self-regulated learning are available to secondary school students. In a similar vein, the extent to which metacognitive strategies are in fact required for optimal learning and performance may also vary as a function of the specific learning situation (e.g., in-class versus homework activities; [13]). Given the impact of the educational context on the potential for self-regulated learning, it has been further hypothesized that retrospective, questionnaire-based measures lack sufficient validity in assessing students' real-life use of metacognitive strategies that, to a large extent, are contingent upon the opportunities for, and advantages of, metacognitive strategies afforded by specific learning tasks and situations [7, 17, 36].

1.3. Measuring Metacognitive Strategies. At present, there exist two major approaches toward the assessment of metacognitive strategy use involving either externally observed measures of behavioral or physiological indicators,

or self-report methods. Whereas behavioral or physiological indicators are highly reliable and less subject to participant bias (e.g., [42]), they are disadvantaged by the inherently subjective nature of metacognitive processes in that the individual has the most direct and reliable access to cognitive experiences [11]. In addition, the use of physiological or behavioral measures relying on evaluations by others may lead to misinterpretation or experimenter bias. As such, despite the aforementioned limitations of self-report measures, such instruments are much more commonly employed to assess metacognitive strategy use as reflected in reviews of research on the assessment of self-regulatory strategies [8, 9, 17, 29].

Among the most popular methods for assessing students' use of metacognitive strategies are questionnaires (e.g., Motivated Strategies for Learning Questionnaire, MSLQ; [20]; Learning and Study Strategies Inventory, LASSI; [43]) and interviews (e.g., Learning Interview Schedule, SRLIS; [31, 32]). Although these self-report measures provide efficient and reliable evaluations of dispositional aspects of metacognitive strategy use [17], questionnaire methods in particular appear to have very little predictive validity with respect to achievement outcomes [8] as well as limited construct validity [44].

In contrast to the above methods, think-aloud protocols (e.g., [45]) allow for the actual process of self-regulated learning to be observed while it occurs. However, a critical limitation of this method is that it permits only short-term problem-solving processes to be assessed (e.g., during a specific learning task) as opposed to more long-term learning processes (e.g., during the week preceding an exam). Although learning diaries do provide a better opportunity to assess students' long-term use of metacognitive strategies [21], such findings are often confounded by the diary itself serving as an intervention encouraging greater structured reflection on one's learning activities (e.g., [41, 46, 47]). Furthermore, just as the effectiveness of open-ended diary assessments is largely dependent on the learners' writing ability as well compliance with the writing protocol, the utility of structured learning diaries is also compromised due to the typical inclusion of self-report questionnaires that involve the weaknesses of questionnaires in general may serve to prompt further reflection on the self-regulatory strategies assessed.

Taken together, the limitations of previous assessment methods suggest that more ecologically and empirically valid methods are required to better evaluate metacognitive strategy use as it naturally occurs in real-life learning situations, thereby contributing to a more dynamic, differentiated, and ecologically valid understanding of the nature of self-regulated learning [48]. To address this research gap, a naturalistic assessment method, referred to as the experience sampling method (ESM; [10, 11]), appears promising and has been successfully employed in related self-regulation research (e.g., coping with stress; [49]). In ESM protocols, participants complete a short questionnaire concerning their learning behavior at randomly selected times over a period of days via a personal digital assistant (PDA). This method thus incorporates some of the advantages of previous methods,

such as the utilization of reliable items from established questionnaires, while avoiding their disadvantages by allowing for situational factors to be assessed in a nonintrusive manner and providing a highly elaborated yet longitudinal data structure. Thus, whereas this method incorporates some of the advantages of learning diaries, they are less intrusive due to their reduced length as well as more candid in soliciting responses at unpredictable intervals throughout the study duration. As such, it is assumed that experience sampling methods provide a less interventional and more ecologically valid means of evaluating learning processes than the self-report methods typically employed in this domain.

1.4. The Present Study. In adapting ESM procedures to the study of metacognitive strategy use, the present study aimed to provide ecologically as well as empirically valid data in support of the effectiveness of metacognitive strategies used by students during the weeks prior to an achievement test. As such, this research aims to contribute to the self-regulation literature by addressing significant limitations of both self-report (questionnaire, interview, think-aloud) and more objective assessment methods (behavioral, physiological). More specifically, the use of ESM protocols allowed for research questions concerning the frequency, timing, as well as achievement benefits of metacognitive strategy use to be evaluated as assessed during the weeks prior to a class test. Consistent with the assertion that students' use of self-regulated learning strategies be evaluated as an ensemble comprising more than just the sum of its parts [50], our research questions further addressed both students' global test-related cognitions as well as the discrete metacognitive strategies of planning, monitoring, and evaluation [5].

2. Research Questions and Hypotheses

The present study aimed to explore students' real-life test-related cognitions (more specifically if and when they thought about the test) and use of metacognitive strategies prior to an achievement test. In so doing, the present study focused on the learning process with respect to changes over time in the frequency of test-related cognitions in general, the involvement of metacognitive strategies, as well as predictive relations between test-related thoughts or specific strategies and performance. Through the use of the experience sampling method [10, 11], students' test-related thoughts and reported metacognitive strategy use was recorded over 14 days prior to an important test in a mathematics course. Considering the goal-oriented nature of self-regulated learning [29], the present research questions were evaluated in a specific academic domain, a mathematics course, to provide for a more accurate, domain-specific analysis of metacognitive strategies employed toward specific learning goals (cf., [51, 52]). Mathematics was selected due to it being a core academic subject worldwide that is typically associated with high scientific and societal value. Students' grades on mathematics tests completed prior to and following the ESM protocol were also obtained. The first step in our analyses evaluated students' global test-related cognitions (if they thought about the test) and their effects on test

TABLE 1: Research Hypotheses.

	Hypothesis 1	Hypothesis 2
	Frequency and effects of test-related cognitions	Frequency and effects of metacognitive strategies
a	Test-related cognitions are linked to learning-related situations and are observed more often as the test date approaches.	Metacognitive strategies are interrelated and used more often as the test date approaches.
b	Test-related cognitions correspond to performance improvement.	Metacognitive strategies correspond to performance improvement.

performance. In a second step, we evaluated the frequency and performance benefits of discrete metacognitive strategies reported by students when thinking about the test, with a specific focus on planning, monitoring, and evaluation as critical self-regulatory processes. An overview of the study hypotheses is presented in Table 1.

2.1. Research Question 1. How often and when do students occupy themselves with thoughts concerning an upcoming test in mathematics during the 14 days preceding the test, and is this cognitive engagement related to their test performance?

2.1.1. Hypothesis 1a. Based on the assumption that optimal self-regulated learning involves the ability to disengage from achievement-related cognitions in situations not consistent with learning and achievement, we anticipate that students will more frequently report test-related cognitions during learning-related situations (e.g., mathematics or learning situations) than during their leisure time. Furthermore, considering the goal-oriented nature of learning behaviors, we expect students to think more often about the test as the test date approaches.

2.1.2. Hypothesis 1b. We expect that the frequency of global test-related cognitions will correspond to performance improvements on the subsequent test based on the assumption that such cognitions are associated with the use of metacognitive strategies.

2.2. Research Question 2. When occupied with thoughts about the upcoming test, which metacognitive strategies do students report, how does the use of these metacognitive strategies develop as the test date approaches, and to what extent are specific strategies related with test performance?

2.2.1. Hypothesis 2a. We anticipate that students' global test-related cognitions will positively correspond with each of the metacognitive strategies of planning, monitoring, and regulation. Furthermore, strong positive correlations are expected between the three strategies, and each metacognitive strategy is expected to demonstrate a increasing growth curve over time consistent with the anticipated curve for global test-related cognitions.

2.2.2. Hypothesis 2b. We hypothesize that students' use of each metacognitive strategy will be related to subsequent improvements on the next class test. Monitoring is expected to be most positively related to test improvement as suggested by the aforementioned research in which this specific strategy is examined.

3. Method

3.1. Participants and Data Collection. Data was collected through the use of the experience sampling method [10, 11] involving personal digital assistant (PDA) devices programmed with PMat software [53]. Participants were students from altogether 22 classes of grade 11 students in Germany. Between two and four students from each class were randomly selected and provided with PDA devices, with each device set to provide alerts based on a combination of event-based and randomized signaling [11]. Thus, a total of 70 students completed the experience sampling study protocol, 41 (58.6%) were female, and the average age was 17.09 years ($SD = .55$ years). Event-based signaling was conducted by instructing participants to register their device at the beginning of mathematics, German, or English classes while in school. The device would then signal at a randomly selected time within the next 40 minutes (each class lasting 45 minutes) and provide the participant with a digital questionnaire to be immediately completed. Teachers were informed of this procedure and had previously consented to allow students to participate. In addition, the randomized signaling protocol required the device to signal at three randomly selected times between 2 PM and 10 PM. On weekends, 6 signals were provided at randomly chosen moments between 10 AM and 10 PM. Before starting the questionnaire, participants reported the activity in which they were engaged at that moment, namely, whether they were attending a class in mathematics or another subject, completing homework, engaged in learning behavior (e.g., reviewing class materials), engaged in leisure activities, or another activity. The questionnaire was accessible for only 5 minutes, after which it was no longer available. A minimum two hour lag was introduced between adjacent signals, and students completed the questionnaire items each day over a period of 14 days prior to their next test in the mathematics class.

Participation in all parts of the study was voluntary and all responses were anonymous. In addition to the ESM data collection, achievement data was collected consisting of students' self-reported grades on their most recent test as well as actual grades on the upcoming test for which they were currently preparing as obtained from the mathematics teachers.

3.2. Study Variables

3.2.1. Achievement in Mathematics. Students' self-reported grades on their most recent mathematics test, as well as their actual grades on the subsequent test, were assessed as measures of academic achievement. In the German school system, grades range from 1 (very good) to 6 (failed).

Students' grades were inverted prior to analysis allowing for higher scores to reflect better academic achievement. The first test in mathematics ($M = 2.94$, $SD = 1.27$) and the second test ($M = 2.87$, $SD = 1.26$) were significantly correlated at $r = .54$ ($P < .01$). To obtain a measure of relative improvement (or relative decline) in mathematics achievement, z-standardized residuals on second mathematics test excluding variance accounted for by the first test were calculated.

It is important to note that by utilizing residual as opposed to difference or gain scores, it is not the improvement in terms of a difference in average change between the two tests that is measured, but rather the extent of improvement on the second test relative to the first test controlling for initial levels on the first test. In other words, residual scores allow for analysis of levels of the second test over and above what can be predicted by first test. As such, the use of residual scores allowed for a more sensitive analysis of changes in grades relative to prior achievement, as opposed to changes in raw achievement scores over time.

3.2.2. State Assessment. To avoid having students complete overly long state-based questionnaires, the test-related cognition and metacognition constructs were assessed using single-item measures. This practice is consistent with similar ESM research on academic emotions [54, 55]. Students responded to the following item concerning global test-related cognitions: "*During the past hour, I engaged in thinking about the mathematics test*" (0 = not at all, 4 = all the time). Students who responded with values 1 through 4, indicating some level of test-related thoughts, were further asked to respond to the following metacognitive strategy items having the same response options: "*During the past hour, I thought about ... what to learn for the mathematics test*" (planning), "*... what I already know for the mathematics test*" (monitoring), "*... whether to change my way of studying for the mathematics test*" (evaluation). These self-report items were adapted from the metacognitive strategies subscale of the MSLQ [20] and were employed due to the two-step process involved wherein responses to the same initial question concerning mathematics-specific cognition were followed by more specific probing questions in which different metacognitive strategies were explicitly assessed.

The contingent assessment of the three strategy questions following the initial test-related cognition item was important in order to more innocuously assess their use among students who were already thinking about the test, as opposed to unintentionally highlighting these strategies and encouraging their use by having students first consider the specific strategies and then reflect on their relevance to their upcoming test. Nonetheless, this two-step process still allowed for a differentiated assessment of metacognitive strategies in that by having highly similar wordings in the second set of questions in which only the strategy type differed, students who were already reflecting on their upcoming test were assumed to be better able to recognize and report on these more specific elements of test-related cognitions. In light of global test-related cognitions being a requisite for the

subsequent reporting of metacognitive strategies, students who indicated no level of thought about the upcoming exam were provided a score of 0 (*not at all*) for each strategy item. Although these three metacognitive strategies are often assumed to occur in a strict temporal order, the three strategies were assessed concurrently due to their potentially overlapping, fast-moving, and cyclical nature [21] all of which was anticipated to make it more difficult to differentiate their specific sequencing, as is postulated in self-regulation theories, in a real-life achievement setting (cf., [19, 20]).

3.3. Statistical Analysis. The study data obtained reflects a two-level structure consisting of points of assessment (Level 1; $N = 2654$ single measurement points, 733 if aggregated over one day) within persons (Level 2; $N = 70$).¹ The analyses presented were conducted via hierarchical linear modeling using HLM 6.07 software [56] allowing for the two-level structure of the data to be accounted for in correlational as well as temporal change analyses. Based on the missing at random assumption (MAR), HLM applies the “full information maximum likelihood” method (FIML; [57]) allowing all Level 1 cases to be assessed even if they include missing data. All reported study findings are thus based on the entire study sample, with HLM allowing for parameter estimation and handling of missing data to be integrated into a single step so as to provide appropriate parameter estimates and standard errors (e.g., [58]). However, as 11 students did not indicate their results on the mathematics test, only 59 students were assessed when this Level 2 variable was included in the analysis. In sum, HLM provides substantial advantages over traditional change models (such as ANOVA) by allowing for modelling of nonlinear growth, evaluating predictors of growth parameters, and also accounting for missing data at individual points of measurement [59].

4. Results

To evaluate the extent to which the experience sampling method interfered with the learning process, the participants’ achievement scores on both tests in mathematics were compared to the scores of other students from the same classes ($N = 382$) who did not take part in the study. No significant difference on either test between the experiences sampling group ($M_1 = 2.94$, $SD_1 = 1.27$; $M_2 = 2.87$, $SD_2 = 1.26$) and control group ($M_1 = 2.82$, $SD_1 = 1.33$; $M_2 = 2.71$, $SD_2 = 1.17$) was observed, with all P values being above .14.

4.1. Hypothesis 1a: Frequency of Test-Related Cognitions. Some level of thought concerning the upcoming mathematics test over the previous hour was reported by students during 23% of the assessment periods. Results also supported our hypothesis that test-related cognitions would most frequently occur in learning situations, with students found to report thinking about the test, at least once, 56% of the time in mathematic classes, 15% of the time in other classes,

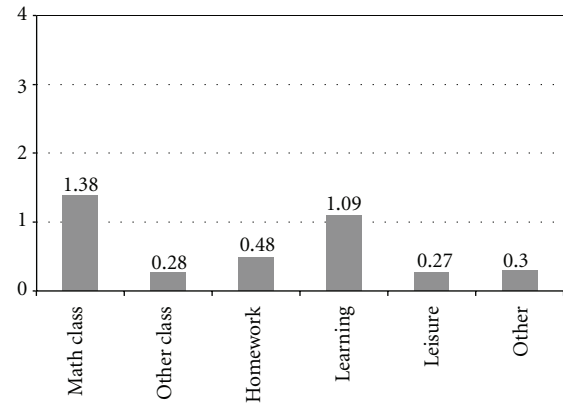


FIGURE 1: Average intensity of test-related cognitions.

24% of the time while completing homework for all subjects, and 40% of the time spent learning for all subjects (e.g., reviewing class materials, preparing for a test at home), as compared to during 14% of leisure time and 15% of the time spent on additional activities. Figure 1 presents the average intensity of test-related cognitions *over the preceding hour* across assessment points for each response option from 0 (*not at all*) to 4 (*all the time*). These response averages more specifically suggest that students think relatively more often about the upcoming test in learning or achievement-related situations (e.g., mathematics or other classes) as opposed to nonachievement settings (e.g., leisure time) over the preceding hour, and further, that some degree of thought concerning the test is evidenced in each of the domains assessed.

Concerning the development of students’ occupation with thoughts about the upcoming test, we calculated the average frequency of students’ test-related cognitions across all assessments administered during a given day. As the questionnaires were completed at randomized intervals, we assumed that the average of all measures completed during each of 14 days prior to the mathematics test should provide a good estimate of the relative intensity of students’ test-related thoughts during that day. To assess change over time in test-related cognitions, each student’s development was represented by an individual polynomial growth trajectory based on a unique set of parameters, as reflected by a polynomial Level 1. By subsequently adding polynomial parameters of a higher order (linear, quadratic, cubic) until the beta weight for the fixed parameter of the highest polynomial predictor was not significant, the most accurate shape for the polynomial growth curve was identified. If the beta weight of the highest polynomial order predictor was not significant, this indicated that no additional within-person variance could be explained by increasing the polynomial degree of the growth function and this parameter was excluded (cf., [59]).

With respect to the manner in which time was coded [60, 61], the present analysis coded time such that the final day before the test was assigned a value of zero and thus evaluated as the reference category (linear change parameter scoring $-13, -12, \dots, -2, -1, 0$; quadratic values scoring

TABLE 2: Analysis of change in test-related cognitions prior to test completion.

	Basic model		Linear model		Quadratic model		MG + quadratic model	
	<i>b</i>	SE	<i>b</i>	SE	<i>b</i>	SE	<i>b</i>	SE
Fixed effects								
Intercept (I)	.52	.05	.87**	.10	1.19**	.15	1.19**	.15
<i>MG</i> × <i>I</i>							.19*	.08
Linear Slope (LS)			.05**	.01	.19**	.04	.19**	.04
<i>MG</i> × <i>LS</i>							.02*	.01
Quadratic Slope (QS)					.01**	.002	.01**	.002
	Var		Var		Var		Var	
Random effects								
Between: I	.100**		.405**		.776**		.695**	
Between: LS			.003**		.033*		.031*	
Between: QS					.000		.000	
Within	.577 ^a		.487 ^a		.443 ^a		.443 ^a	
Model statistics								
Deviance	1503.04		1420.01		1381.23		1377.15	
No. parameters	3		6		10		12	

⁺ $P < .1$; * $P < .05$; ** $P < .01$; ^ano significance test available.

Note: Intercept represents the values one day before the test. State measures across one day were aggregated. $N_{\text{level } 1} = 733$; $N_{\text{level } 2} = 59$; *MG* = *z*-standardized residuals on second mathematics test excluding variance accounted for by first test.

169, 144, ..., 4, 1, 0). Each growth parameter was entered as both fixed and random effects, based on the assumption that individual patterns of change may be observed (Table 2). The null model was evaluated first to assess variability in test-related cognition values across all points of assessment according to its within-individual and between-individual components [61].

Results showed the majority of the variability to reside within individuals (85%) as would be expected due to the situation-dependent, state-based measures employed. Nonetheless, a significant amount of variability was also found to occur between individuals (15%) showing that in addition to substantial change in test-related cognitions over time, a trait-based conceptualization of the degree to which students engage in test-related thinking is warranted (Table 2, Column 1). Findings also revealed a nonsignificant beta weight for the fixed parameter of the cubic predictor, suggesting that the significant quadratic curve provided the best polynomial model for the present data (cf., [59]).

Based on the manner in which time was coded, the growth parameters are interpretable in the following manner (Table 2, Column 3). First, the intercept represents the mean frequency of test-related thoughts one day before the test (1.19 on a scale of 0 to 4). Second, the coefficient of the linear change reflects the slope of the curvilinear trajectories in test-related thinking one day before the test, indicating that the frequency of test-related cognitions increased by .19 on the final day. Third, the coefficient of quadratic change indicates the curvature of this quadratic function at any given point in time (see [62, 63]). The observed quadratic value of .01 suggests that change over time in test-related thoughts

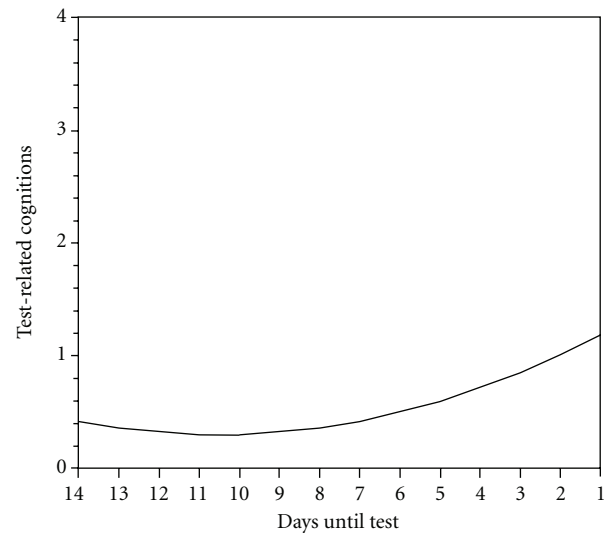


FIGURE 2: Average intensity of test-related cognitions.

followed a positive and notably wide parabolic curve. This finding can be interpreted as showing students' test-related cognitions to significantly increase quadratically in frequency as the test date approached (Figure 2).

The random effects assessed in the model provide information as to whether the shape and position of the parabola varied between participants. Results showed significant variability in the intercept (variance of .415 one day before the test) and in the linear component of the trajectory (variance of .034 one day before the test). In contrast, the variability of the quadratic component of the trajectory was

not significant, implying substantial variability in the relative position of the growth curve across students, and relatively little variability in its curvature.

4.2. Hypothesis 1b: Test-Related Cognitions and Test Performance. Having identified the polynomial function that best represents change over time in test-related cognitions, test performance improvement from the first to the second test was entered as a predictor of the different growth coefficients, namely, the individual intercept and the linear slope on Level 2 (Table 2, Columns 4 and 5).² The quadratic growth parameter was not predicted due to nonsignificant variability across students leaving little variance to be explained by any predictors.³

Based on the manner in which time was coded [60], a positive effect of test improvement on the intercept would indicate that test improvement corresponded to higher overall values of test-related cognitions, whereas a positive effect on the slope would imply that test improvement corresponded to a stronger increase in test-related cognitions (the intercept employed was the day before the next test). However, as these effects could differ if assessed at other points in time prior the second test, only the achievement effects on the highest polynomial parameter (i.e., quadratic growth) are independent from the coding of time [62, 63]. Therefore, a visual inspection of the estimated growth trajectories is necessary to ensure a meaningful interpretation of the fixed and random effects in the HLM models [60].

A significant effect of .19 for test improvement on the intercept was found suggesting that the more students thought about test, especially one day prior to the test date, the more they improved on this test in relation to the preceding test. A significant effect of .02 for test improvement was also found on the linear slope suggesting, in combination with a visual inspection, that increased frequency over time in test-related cognitions, particularly on the final day before the test, contributed to improved test performance. As the magnitude of this effect is contingent upon the coding of time (x axis), it can, therefore, not be interpreted. Figure 3 reflects the growth curve for the averaged lower quartile (students who declined between the two tests) and for the averaged upper quartile (students who improved between the two tests) on the test improvement measure. The figure shows that although both groups appear to demonstrate a similar frequency of test-related thoughts 14 days prior to the test, differences between the two groups on changes in test-related cognitions become more apparent as the test date approaches with the frequency of test-related cognitions increasing at a faster rate for students who improved from one test to the next.

4.3. Hypothesis 2a: Frequency of Metacognitive Strategy Use. Results showed that students' thoughts about the upcoming mathematics test over the previous hour were accompanied by reports of engaging in at least one of the metacognitive strategies 86% of the time. More specifically, time spent thinking about the test was accompanied 75% of the time by students reminding themselves of what they had to learn for

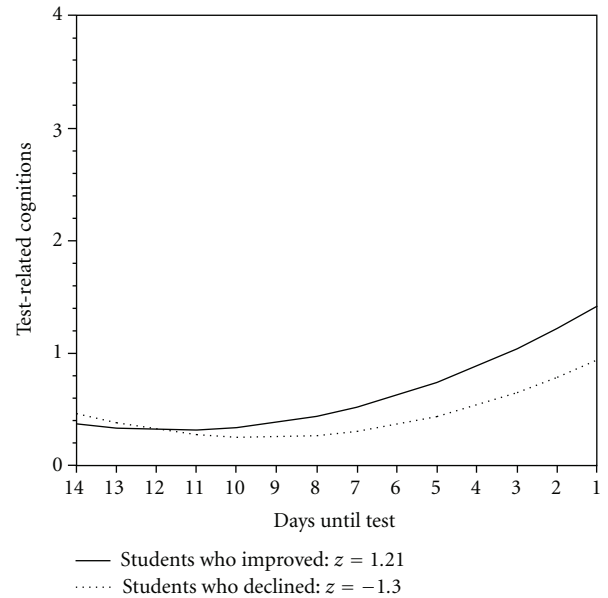


FIGURE 3: Average intensity of test-related cognitions by improvement classification. *Note:* The two lines reflect the growth curves of the parameters for averaged lower and upper quartiles considering the relative improvement between the two tests.

the test (planning), 63% of the time by students monitoring their learning, and 41% of the time by students thinking at least once about whether to change their learning process (evaluation) over the previous hour. The average intensity of metacognitive strategies when test-related thoughts were reported indicate that planning ($M = 1.70$) was not only reported more often across assessment points (i.e., at least once), but also more frequently *over the past hour* than monitoring ($M = 1.33$) or evaluation ($M = 0.77$). In contrast, evaluation was least often reported across assessment points and also showed the lowest intensity over the preceding hour. Table 3 outlines the correlations between the metacognitive strategies and test-related cognitions (only for scores above 0), showing planning and monitoring to more positively correspond with test-related cognitions than evaluation. Moreover, planning and monitoring were highly intercorrelated and showed lower correlations with evaluation.

To further analyze change over time in these metacognitive strategies over the 14 days prior to the mathematics test (Tables 4, 5, and 6), all measurement points, including those during which students did not report thinking about the test, were taken into account. Similar to the curve identified for test-related cognitions, quadratic growth curves were also significant for each metacognitive strategy. Due to the equivalent manner in which time was coded, the growth parameters of these curves are comparable. The quadratic growth curves for planning (Table 4, Figure 4) and monitoring (Table 5, Figure 4) were very similar, each having intercepts near 1, nearly equivalent linear parameters (planning: .18; monitoring: .16), and a small quadratic parameter indicating a wide curvature. In contrast, the

TABLE 3: Correlations between test-related cognitions and metacognitive strategies.

	Planning	Monitoring	Evaluation
Monitoring	.57***	—	
Evaluation	.37***	.45***	—
Test-related cognitions	.59***	.50***	.25***

Note: $N_{\text{level 1}} = 2654$; $N_{\text{level 2}} = 70$. *** $P < .001$.

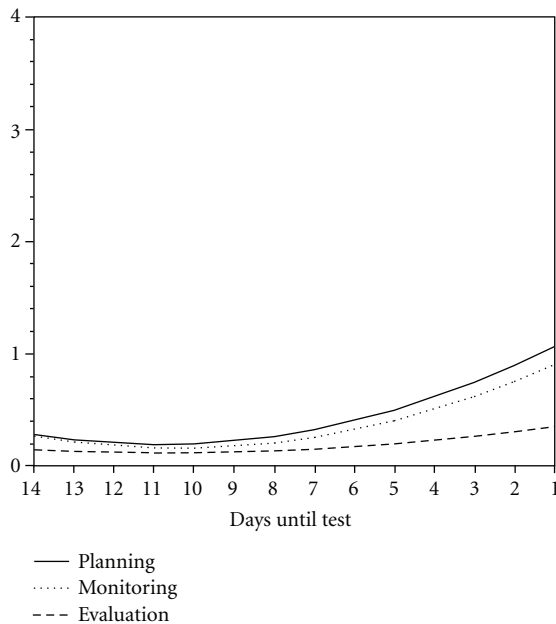


FIGURE 4: Change in frequency of planning, monitoring, and evaluation strategy use.

growth curve for evaluation (Table 6, Figure 4) was relatively even, having a very low intercept (.35) as well as linear slope (.05) assessed at one day before the test, together suggesting that this strategy was rarely used and increased only slightly in the days prior to the test.

4.4. Hypothesis 2b: Metacognition and Test Performance. Consistent with the achievement analysis for test-related cognitions, z -standardized results of test improvement residuals were entered as predictors in the model (Tables 4, 5, and 6; Columns 4 and 5). The test improvement residual was found to be significantly related to the intercept as well as the linear slope for the metacognitive strategy of monitoring. This finding suggests that the more monitoring was reported by students, particularly during the days just prior to the test, the better they performed on that test relative to their previous test performance. Further, this result indicates that increases in monitoring, especially during the days just prior to the test, corresponded with better performance on that test in comparison to their previous test (Figure 5). No such relations were observed for the metacognitive strategies of planning and evaluation.

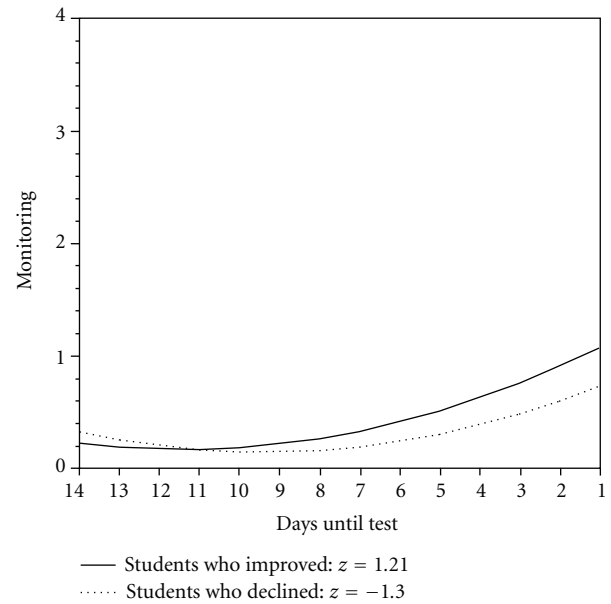


FIGURE 5: Change in frequency of monitoring strategy use by improvement classification. Note: The two lines reflect the growth curves of the parameters for averaged lower and upper quartiles considering the relative improvement between the two tests.

5. Discussion

5.1. Hypotheses 1a and 1b: Test-Related Cognitions

5.1.1. Frequency of Test-Related Cognitions. The present findings suggest that students did indeed think about their upcoming test in mathematics during nearly a quarter of the experience sampling assessments obtained during the 14 days prior to test completion. Further, this result provides evidence in support of evaluating domain-specific test completion, in the present case with respect to mathematics class, as a specific achievement goal that students are acutely aware of and explicitly think about on a regular basis [29]. Regardless of their performance expectations, the observed frequency of test-related thoughts in this study suggests that students perceived their test performance as important which likely contributed to efforts aimed at regulating their learning behavior to this end.

Nonetheless, students were also found to report not thinking as often about the test during their leisure time, indicating that students are also able to disengage from test-related thoughts in situations not related to academic achievement. According to Boekaerts [5], the minimizing of achievement-related cognitions in nonachievement settings is indicative of an adaptive self-regulatory ability to preserve motivational and cognitive resources for future achievement striving. Conversely, this finding also indicates that, even during their leisure time, some students are occupied with thoughts about their test, thus underscoring the importance of this test as well as the relative self-regulatory inability of some students to optimally preserve their motivational and cognitive resources in situations not related to learning.

TABLE 4: Analysis of change in planning strategies prior to test completion.

	Basic model		Linear model		Quadratic model		MG + quadratic model	
	<i>b</i>	SE	<i>b</i>	SE	<i>b</i>	SE	<i>b</i>	SE
Fixed effects								
Intercept (I)	.41**	.04	.78**	.09	1.06**	.13	1.06**	.13
<i>MG</i> × <i>I</i>							.09	.07
Linear Slope (LS)			.06**	.01	.18**	.03	.18**	.03
<i>MG</i> × <i>LS</i>							.01	.01
Quadratic Slope (QS)					.01**	.001	.01***	.002
	Var		Var		Var		Var	
Random effects								
Between: I	.068**		.296**		.595**		.578**	
Between: LS			.002**		.023 ⁺		.023 ⁺	
Between: QS					.000		.000	
Within	.497 ^a		.420 ^a		.386 ^a		.386 ^a	
Model statistics								
Deviance	1400.07		1311.45		1276.15		1274.64	
No. parameters	3		6		10		12	

⁺ $P < .1$; * $P < .05$; ** $P < .01$; ^ano significance test available.

Note: Intercept represents the values one day before the test. State measures across one day were aggregated. $N_{\text{level } 1} = 733$; $N_{\text{level } 2} = 59$; MG = z-standardized residuals on second mathematics test excluding variance accounted for by first test.

TABLE 5: Analysis of change in monitoring strategies prior to test completion.

	Basic model		Linear model		Quadratic model		MG + quadratic model	
	<i>b</i>	SE	<i>b</i>	SE	<i>b</i>	SE	<i>b</i>	SE
Fixed effects								
Intercept (I)	.34**	.04	.64**	.08	.90**	.12	.90**	.12
<i>MG</i> × <i>I</i>							.14*	.07
Linear Slope (LS)			.05**	.01	.16**	.03	.16**	.03
<i>MG</i> × <i>LS</i>							.01*	.001
Quadratic Slope (QS)					.01**	.002	.01**	.001
	Var		Var		Var		Var	
Random effects								
Between: I	.052**		.252**		.524**		.463**	
Between: LS			.002**		.021 ⁺		.019 ⁺	
Between: QS					.000		.000	
Within	.381 ^a		.322 ^a		.293 ^a		.292 ^a	
Model statistics								
Deviance	1232.34		1152.43		1111.16		1107.86	
No. parameters	3		6		10		12	

⁺ $P < .1$; * $P < .05$; ** $P < .01$; ^ano significance test available.

Note: Intercept represents the values one day before the test. State measures across one day were aggregated. $N_{\text{level } 1} = 733$; $N_{\text{level } 2} = 59$; MG = z-standardized residuals on second mathematics test excluding variance accounted for by first test.

and achievement. Further, the finding that students reported thinking about the test during more than half of the time spent in mathematics classes suggests that test preparation may be an explicit topic of classroom instruction and thus externally regulated by the teacher by way of the students' learning environment (see [64]). Thus, depending on the extent to which "teaching to the test" characterizes the class curriculum, the question arises as to the suitability of externally regulated learning environments for fostering and rewarding the self-regulatory competencies explored in this

study [5, 26]. On the other hand, further research on this assumption is warranted given that explicit instruction by the teacher on metacognitive strategy use concerning the upcoming test may also be responsible for students' higher levels of general test-related cognitions during class.

Concerning changes in students test-related thoughts as the test date approached, the development of test-related cognitions was best reflected by a quadratic curve implying that not only did students think more about the test over time, this growth in the frequency of test-related thoughts

TABLE 6: Analysis of change in evaluation strategies prior to test completion.

	Basic model		Linear model		Quadratic model		MG + quadratic model	
	<i>b</i>	SE	<i>b</i>	SE	<i>b</i>	SE	<i>b</i>	SE
Fixed effects								
Intercept (I)	.17**	.03	.27**	.05	.35**	.08	.35**	.08
<i>MG</i> × <i>I</i>							.07	.04
Linear Slope (LS)			.02**	.005	.05*	.02	.05*	.02
<i>MG</i> × <i>LS</i>								
Quadratic Slope (QS)					.003*	.001	.003*	.004
	Var		Var		Var		Var	
Random effects								
Between: I	.029**		.081**		.200**		.188**	
Between: LS			.000*		.010*		.009*	
Between: QS					.000		.000	
Within	.142 ^a		.133 ^a		.125 ^a		.353 ^a	
Model statistics								
Deviance	628.66		604.49		582.21		579.14	
No. parameters	3		6		10		12	

⁺ $P < .1$; * $P < .05$; ** $P < .01$; ^ano significance test available.

Note: Intercept represents the values one day before the test. State measures across one day were aggregated. $N_{\text{level } 1} = 733$; $N_{\text{level } 2} = 59$; *MG* = *z*-standardized residuals on second mathematics test excluding variance accounted for by first test.

also further increased as the test date neared. This pattern of change in test-related thinking is perhaps not surprising as the incremental cognitive investment toward achievement goals allowing for goals with more immediate deadlines to be most fully pursued (i.e., tests in other classes) is assumed to be a critical component of self-regulated learning [5]. Notably greater variance for the intercepts as compared to the linear and especially the quadratic slopes was also observed, indicating that despite considerable variability between students in terms of their frequency of test-related thoughts, the shape of the growth curve was nonetheless very similar across students. Taken together, these findings provide empirical support for our first hypothesis in demonstrating greater test-related thinking in achievement or learning situations, and further, in showing that increases in the frequency of test-related thoughts over time began to rise as the test date approached. However, given the possibility that the present data collection method requiring three to six assessments per day may have also inflated test-related cognitions, further ESM studies in which less frequent assessments are employed are needed to more fully address this potential alternative explanation for the prevalence of test-related thoughts.

5.1.2. Test-Related Cognitions and Test Performance. The intercept as well as slope for the frequency of test-related thinking over time were found to positively correspond with improvements in test performance. This finding suggests that despite the highly generalized nature of the present test-related cognition measure, it nonetheless appears to be consistently associated with more specific cognitions that more directly contribute to improvements in test performance, namely, metacognitive learning strategies (cf., [50]), and as

such, provides some empirical support for the use of global test-related cognition items in ESM research. However, it is important to also note that this global assessment of test-related thoughts may also have reflected negative emotions concerning the test (e.g., anxiety) or related coping strategies aimed at minimizing preparation for the test. Nevertheless, that test-related cognitions were still found to predict better subsequent test performance overall, despite the possibility of this measure indicating maladaptive emotions or cognitions for some students, provides further support for our assumption that test-directed cognitions were associated with effective test-related metacognitive strategies.

5.2. Hypotheses 2a and 2b: Metacognitive Strategies

5.2.1. Frequency of Metacognitive Strategy Use. To more explicitly address the above assumption that test-related cognitions imply a potential ensemble of learning strategies, analyses further revealed that at least one of the three metacognitive strategies assessed in this study was typically reported if test-related thinking was indicated. More specifically, the strategy of planning, defined as thinking about what to learn for the test, was most frequently reported, followed by the strategy of monitoring, operationalized as reviewing existing knowledge as it applies to the upcoming test. In contrast, evaluation was applied far less often and also reported to occur less frequently over the course of the previous hour than were the other two strategies. Correlations further showed planning and monitoring to have a strong positive relationship, with evaluation also showing positive albeit notably weaker relations with these two strategies.

These findings are in line with previous findings showing these metacognitive strategies to be highly interdependent [19, 20]. They further show that although students regularly engage in the planning and monitoring of their learning behavior when preparing for an upcoming test, they may fail to regularly adapt their learning behavior based on evaluated progress toward their learning goal. Concerning potential explanations for this pattern of results, it is possible that students may opt to forego self-regulatory strategies aimed at improving their learning behavior and rely instead on teachers to correct and externally regulate their learning behavior [64]. Alternatively, it is also possible that many students may already have well-established ways of learning that they do not wish to change, that students do not find the evaluation methods or feedback provided by instructors during the learning process to be useful or necessary, or that they in fact lack the ability to effectively use the metacognitive strategy of evaluation regardless of the type of evaluative feedback available given that this strategy is significantly more complex than either planning or monitoring.

The observed changes in the use of these three metacognitive strategies parallel those observed for test-related cognitions in that all three growth trajectories were best reflected by a quadratic growth curve. However, whereas the linear and quadratic slope parameters for the strategies of planning and monitoring were very similar to each other, the development of evaluation strategies over time was far more flat and constant in nature. This finding is further consistent with previous findings showing these metacognitive strategies to not be related to each other in a hierarchical or sequential manner, but rather to occur simultaneously toward the completion of an achievement goal [19].

5.2.2. Metacognitive Strategy Use and Test Performance. In contrast, the test improvement measure was found to positively correspond with the intercept as well as linear slope for the strategy of monitoring. This finding suggests that the more monitoring was evidenced by students especially during the final days before the test, and the more monitoring increased during the final days before the test, the better they performed on this test relative to their previous test in mathematics class. Similar relations were not found for the strategies of planning and evaluation.

These significant results underscore the importance of monitoring as a critical metacognitive strategy that, according to Winne and Hadwin [7], represents an omnipresent metacognitive process responsible for facilitating various aspects of self-regulated learning. More specifically, the present findings suggest that student engagement in the monitoring process is significantly related to subsequent academic achievement and that as a test date approaches, increased emphasis on monitoring one's knowledge acquisition is essential to ensuring that the learning process remains efficient and effective. With respect to the lack of significant relations between test improvement and planning, it is possible that planning as a more distal strategy, initiated at the outset of the learning process, may show lower relations with subsequent performance than more proximal strategies such as monitoring that entail the continued engagement in

planned activities up to test completion. As noted above, it is also possible that no relations between evaluation and achievement gains were observed due to the externally regulated nature of the classroom setting in which the learning progress is also monitored and corrected by the teacher in lieu of optimal test performance, but the lack of a relationship between evaluation and improvement might also lie in the lack of the use of evaluation strategies.

The present study findings thus demonstrate a significant relationship between students' engagement in metacognitive strategy use with respect to an upcoming test in mathematics and their actual performance on this test. Further, these results highlight the importance of evaluating not only the overall frequency with which such strategies are employed, but also change over time of the use of metacognitive strategies during the learning process prior to test completion. Findings revealed the intercept as well as growth in the frequency of monitoring the learning process to be significantly related with improvements in test grades relative to prior test performance. Taken together, the results of the present study provide empirical support for the temporal relationship between metacognitive strategy use and achievement over and above the findings of correlational questionnaire studies (cf. [8]). Moreover, the present findings also highlight the utility of experience sampling methods to better evaluate students' use of metacognitive strategies as not only global dispositions but situated learning behaviors that directly correspond to actual performance gains in a real-life classroom setting.

6. Implications for Future Research and Practice

As suggested by the present findings showing the experience sampling method to be a worthwhile instrument for evaluating self-regulated learning processes, future studies in which such elaborated assessment methods are employed are recommended to further explore students' learning behaviors in real-life achievement settings. Similarly, greater methodological research in which better and additional measures of validity and reliability for data obtained from experience sampling methods is required. It is anticipated that future research in which additional learning strategies are explored using such data collection and statistical methods can provide considerable insight into students' use of such strategies in actual learning situations. For example, these studies could evaluate metacognitive as well as more specific cognitive learning strategies used by students when presented with learning material [5] that are known to predict academic achievement (e.g., elaborative learning; [65]) as well as motivation- and emotion-regulation strategies used by students in response to academic stress, failure, and negative emotions (e.g., anxiety, boredom; [51, 66, 67]). Coding of open-ended ESM question formats, as well as exploring interactions between specific self-regulated learning strategies, may also help to further elucidate the types of strategies being employed in real-life learning settings and further refine existing models of self-regulated learning.

Future ESM research in which objective evaluations of specific elements of the students' learning environment are also assessed should serve significantly to complete our understanding of how and why metacognitive and learning strategies are employed by students in specific learning situations. By assessing not only how students regulate their learning process but also the extent to which classroom teachers as well as goal structures promote mastery of learning material, as opposed to performance on achievement tests (see [68]), the degree to which specific learning and metacognitive strategies are endorsed can be more thoroughly analyzed in relation to the actual opportunities afforded to students to learn these strategies and have them pay off on performance evaluations. In other words, our findings suggest that by extending the present situated method of analysis to also include classroom and teacher variables, a more comprehensive understanding of students' learning behavior can be achieved by researchers and educators alike. Efforts to combine such self-report measures with observational methods such as video-based assessment should also help to provide further insight into students' learning behaviors.

Similarly, it is anticipated that such findings from studies in which situational factors are more fully explored should serve to inform efforts to improve classroom environments as well as develop effective intervention programs for struggling students. Upon identifying aspects of the classroom environment that facilitate or prevent the use and effectiveness of self-regulated learning strategies, such findings can be readily incorporated into teacher education programs. As such, it should be possible to inform preservice and practicing teachers of the importance of creating optimal classroom settings for fostering students' self-regulatory competences and increase their awareness of the necessity and utility of students' self-regulatory efforts (see [13]).

Finally, further research on the moderating effect of situational factors on the use and achievement benefits of metacognitive and cognitive learning strategies should also serve to enhance the effectiveness of related intervention programs. For example, the lack of change in the use of evaluation as a metacognitive strategy needs further research with respect to the impact of classroom features on students' use of this strategy such as test format, lesson structure, or explicit guidance concerning the effective use of self-regulated learning strategies by the instructor. By better delineating the effects of classroom and instructional dynamics, as compared to dispositional factors (e.g., individual differences in motivational strategies) on students' metacognitive strategy use, we can design better intervention programs that target the critical sources of maladaptive learning approaches (e.g., having a focus on teacher training versus improving student motivation). In summary, the present study illustrates that students do indeed have the ability to effectively use metacognitive strategies, particularly the strategy of monitoring, and in so doing, can positive impact their subsequent achievement. Nevertheless, this research also suggests further avenues of research into the intentionality of changes in students' self-regulatory strategy use over time, possible changes in related

self-regulatory strategies over time (e.g., to regulate one's motivation, emotions), exploring and improving the use and effectiveness of specific learning strategies (e.g., evaluation, elaboration), as well as the potential moderating effects of instructional methods and classroom features.

Disclosure

This project is based on data from the doctoral dissertation of the first author.

Endnotes

1. It should be noted that the present sample affords a three-level data structure consisting of measurement points within individuals within classes. The intraclass correlations (ρ) between the individual level and the class level for the variables were $\rho_{\text{test-related cognitions}} = .19$, $\rho_{\text{planning}} = .23$, $\rho_{\text{monitoring}} = .02$, $\rho_{\text{evaluation}} = .02$. As a result of the small cluster group ($M_c \leq 4$), the design effect ($\text{DEFF} = 1 + \rho(M_c - 1)$) of individuals within classes was below 2 for all variables. Thus, the third class level was not assessed in subsequent models (see Muthén, L.; 1999, October 29. Intraclass correlations [Msg 2]. Message posted to <http://www.statmodel.com/discussion/messages/12/18.html?1253640879>). On average, 10.5 days out of a maximum of 14 days, were assessed per student. Whole missing days might be due a student forgetting the device at home or forgetting to recharge the device on time or similar reasons. On the days were data was assessed, on average 3.6 single measurement points were gained out of an expected possible (but not necessarily given) maximum of in average 5 single measurement points. No systematic in the pattern of missings could be found, thus we assume, that all missings are at random.
2. Achievement on the second test in mathematics is largely predicted by previous test performance, and this relationship may, in turn, influence the regression weights of the growth parameters. As such, what we assessed were the residuals of performance on the second test obtained after variance explained by the first test was removed. This score is thus referred to as test improvement and describes the relative improvement over time from the first test to the second, with students who performed better than expected considering their previous test score obtaining a score above zero, and those who performed worse than expected considering their previous test score obtaining a negative score. Integrating performance on the first test in the model as a predictor follows from the assumption that the prior achievement may influence learning behavior. Although including test improvement as a predictor of prior learning behavior is counterintuitive, it was nonetheless included as such due to the significance of the relationship between the individual growth parameters and test improvement being accurately reflected

by the regression results despite an inverse direction of causality being assumed.

3. Applying the HLM notation, the resulting equations were as follows:

$$\begin{aligned} Y &= \pi_0 + \pi_1 x + \pi_2 x^2 + e, \\ \pi_0 &= \beta_{00} + \beta_{01} z + r_0, \\ \pi_1 &= \beta_{10} + \beta_{11} z + r_1, \\ \pi_2 &= \beta_{20} + r_2. \end{aligned} \quad (1)$$

References

- [1] B. Schmitz, M. Schmidt, M. Landmann, and C. Spiel, "New developments in the field of self-regulated learning," *Journal of Psychology*, vol. 215, no. 3, pp. 153–156, 2007.
- [2] P. H. Winne, "A perspective on state-of-the-art research on self-regulated learning," *Instructional Science*, vol. 33, no. 5-6, pp. 559–565, 2005.
- [3] J. H. Flavell, "Metacognition and cognitive monitoring: a new area of cognitive-developmental inquiry," *American Psychologist*, vol. 34, no. 10, pp. 906–911, 1979.
- [4] M. V. J. Veenman, B. H. A. M. van Hout-Wolters, and P. Afflerbach, "Metacognition and learning: conceptual and methodological considerations," *Metacognition and Learning*, vol. 1, no. 1, pp. 3–14, 2006.
- [5] M. Boekaerts, "Self-regulated learning: where we are today," *International Journal of Educational Research*, vol. 31, no. 6, pp. 445–457, 1999.
- [6] S. G. Paris and A. H. Paris, "Classroom applications of research on self-regulated learning," *Educational Psychologist*, vol. 36, no. 2, pp. 89–101, 2001.
- [7] P. H. Winne and A. F. Hadwin, "Studying as self-regulated learning," in *Metacognition in Educational Theory and Practice*, D. J. Hacker, J. Dunlosky, and A. C. Graesser, Eds., pp. 279–306, Erlbaum, Hillsdale, NJ, USA, 1998.
- [8] M. V. J. Veenman, "The assessment of metacognitive skills: what can be learned from multi-method designs?" in *Lernstrategien und Metacognition: Implikationen für Forschung und Praxis*, C. Artelt and B. Moschner, Eds., pp. 77–99, Waxmann, Münster, Germany, 2005.
- [9] P. H. Winne and N. E. Perry, "Measuring self-regulated learning," in *Handbook of Self-Regulation*, M. Boekaerts, P. R. Pintrich, and M. Zeidner, Eds., pp. 531–566, Elsevier, Burlington, Mass, USA, 2000.
- [10] M. Csikszentmihalyi and R. Larson, "Validity and reliability of the experience-sampling method," *Journal of Nervous and Mental Disease*, vol. 175, no. 9, pp. 526–536, 1987.
- [11] J. M. Hektner, J. A. Schmidt, and M. Csikszentmihalyi, *Experience Sampling Method: Measuring the Quality of Everyday Life*, Sage, Thousand Oaks, Calif, USA, 2007.
- [12] S. Rezvan, S. A. Ahmadi, and M. R. Abedi, "The effects of metacognitive training on the academic achievement and happiness of Esfahan University conditional students," *Counselling Psychology Quarterly*, vol. 19, no. 4, pp. 415–428, 2006.
- [13] M. Boekaerts and M. Niemivirta, "Self-regulated learning: finding a balance between learning goals and ego-protective goals," in *Handbook of Self-Regulation*, M. Boekaerts, P. R. Pintrich, and M. Zeidner, Eds., pp. 417–451, Elsevier, Burlington, Mass, USA, 2005.
- [14] P. R. Pintrich, "The role of goal-orientation in self-regulated learning," in *Handbook of Self-Regulation*, M. Boekaerts, P. R. Pintrich, and M. Zeidner, Eds., pp. 451–502, Elsevier, Burlington, Mass, USA, 2005.
- [15] B. J. Zimmerman, "Attaining self-regulation: a social cognitive perspective," in *Handbook of Self-Regulation*, M. Boekaerts, P. R. Pintrich, and M. Zeidner, Eds., pp. 13–41, Elsevier, Burlington, Mass, USA, 2005.
- [16] G. Schraw, "Promoting general metacognitive awareness," *Instructional Science*, vol. 26, no. 1, pp. 113–125, 1998, <http://andrewws.blogs.com/usu/files/promoting-general-metacognitive-awareness-schraw.pdf>.
- [17] N. Spoerer and J. C. Brunstein, "Erfassung selbstregulierten Lernens mit Selbstberichtsverfahren: Ein Überblick zum Stand der Forschung," *Zeitschrift für Pädagogische Psychologie*, vol. 20, no. 2, pp. 147–160, 2006.
- [18] P. H. Winne, "Self-regulation is ubiquitous but its forms vary with knowledge," *Educational Psychologist*, vol. 30, no. 4, pp. 223–228, 1995.
- [19] R. Azevedo, "Theoretical, conceptual, methodological, and instructional issues in research on metacognition and self-regulated learning: a discussion," *Metacognition and Learning*, vol. 4, no. 1, pp. 87–95, 2009.
- [20] P. R. Pintrich, D. A. F. Smith, T. Garcia, and W. J. McKeachie, "Reliability and predictive validity of the motivated strategies for learning questionnaire (MSLQ)," *Educational & Psychological Measurement*, vol. 53, no. 3, pp. 801–813, 1993.
- [21] B. Schmitz, "Advantages of studying processes in educational research," *Learning and Instruction*, vol. 16, no. 5, pp. 433–449, 2006.
- [22] A. Koriart and R. Nussinson, "Attributing study effort to data-driven and goal-driven effects: implications for metacognitive judgments," *Journal of Experimental Psychology*, vol. 35, no. 5, pp. 1338–1343, 2009.
- [23] K. Lockl and W. Schneider, "Developmental trends in children's feeling-of-knowing judgements," *International Journal of Behavioral Development*, vol. 26, no. 4, pp. 327–333, 2002.
- [24] T. O. Nelson, L. Narens, and J. Dunlosky, "A revised methodology for research on metamemory: pre-judgment recall and monitoring (PRAM)," *Psychological Methods*, vol. 9, no. 1, pp. 53–69, 2004.
- [25] J. A. Greene and R. Azevedo, "A macro-level analysis of SRL processes and their relations to the acquisition of a sophisticated mental model of a complex system," *Contemporary Educational Psychology*, vol. 34, no. 1, pp. 18–29, 2009.
- [26] E. de Corte, L. Verschaffel, and C. Masui, "The CLIA-model: a framework for designing powerful learning environments for thinking and problem solving," *European Journal of Psychology of Education*, vol. 19, no. 4, pp. 365–384, 2004.
- [27] B. J. Zimmerman and A. Bandura, "Impact of self-regulatory influences on writing course attainment," *American Educational Research Journal*, vol. 31, no. 4, pp. 845–862, 1994, <http://www.des.emory.edu/mfp/Bandura1994AERJ.pdf>.
- [28] M. C. Wang, G. D. Haertel, and H. J. Walberg, "What influences learning? A content analysis of review literature," *Journal of Educational Research*, vol. 84, no. 1, pp. 30–43, 1990.
- [29] M. Boekaerts and L. Corno, "Self-regulation in the classroom: a perspective on assessment and intervention," *Applied Psychology*, vol. 54, no. 2, pp. 199–231, 2005.
- [30] R. A. Sperling, B. C. Howard, R. Staley, and N. DuBois, "Metacognition and self-regulated learning constructs," *Educational Research and Evaluation*, vol. 10, no. 2, pp. 117–139, 2004.
- [31] B. J. Zimmerman and M. Martinez-Pons, "Development of a structured interview for assessing student use of self-regulated

- learning strategies," *American Educational Research Journal*, vol. 23, no. 4, pp. 614–628, 1986.
- [32] B. J. Zimmerman and M. Martinez-Pons, "Construct validation of a strategy model of student self-regulated learning," *Journal of Educational Psychology*, vol. 80, no. 3, pp. 284–290, 1988.
- [33] C. Dignath, G. Buettner, and H. P. Langfeldt, "How can primary school students learn self-regulated learning strategies most effectively? A meta-analysis on self-regulation training programmes," *Educational Research Review*, vol. 3, no. 2, pp. 101–129, 2008.
- [34] C. Dignath and G. Büttner, "Components of fostering self-regulated learning among students. A meta-analysis on intervention studies at primary and secondary school level," *Metacognition and Learning*, vol. 3, no. 3, pp. 231–264, 2008.
- [35] J. Hattie, J. Biggs, and N. Purdie, "Effects of learning skills interventions on student learning: a meta-analysis," *Review of Educational Research*, vol. 66, no. 2, pp. 99–136, 1996.
- [36] C. Artelt, *Strategisches Lernen*, Waxmann, Münster, Germany, 2000.
- [37] F. J. Sinkavich, "Metamemory, attributional style, and study strategies: predicting classroom performance in graduate students," *Journal of Instructional Psychology*, vol. 21, no. 2, pp. 172–182, 1994.
- [38] S. Coutinho, "Self-efficacy, metacognition, and performance," *North American Journal of Psychology*, vol. 10, no. 1, pp. 165–172, 2008.
- [39] M. V. J. Veenman and B. H. A. M. van Hout-Wolters, "Het meten van metacognitieve vaardigheden," in *Onderwijsonderzoek in Nederland en Vlaanderen. Proceedings van de 29e Onderwijs Research Dagen 2002 te Antwerpen*, F. Daems, R. Rymenans, and G. Rogiest, Eds., pp. 102–103, Universiteit Antwerpen, 2002.
- [40] J. D. Karpicke, "Metacognitive control and strategy selection: deciding to practice retrieval during learning," *Journal of Experimental Psychology*, vol. 138, no. 4, pp. 469–486, 2009.
- [41] M. Nueckles, R. Schwonke, K. Berthold, and A. Renkl, "The use of public learning diaries in blended learning," *Journal of Educational Media*, vol. 29, no. 1, pp. 49–66, 2004.
- [42] D. Jamieson-Noel and P. H. Winne, "Comparing self-reports to traces of studying behavior as representations of students' studying and achievement," *Zeitschrift für Pädagogische Psychologie*, vol. 17, no. 3–4, pp. 159–171, 2003.
- [43] C. E. Weinstein, "Assessment and training of student learning strategies," in *Learning Strategies and Learning Styles*, R. R. Schmeck, Ed., pp. 291–316, Plenum Press, New York, NY, USA, 1988.
- [44] A. Minnaert and P. J. Janssen, "Bias in the assessment of regulation activities in studying at the level of higher education," *European Journal of Psychological Assessment*, vol. 13, no. 2, pp. 99–108, 1997.
- [45] M. V. J. Veenman, F. J. Prins, and J. Verheij, "Learning styles: self-reports versus thinking-aloud measures," *British Journal of Educational Psychology*, vol. 73, no. 3, pp. 357–372, 2003.
- [46] S. Hübner, M. Nückles, and A. Renkl, "Writing learning journals: instructional support to overcome learning-strategy deficits," *Learning and Instruction*, vol. 20, no. 1, pp. 18–29, 2010.
- [47] M. Nückles, S. Hübner, and A. Renkl, "Enhancing self-regulated learning by writing learning protocols," *Learning and Instruction*, vol. 19, no. 3, pp. 259–271, 2009.
- [48] M. Puustinen and L. Pulkkinen, "Models of self-regulated learning: a review," *Journal of Educational Research*, vol. 45, no. 3, pp. 269–286, 2001.
- [49] J. E. Schwartz, J. Neale, C. Marco, S. S. Shiffman, and A. A. Stone, "Does trait coping exist? A momentary assessment approach to the evaluation of traits," *Journal of Personality and Social Psychology*, vol. 77, no. 2, pp. 360–369, 1999.
- [50] J. B. Vancouver and D. V. Day, "Industrial and organisation research on self-regulation: from constructs to applications," *Applied Psychology*, vol. 54, no. 2, pp. 155–185, 2005.
- [51] T. Goetz, A. C. Frenzel, R. Pekrun, and N. C. Hall, "The domain specificity of academic emotional experiences," *Journal of Experimental Education*, vol. 75, no. 1, pp. 5–29, 2006.
- [52] T. Goetz, A. C. Frenzel, R. Pekrun, N. C. Hall, and O. Lüdtke, "Between- and within-domain relations of students' academic emotions," *Journal of Educational Psychology*, vol. 99, no. 4, pp. 715–733, 2007.
- [53] H. M. Weiss, D. J. Beal, S. L. Lucy, and S. M. MacDermid, "Constructing EMA studies with PMAT: the purdue momentary assessment tool user's manual," 2004.
- [54] T. Goetz, A. C. Frenzel, H. Stoeger, and N. C. Hall, "Antecedents of everyday positive emotions: an experience sampling analysis," *Motivation and Emotion*, vol. 34, no. 1, pp. 49–62, 2010.
- [55] U. E. Nett, T. Goetz, and N. C. Hall, "Coping with boredom in school: an experience sampling perspective," *Contemporary Educational Psychology*, vol. 36, no. 1, pp. 49–59, 2011.
- [56] S. W. Raudenbush, A. S. Bryk, and R. Congdon, "HLM 6.04. hierarchical linear and nonlinear modeling," Lincolnwood: Scientific Software International, 2007.
- [57] J. L. Schafer and J. W. Graham, "Missing data: our view of the state of the art," *Psychological Methods*, vol. 7, no. 2, pp. 147–177, 2002.
- [58] P. D. Allison, *Missing Data*, Sage, Thousand Oaks, Calif, USA, 2001.
- [59] A. C. Frenzel, T. Goetz, R. Pekrun, and H. M. G. Watt, "Development of mathematics interest in adolescence: influences of gender, family, and school context," *Journal of Research on Adolescence*, vol. 20, no. 2, pp. 507–537, 2010.
- [60] J. C. Biesanz, N. Deeb-Sossa, A. A. Papadakis, K. A. Bollen, and P. J. Curran, "The role of coding time in estimating and interpreting growth curve models," *Psychological Methods*, vol. 9, no. 1, pp. 30–52, 2004.
- [61] S. W. Raudenbush and A. S. Bryk, *Hierarchical Linear Models. Applications and Data Analysis Methods*, Sage, Thousand Oaks, Calif, USA, 2nd edition, 2002.
- [62] L. S. Aiken and S. G. West, *Multiple Regression: Testing and Interpreting Interactions*, Sage, Newbury Park, Calif, USA, 1991.
- [63] J. Cohen, "Partialled products are interactions; partialled powers are curve components," *Psychological Bulletin*, vol. 85, no. 4, pp. 858–866, 1978.
- [64] U. Schiefele and R. Pekrun, "Psychologische modelle des fremdgesteuerten und selbstgesteuerten Lernens," in *Enzyklopädie der Psychologie*, F. E. Weinert, Ed., vol. 2, pp. 249–278, Hogrefe, Göttingen, Germany, 1996.
- [65] N. C. Hall, R. P. Perry, T. Goetz, J. C. Ruthig, R. H. Stupnisky, and N. E. Newall, "Attributional retraining and elaborative learning: improving academic development through writing-based interventions," *Learning and Individual Differences*, vol. 17, no. 3, pp. 280–290, 2007.
- [66] N. C. Hall, "Self-regulation of primary and secondary control in achievement settings: a process model," *Journal of Social and Clinical Psychology*, vol. 27, no. 10, pp. 1126–1164, 2008.
- [67] R. Pekrun, T. Goetz, W. Titz, and R. P. Perry, "Academic emotions in students' self-regulated learning and achievement:

a program of qualitative and quantitative research,” *Educational Psychologist*, vol. 37, no. 2, pp. 91–105, 2002.

- [68] E. M. Anderman and C. Wolters, “Goals, values, and affect,” in *Handbook of Educational Psychology*, P. Alexander and P. H. Winne, Eds., pp. 369–390, Lawrence Erlbaum Associates, Mahway, NJ, USA, 2nd edition, 2006.

